Habilitation à diriger des recherches

Université Paris-Sud 11

Spécialité : Physique

Présentée par

Alexandre Renaux

Contributions aux bornes inférieures de l'erreur quadratique moyenne en traitement du signal

Soutenue le 7 décembre 2011 devant le jury composé de :

Mr.	Chevalier Pascal	CNAM Paris	Président
Mr.	Abed-Meraïm Karim	Télécom ParisTech	Rapporteur
Mr.	Besson Olivier	ISAE Toulouse	Rapporteur
Mr.	Ferrari André	Université de Nice	Rapporteur
Mr.	Berthoumieu Yannick	Université de Bordeaux	Examinateur
Mr.	Delmas Jean-Pierre	Telecom Sud Paris	Examinateur
Mr.	Marcos Sylvie	CNRS L2S	Examinateur

Laboratoire des signaux et systèmes
Supelec, 3 rue Joliot Curie, 91, Gif-sur-Yvette, France

On sait tous que les scientifiques gâchent leur vie.
Homer J. Simpson. Le safari des Simpson. Saison 12. Episode 17.

Le temps ne respecte pas ce qui se fait sans lui.
Cantillon.

Table des matières

Table des figures

Liste des tableaux

Chapitre 1

Curriculum Vitæ

1.1 Présentation générale

Alexandre Renaux 32 ans, né le 17 décembre 1978 à Nancy (54)

Maître de conférences classe normale (échelon 4)
Université Paris-Sud 11 (département de physique)
Section 61 : Génie informatique, automatique et traitement du signal

Laboratoire des Signaux et Systèmes (Université Paris-Sud 11 / CNRS UMR-8506 / Supelec)
Supelec, 3, rue Joliot Curie 91192 Gif-sur-Yvette cedex
Téléphone : +33 1 69 85 17 58
Adresse électronique : `renaux@lss.supelec.fr`
Page web : `http://www.lss.supelec.fr/perso/alexandre.renaux/`

1.2 Parcours académique

– *2003-2006* **Thèse de doctorat** en Traitement du Signal, École Normale Supérieure de Cachan, laboratoire SATIE (Systèmes et Applications des Technologies de l'Information et de l'Énergie) (ENS Cachan / CNRS UMR-8029). Soutenue le 7 juillet 2006.

Contribution à l'étude des performances d'estimation en traitement statistique du signal.

Composition du jury :

- Bernard Fleury, Professeur des Universités, Aalborg University, Danemark (Président)
- Jean-Yves Tourneret, Professeur des Universités, Institut National Polytechnique de Toulouse (Rapporteur)
- Luc Vandendorpe, Professeur des Universités, Université Catholique de Louvain, Belgique (Rapporteur)
- Jean-Pierre Delmas, Professeur des Universités, Institut National des Télécommunications Evry (Examinateur)
- Eric Boyer, Docteur, École Normale Supérieure de Cachan (Invité)
- Philippe Forster, Professeur des Universités, Université Paris 10 (Co-directeur de thèse)

  - Pascal Larzabal, Professeur des Universités, Université Paris-Sud 11/École Normale Supérieure de Cachan (Directeur de thèse)

- *2002-2003* Diplôme d'Etudes Approfondies (**DEA**) en Automatique et Traitement du Signal, Université Paris-Sud 11/Supelec/École Normale Supérieure de Cachan. Mention Bien.
- *2001-2002* **Agrégation** de Génie Électrique. Rang : $7^{ème}$.
- *2000-2001* **Maîtrise** EEA, Université Paris-Sud 11/École Normale Supérieure de Cachan. Mention Assez Bien.
- *1999-2000* **Licence** Ingénierie Électrique, Université Paris-Sud 11/École Normale Supérieure de Cachan. Mention Assez Bien.
- *1996-1999* Classe Préparatoire aux Grandes Écoles (**CPGE**). Admission à l'École Normale Supérieure de Cachan.
- *1996* **Baccalauréat** STI. Mention Très Bien.

1.3 Activités professionnelles

- *Depuis septembre 2007* **Maître de conférences** classe normale ($4^{ème}$ échelon) à l'Université Paris-Sud 11 (département de physique) / Laboratoire des Signaux et Systèmes (Université Paris-Sud 11 / CNRS UMR-8506 / Supelec).
- *2006-2007* **Post-doctoral research associate** à Washington University in St. Louis, USA / Department of Electrical and Systems Engineering.
- *2003-2006* **Allocataire moniteur** Doctorant au laboratoire SATIE (CNRS / École Normale Supérieure de Cachan) et Moniteur à l'IUT de Cachan.

1.4 Publications

La totalité des articles publiés ainsi que quelques présentations et mes polycopiés de cours sont disponibles sur ma page web : `http://www.lss.supelec.fr/perso/alexandre.renaux/`

Revues :

- 15 articles de revue internationale avec comité de lecture publiés ou acceptés (depuis décembre 2006).
- 5 articles de revue internationale avec comité de lecture soumis.

Conférences :

- 19 articles de conférence internationale avec comité de lecture et actes publiés.
- 9 articles de conférence nationale (GRETSI) avec comité de lecture et actes publiés.

Les articles [CI7] et [CI12] concernent des conférences où j'ai personnellement été invité.

Remarque : les articles issus de mes travaux de thèses sont [J1], [J2], [J3], [J4], [CI1], [CI2], [CI3], [CI4], [CN1], et [CN2]. Les articles issus de mes travaux durant mon année de post-doctorat sont [CI5], [CI6], [CI7], et [CN3]. Les articles issus de mes travaux depuis mon intégration en temps que maître de conférences sont [J5], [J6], [J7], [J8], [J9], [J10], [J11], [J12], [J13], [J14], [J15], [CI8], [CI9], [CI10], [CI11], [CI12], [CI13], [CI14], [CI15], [CI16], [CI17], [CI18], [CI19], [CN4], [CN5], [CN6], [CN7], [CN8] et [CN9] (ainsi que [J16], [J17], [J18], [J19], [J20] soumis).

Les articles ou communications marqués d'une étoile sont donnés en annexe.

1.4.1 Articles réimprimés dans des ouvrages

[J4] et [CI4] ont été sélectionnés pour paraître dans l'ouvrage : *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering and Tracking* écrit par Harry L. Van Trees et Kristine L. Bell parut en 2007. Cet ouvrage édité chez Wiley regroupe 80 articles sur le domaine des bornes minimales ainsi qu'une introduction d'environ 100 pages.

1.4.2 Revues internationales avec comité de lecture

[J1] **A. Renaux**, P. Forster, E. Chaumette et P. Larzabal, "On the high SNR conditional maximum likelihood estimator full statistical characterization", *IEEE* Transactions on Signal Processing, Volume : 54, Issue : 12, Dec. 2006, pp. 4840-4843.

[J2] **A. Renaux**, P. Forster, E. Boyer et P. Larzabal, "Unconditional maximum likelihood performance at finite number of samples and high signal to noise ratio", *IEEE* Transactions on Signal Processing, Volume : 55, Issue : 5, Part 2, May 2007, pp. 2358-2364.

[J3] **A. Renaux**, L. Najjar-Atallah, P. Larzabal et P. Forster, "A useful form of the Abel bound and its application to estimator threshold prediction", *IEEE* Transactions on Signal Processing, Volume : 55, Issue : 5, Part 2, May 2007, pp. 2365-2369.

[J4] **A. Renaux**, "Weiss-Weinstein bound for data aided carrier estimation", *IEEE* Signal Processing Letters, Volume : 14, Issue : 4, Apr. 2007, pp. 283-286.

[J5*] C. Herzet, V. Ramon, **A. Renaux**, et L. Vandendorpe, "Characterization of the EM algorithm mean speed of convergence based on Cramér-Rao bounds", *IEEE* Transactions on Signal Processing, Volume : 56, Issue : 6, Jun. 2008, pp. 2218-2228.

[J6*] S. Bay, B. Geller, **A. Renaux**, J-P Barbot et J-M Brossier, "On the Hybrid Cramér-Rao bound and its application to dynamical phase estimation", *IEEE* Signal Processing Letters, Volume : 15, 2008, pp. 453-456.

[J7*] **A. Renaux**, P. Forster, P. Larzabal, C. Richmond, et A. Nehorai, "A fresh look at the Bayesian bounds of the Weiss-Weinstein family", *IEEE* Transactions on Signal Processing, Volume : 56, Issue : 11, Nov. 2008, pp. 5334-5352.

[J8*] F. Pascal et **A. Renaux**, "Statistical analysis of the covariance matrix MLE in K-distributed clutter", Elsevier Signal Processing, Volume : 90, Issue : 4, Apr. 2010, pp. 1165-1175.

[J9] M. N. El Korso, R. Boyer, **A. Renaux** et S. Marcos, "Conditional and unconditional Cramér-Rao bounds for near-field source localization", *IEEE* Transactions on Signal Processing, Volume : 58, Issue : 5, May 2010, pp. 2901-2907.

[J10*] P. S. La Rosa, **A. Renaux**, A. Nehorai et C. H. Muravchik, "Barankin-type lower bound on multiple change-point estimation", *IEEE* Transactions on Signal Processing, Volume : 58, Issue : 11, Nov. 2010, pp. 5534-5549.

[J11] M. N. El Korso, R. Boyer, **A. Renaux** et S. Marcos, "Statistical resolution limit of the uniform linear cocentered orthogonal loop and dipole array", *IEEE* Transactions on Signal Processing, Volume : 59, Issue : 1, Jan. 2011, pp. 425-431.

[J12*] M. N. El Korso, R. Boyer, **A. Renaux** et S. Marcos, "Multidimensional statistical resolution limit for the multidimensional harmonic retrieval model : hypothesis test and Cramér-Rao bound approaches", à paraître dans Eurasip Journal on Advances on Signal Processing special issue on Advances in Angle-of-Arrival and Multidimensional Signal Processing for Localization and Communications.

[J13*] D. T. Vu, **A. Renaux**, R. Boyer et S. Marcos, "A study of various 3D array geometry using ULA branches : a Cramér-Rao bound approach", à paraître dans Multidimensional Systems and Signal Processing, Springer.

[J14] M. N. El Korso, R. Boyer, **A. Renaux** et S. Marcos, "Statistical resolution limit for source localization with clutter interference in a MIMO radar context", à paraître dans *IEEE* Transactions on Signal Processing.

[J15] M. N. El Korso, R. Boyer, **A. Renaux** et S. Marcos, "Statistical analysis of achievable resolution limit in the near field context", à paraître dans Elsevier Signal Processing.

Revues internationales avec comité de lecture soumises :

[J16] N. D. Tran, **A. Renaux**, R. Boyer, S. Marcos et P. Larzabal, "Weiss-Weinstein bound for MIMO radar with colocated linear arrays for SNR threshold prediction", soumis à Elsevier Signal Processing.

[J17*] N. D. Tran, **A. Renaux**, R. Boyer, S. Marcos et P. Larzabal, "Performance Bounds for the pulse phase estimation of X-ray pulsars", soumis à *IEEE* Transactions on Signal Processing.

[J18] M. N. El Korso, R. Boyer, **A. Renaux** et S. Marcos, "On the statistical resolvability of point sources in subspace interference using a GLRT-based framework", soumis à Elsevier Signal Processing.

[J19*] D. T. Vu, **A. Renaux**, R. Boyer et S. Marcos, "Some results on the Weiss-Weinstein bound in array processing", soumis à *IEEE* Transactions on Signal Processing.

[J20] M. N. El Korso, **A. Renaux**, R. Boyer et S. Marcos, "Deterministic lower bounds on the mean square error for near field source localization", soumis à *IEEE* Transactions on Signal Processing.

1.4.3 Conférences internationales avec comité de lecture et actes

[CI1] **A. Renaux**, E. Boyer, P. Forster et P. Larzabal, "Non-efficiency and non-Gaussianity of a maximum likelihood estimator at high signal-to-noise ratio and finite number of samples", in Proc. of *IEEE* International Conference on Acoustics, Speech, and Signal Processing, ICASSP-04, Montreal, Canada.

[CI2] **A. Renaux**, P. Forster et E. Boyer, "Non asymptotic efficiency of a maximum likelihood estimator at finite number of samples", in Proc. of European Signal Processing Conference, EUSIPCO-04, Vienna, Austria.

[CI3] **A. Renaux**, P. Forster, et P. Larzabal, "A new derivation of the Bayesian bounds for parameter estimation", in Proc. of *IEEE* Workshop on Statistical Signal Processing, SSP-05, Bordeaux, France.

[CI4] **A. Renaux**, P. Forster, P. Larzabal, et C. Richmond "The Bayesian Abel bound on the mean square error", in Proc. of *IEEE* International Conference on Acoustics, Speech, and Signal Processing, ICASSP-06, Toulouse, France.

[CI5] E. Chaumette, J. Galy, F. Vincent, **A. Renaux**, et P. Larzabal "MSE bounds conditioned by the energy detector", in Proc. of the European Signal Processing Conference, EUSIPCO-07, Poznan, Pologne.

[CI6] S. Bay, C. Herzet, JM. Brossier, JP. Barbot, **A. Renaux** et B. Geller, "Derivation of a Bayesian Cramér-Rao bound for dynamical offset estimation", in Proc. of VIII *IEEE* Workshop on Signal Processing Advances in Wireless Communications, SPAWC-07, Helsinki, Finland.

[CI7] P. LaRosa, **A. Renaux**, et A. Nehorai, "Barankin bounds for multiple change points estimation : computational aspects", in proc. of the Second *IEEE* International Workshop on

Computational Advances in Multi-Sensor Adaptive Processing, CAMSAP-07, St. Thomas, U.S. Virgin Islands. (invited paper special session on minimal bounds)

[CI8*] E. Chaumette, **A. Renaux** et P. Larzabal, "Lower bounds on mean square error derived from mixture of linear and non-linear transformations of the unbiasedness definition", in Proc. of *IEEE* International Conference on Acoustics, Speech, and Signal Processing, ICASSP-09, Taipei, Taiwan.

[CI9] M. N. El Korso, R. Boyer, **A. Renaux** et S. Marcos, "Non-matrix closed form expressions of the Cramér-Rao bounds for near-field localization parameters", in Proc. of *IEEE* International Conference on Acoustics, Speech, and Signal Processing, ICASSP-09, Taipei, Taiwan.

[CI10] M. N. El Korso, R. Boyer, **A. Renaux** et S. Marcos, "Statistical resolution limits for multiple parameters of interest and for multiple signals", in Proc. of *IEEE* International Conference on Acoustics, Speech, and Signal Processing, ICASSP-10, Dallas, TX, USA.

[CI11] D. T. Vu, **A. Renaux** , R. Boyer et S. Marcos, "Performance analysis of 2D and 3D antenna arrays for source localization" in Proc. of the European Signal Processing Conference, EUSIPCO-10, Aalborg, Denmark.

[CI12] D. T. Vu, **A. Renaux** , R. Boyer et S. Marcos, "Closed-form expression of the Weiss-Weinstein bound for 3D source localization : the conditional case", in Proc. of *IEEE* Sensor Array Multichannel Workshop SAM-2010 , Kibutz Ma'ale Hahamisha, Israel.
(invited paper special session on lower bound in array processing)

[CI13*] E. Chaumette, **A. Renaux** et P. Larzabal, "New trends in deterministic lower bounds and SNR threshold estimation : from derivable bounds to conjectural bounds ", in Proc. of *IEEE* Sensor Array Multichannel Workshop SAM-2010 , Kibutz Ma'ale Hahamisha, Israel. (invited paper special session on lower bound in array processing)

[CI14] M. N. El Korso, R. Boyer, **A. Renaux** et S. Marcos, "Statistical resolution limit : application to passive polarized sources localization", in Proc. of Séminaire sur les Systèmes de Détection : Architectures et Technologies DAT-2011, Alger, Algérie.

[CI15] M. N. El Korso, R. Boyer, **A. Renaux** et S. Marcos, "Statistical resolution limit for source localization in a MIMO context", in Proc. of *IEEE* International Conference on Acoustics, Speech, and Signal Processing, ICASSP-11, Praha, Czech Republic.

[CI16] N. D. Tran, **A. Renaux**, R. Boyer, S. Marcos et P. Larzabal, "MIMO radar in the presence of modeling errors : A Cramér-Rao bound investigation", in Proc. of *IEEE* International Conference on Acoustics, Speech, and Signal Processing, ICASSP-11, Praha, Czech Republic.

[CI17] D. T. Vu, M. N. El Korso, R. Boyer, **A. Renaux** et S. Marcos, "Angular resolution limit for vector sensor arrays : detection and information theory approaches", in Proc. of *IEEE* Workshop on Statistical Signal Processing, SSP-11, Nice, France.
(invited paper special session on polarized signal processing)

[CI18] D. T. Vu, **A. Renaux** , R. Boyer et S. Marcos, "Weiss-Weinstein bound and SNR threshold analysis for DOA estimation with a COLD array", in Proc. of *IEEE* Workshop on Statistical Signal Processing, SSP-11, Nice, France.
(invited paper special session on polarized signal processing)

[CI19] M. N. El Korso, R. Boyer, **A. Renaux** et S. Marcos, "A GRLT-based framework for the multidimensional statistical resolution limit", in Proc. of *IEEE* Workshop on Statistical Signal Processing, SSP-11, Nice, France.

1.4.4 Conférences nationales (GRETSI) avec comité de lecture et actes

[CN1] P. Forster, E. Boyer, P. Larzabal et **A. Renaux**, "Non-efficacité et non-gaussianité asymptotiques d'un estimateur du maximum de vraisemblance à fort rapport signal sur bruit", Actes du 19ème Colloque GRETSI 2003, Paris, France.

[CN2] **A. Renaux**, P. Forster, et P. Larzabal, "Une nouvelle approche des bornes bayésiennes", Actes du 20ème Colloque GRETSI 2005, Louvain-la-Neuve, Belgium.

[CN3] S. Bay, C. Herzet, JM. Brossier, JP. Barbot, **A. Renaux** et B. Geller, "Bornes bayésiennes pour l'estimation de phase évoluant au cours du temps", Actes du 21ème Colloque GRETSI 2007, Troyes, France.

[CN4] D. T. Vu, **A. Renaux** , R. Boyer et S. Marcos, "Analyse des performances de réseaux de capteurs 2D et 3D pour la localisation de source", Actes du 22ème Colloque GRETSI 2009, Dijon, France.

[CN5] M. N. El Korso, R. Boyer, **A. Renaux** et S. Marcos, "Expressions non-matricielles des bornes de Cramér-Rao pour la localisation de source en champ proche", Actes du 22ème Colloque GRETSI 2009, Dijon, France.

[CN6] N. D. Tran, **A. Renaux**, R. Boyer, S. Marcos et P. Larzabal, "Erreurs de modèle pour les radars MIMO : une étude par la borne de Cramér-Rao", Actes du 23ème Colloque GRETSI 2011, Bordeaux, France.

[CN7] M. N. El Korso, **A. Renaux**, R. Boyer et S. Marcos, "Bornes inférieures de l'erreur quadratique moyenne pour la localisation de sources en champs proche", Actes du 23ème Colloque GRETSI 2011, Bordeaux, France.

[CN8] D. T. Vu, **A. Renaux** , R. Boyer et S. Marcos, "Borne de Weiss-Weinstein pour la localisation de source polarisé l'aide d'un réseau de capteurs COLD", Actes du 23ème Colloque GRETSI 2011, Bordeaux, France.

[CN9] D. T. Vu, M. N. El Korso, R. Boyer, **A. Renaux** et S. Marcos, "Résolution limite angulaire : approches basées sur la théorie de l'information et sur la théorie de la détection", Actes du 23ème Colloque GRETSI 2011, Bordeaux, France.

1.4.5 Polycopiés pour l'enseignement

[P1] Traitement numérique des signaux, IFIPS 1ère année, (61 pages).

[P2] Statistical signal processing, Vietnam National University, Hanoi (VNU), (106 pages).

1.4.6 Séminaires

[S1] "Minimal bounds on the mean square error", École Normale Supérieure de Cachan, Septembre 2004, Cachan, France.

[S2] "Estimation theory and minimal bounds : application to spectral analysis and passive Radar", ONERA (The French Aerospace Laboratory) Mars 2005, Palaiseau, France.

[S3] "Estimation lower bounds and synchronization issue in single carrier system", Newcom Autumn School on "Estimation Theory for wireless communications" 24-28 Octobre 2005, Paris, France.

[S4] "Estimation lower bounds : links between deterministic and Bayesian bounds", Newcom Workshop at Aalborg University, 18-19 Novembre 2005, Aalborg, Denmark.

[S5] "Minimal bounds on the mean square error : a tutorial", Washington University in St. Louis, Février 2007, St. Louis, MO, USA

[S6] "Bornes inférieures de l'erreur quadratique moyenne : c'est pratique !!!", Séminaire SONDRA, Février 2010, Supelec, Gif-sur-Yvette, France

[S7] "Multidimensional statistical resolution limit", Washington University in St. Louis, Mars 2010, St. Louis, MO, USA

[S8] "The lower bounds on the mean square error - Application to radar ", National University of Singapore, Septembre 2010, Singapore

1.5 Participation à des contrats

– Projet Digiteo TIMuCa (Traitement de l'Information MUlti-CApteurs) 2008-2012. Sur certains thèmes de recherche, qu'elle a labellisés Domaines d'Intérêt Majeur (DIM), la Région Île-de-France a confié l'organisation des appels à propositions et l'attribution des financements à des structures fédératrices. C'est ainsi que la Fondation Digiteo est chargée du DIM logiciels et systèmes complexes. Dans ce contexte, lors de mon arrivé en poste au Laboratoire des Signaux et Systèmes, nous avons répondu à un appel à projet sous la thèmatique capteurs logiciels conformables, adaptables, fusion de données. Il s'agissait de mettre en place une collaboration entre le Laboratoire des Signaux et Systèmes (S. Marcos, R. Boyer, et A. Renaux) et le laboratoire SATIE de l'École Normale Supérieure de Cachan (P. Larzabal). Ce projet de recherche est basé sur les travaux de recherches que j'ai effectués lors de mon post-doctorat. Ce projet se focalise sur des systèmes qui intègrent un réseau de capteurs pour la connaissance et la surveillance de l'environnement. Déjà utilisé dans les domaines de la téléphonie mobile, de l'internet (réseaux WiFi) et de la localisation (GPS), les systèmes multi-capteurs tendent également à se développer à d'autres domaines tels que l'imagerie médicale, l'aéronautique, et le nucléaire (déploiement d'une constellation de robots en milieu hostile). Nous proposons d'utiliser les degrés de libertés mis à notre disposition, à savoir les classiques diversités temporelles et spatiales (environnement multi-capteurs) mais surtout la diversité des formes d'ondes et la possible mobilité du réseau de capteurs afin d'augmenter les performances des systèmes actuels. Ces deux derniers points font actuellement l'objet de recherches intensives de part le monde du fait de leur potentiel pour l'amélioration des systèmes multi-capteurs du futur.
  Le montant reçu s'élève à 220000 euros pour le financement, entre autre, de deux thèses (Dinh Thang VU et Duy Tran NGUYEN) dont je suis co-directeur.

– De manière non-officielle, lors de mon séjour post-doctoral à Washington University in St. Louis (2006-2007), j'ai participé au BAA-07-02-IFKA, "Sensors as Robots", proposé par l'U.S. Air Force (Budget total 4.9M$) sous la supervision du Professeur Arye Nehorai. `http://www.fbo.gov/spg/USAF/AFMC/AFRLRRS/Reference-Number-BAA-07-02-IFKA/SynopsisP.html`. Le thème de ce projet concerne l'étude d'un système multi-Radars intelligent (polarimétrique, SAR, sonar, MIMO, cognitifs) déployé sur différents supports (air, terre et mer). J'ai participé à la rédaction du "white paper" et du "proposal" du point de vue technique (rédactions de la partie du projet concernant le design de formes d'ondes radar, l'optimisation de géométries d'antennes, les calculs de performances et les stratégies de contre contre mesures) et administratif (management des collaborateurs—11 personnes—, prévisions de budget, et diagramme de travail prévisionnel des différentes tâches).

1.6 Activités d'encadrements

Les détails scientifiques de ces encadrements seront développés dans le chapitre 2.

1.6.1 Encadrements de doctorants

– **Mohammed Nabil El Korso**, Thèse de l'Université Paris-Sud 11. Bourse ministérielle. Encadrants : S. Marcos (Directrice de thèse - 10%), R. Boyer (Co-directeur de thèse - 50%) et **A. Renaux (encadrant - 40%)**.
Thèse commencée en septembre 2008 et soutenue le 7 Juillet 2011.

Analyse de performances en traitement d'antenne. Bornes inférieures de l'erreur quadratique moyenne et seuil de résolution limite.

Composition du jury :

- Pascal Larzabal, Professeur des Universités, Université Paris-Sud 11/École Normale Supérieure de Cachan (Président)
- Jean-Marc Brossier, Professeur des Universités, Institut National Polytechnique de Grenoble (Rapporteur)
- Jean-Yves Tourneret, Professeur des Universités, Institut National Polytechnique de Toulouse (Rapporteur)
- Karim Abed-Meraim, Professeur assistant, Télécom ParisTech (Examinateur)
- Gérard Favier, Directeur de Recherche CNRS, I3S, Nice (Examinateur)
- Rémy Boyer, Maître de Conférences, Université Paris-Sud 11 (Co-directeur de thèse)
- Sylvie Marcos, Directeur de Recherche CNRS, L2S, Gif-sur-Yvette (Directrice de thèse)
- Alexandre Renaux, Maître de Conférences, Université Paris-Sud 11 (Co-directeur de thèse)

Publications associées : [J09][J11][J12][J14][J15] - [CI9][CI10][CI14][CI15][CI17][CI19][CN5][CN7][CN9] (+ [J18][J20] soumis).

– **Dinh Thang Vu**, Thèse de l'Université Paris-Sud 11. Bourse Digiteo.
Encadrants : S. Marcos (Directrice de thèse - 10%), **A. Renaux (Co-directeur de thèse - 60%)** et R. Boyer (encadrant - 30%). Thèse commencée en septembre 2008, soutenance prévue octobre 2011.

Outils statistiques pour le positionnement optimal de capteurs dans le contexte de la localisation de sources.

Composition du jury :

- Philippe Forster, Professeur des Universités, Université Paris 10 (Rapporteur)
- Yide Wang, Professeur des Universités, Université de Nantes (Rapporteur)
- Pascal Chevalier, Professeur des Universités, CNAM, Paris (Examinateur)
- Jean-Philippe Ovarlez, Maître de Recheche ONERA, Palaiseau (Examinateur)
- Rémy Boyer, Maître de Conférences, Université Paris-Sud 11 (Co-directeur de thèse)

- Sylvie Marcos, Directeur de Recherche CNRS, L2S, Gif-sur-Yvette (Directrice de thèse)
- Alexandre Renaux, Maître de Conférences, Université Paris-Sud 11 (Co-directeur de thèse)

Publications associées : [J13] - [CI11][CI12][CI17][CI18][CN4][CN8][CN9] (+ [J19] soumis).

– **Nguyen Duy Tran**, Thèse de l'École Normale Supérieure de Cachan. Bourse Digiteo. Encadrants : P. Larzabal (Directeur de thèse - 10%), **A. Renaux (Co-directeur de thèse - 50%)** S. Marcos (encadrant - 10%) et R. Boyer (encadrant - 30%). Thèse commencée en septembre 2009, soutenance prévue en septembre 2012.

Titre provisoire : Performances d'estimation pour les radar MIMO et la localisation de sources

Publications associées : [CI16][CN6] (+[J16][J17] soumis).

1.6.2 Participations non officielles à des encadrements de doctorants

Durant mon année de post-doctorat (2006/2007), j'ai participé de manière informelle à l'encadrement de deux doctorants.

– Stéphanie Bay (ENS Cachan/laboratoire SATIE, directeurs de thèse : B. Geller et J.-P. Barbot) sur le thème des bornes minimales pour les systèmes dynamiques. Nous avons appliqué ces bornes à un problème de synchronisation en communications numériques non entraîné. Ce travail a donné lieu à la publication d'un article de revue internationale [J6] ainsi que deux articles de conférence [CI6][CN3].
– Patricio LaRosa (Washington University in St. Louis, directeur de thèse : A. Nehorai) sur le thème de la détection de point de rupture avec application à la détection de contractions chez la femme enceinte. Nous avons travaillé sur la caractérisation des performances des estimateurs de point de rupture. A ce sujet, nous avons publié un article de revue internationale [J10] et un article de conférence internationale [CI7]. Patricio LaRosa est désormais post-doctorant et nous continuons cette étude dans le context bayésien (séjour de 2 semaines à Washington University in St. Louis en mars 2010).

1.6.3 Encadrements de stage de M2-R

– Tran Hoang Tung, Université Paris-Sud 11 et Vietnam National University, "Modèles de bruit hétéroskedastiques non-gaussien en traitement d'antenne".
Encadrants : **A. Renaux (90%)** et L.T. Nguyen (10%). Soutenue en décembre 2009.
Cette étude n'a pas donné lieu à des publications mais a permis de mettre à jour certaines inexactitudes dans les résultats actuels de la littérature. Nous continuons nos travaux avec L.T. Nguyen (séjour d'une semaines à l'université national du Vietnam en juin 2010).

1.6.4 Encadrements de stage de Magistère ENS

Il s'agit de stages qui ont eu lieu durant 8 semaines en juin et juillet 2010. J'ai été l'encadrant des étudiants suivants, tous deux étudiants fonctionnaires stagiaires à l'ENS Cachan en M1-IST :
– Adrien Mercier sur le thème des performances d'estimation.
– Chengfang Ren sur le thème des performances d'estimation.

1.6.5 Encadrements de stage de TER M2 et M1

Il s'agit d'un Travail d'Etude et de Recherche (TER) de deux semaines que j'ai encadré en Mai 2009. Les deux étudiants du M2-P réseaux et télécom (Université Paris-Sud 11) ont étudié l'implementation de l'algorithme du maximum de vraisemblance pour la localisation de sources en champ proche.

- Fatou Ndyaye.
- Tojo Nirina Raboanarivola.

1.7 Séjours à l'étranger

1.7.1 Dans le cadre de mes activités de recherches

- **Octobre 2006 - Aout 2007 (11 mois)** : Post-doc à Washington University in St. Louis, USA / Department of Electrical and Systems Engineering. Collaboration avec Arye Nehorai.
- **Décembre 2006 (2 semaines)** : University of California Berkeley. Collaboration avec Cédric Herzet (alors post-doctorant à Berkeley et désormais chargé de recherche à l'INRIA) sur la convergence de l'algorithme EM (au travers de la publication [J5]).
- **Mars 2010 (2 semaines)** : Visiting professor à Washington University in St. Louis, USA / Department of Electrical and Systems Engineering. Collaboration avec Arye Nerohai.
- **Juin 2010 (1 semaine)** : Visiting professor à l'Université national du Vietnam, Hanoi. Collaboration avec L.T. Nguyen.
- **Août 2010 (2 semaines)** : National University of Singapore. Collaboration avec J.P. Ovarlez.

1.7.2 Dans le cadre de mes activités d'enseignements

- **Janvier 2008 (2 semaines)** : Université national du Vietnam, Hanoi.
- **Janvier 2009 (2 semaines)** : Université national du Vietnam, Hanoi.
- **Juin 2010 (1 semaine)** : Université national du Vietnam, Hanoi.

1.8 Organisation de conférences

- Organisation d'une journée GDR-ISIS, "Bornes minimales en estimation" avec Rémy Boyer et Jean-Yves Tourneret. 08 Sep. 2008. (`http://gdr-isis.org/rilk/gdr/ReunionListe-488`).
- Membre du comité des programmes pour l'*IEEE* International Conferences on Advanced Technologies for Communications (ATC 2009), Hai Phong, Vietnam (`http://www.atc09.org/`).

1.9 Participation à des jury de thèses et M2R

- **Qi Cheng**, Thèse de l'Université Paris-Sud 11. Thèse soutenue le 9 juillet 2009.

  Contribution à l'estimation d'état dans les modèles non linéaires.

  Composition du jury :

- Messaoud Benidir, Professeur des Universités, Université Paris-Sud 11 (Président)
- François Dufour, Professeur des Universités, Université de Bordeaux 1 (Rapporteur)
- Pierre-Olivier Amblard, Directeur de Recherche CNRS, GIPSA Lab, Grenoble (Rapporteur)
- Jean-Philippe Ovarlez, Maître de Recheche ONERA, Palaiseau (Examinateur)
- Alexandre Renaux, Maître de Conférences, Université Paris-Sud 11 (Examinateur)
- Pascal Bondon, Directeur de Recherche CNRS, L2S, Gif-sur-Yvette (Directeur de thèse)

– **Stéphanie Bay**, Thèse de l'École Normale Supérieure de Cachan. Thèse soutenue le 6 avril 2010.

Contribution à l'étude des bornes de Cramér-Rao. Application à la synchronisation de phase.

Composition du jury :

- Pascal Larzabal, Professeur des Universités, Université Paris-Sud 11 / École Normale Supérieure de Cachan (Président)
- Jean-Pierre Cances, Professeur des Universités, Université de Limoges (Rapporteur)
- Olivier Rioul, Maître de Conférences HDR, Télécom ParisTech (Rapporteur)
- Jean-Marc Brossier, Professeur des Universités, Institut National Polytechnique de Grenoble (Rapporteur)
- Alexandre Renaux, Maître de Conférences, Université Paris-Sud 11 (Invité)
- Benoit Geller, Maître de Conférences HDR, ENSTA ParisTech (Directeur de thèse)
- Jean-Pierre Barbot, Maître de Conférences, École Normale Supérieure de Cachan (Co-directeur de thèse)

En outre, depuis 2008, je participe à un jury de stage de M2R ATSI par an.

1.10 Evaluations d'articles et de projets de recherches

Revues internationales et nationales : total : 59 relectures pour 10 revues
– *IEEE* Transactions on Signal Processing (29 relectures).
– *IEEE* Signal Processing Letters (4 relectures).
– *IEEE* Transactions on Information Theory (4 relectures).
– Elsevier Signal Processing Journal of the European Association for Signal Processing (EURASIP) (11 relectures).
– Springer Signal Image and Video Processing (3 relectures).
– *IEEE* Sensors Journal (2 relectures)
– *IEEE* Journal of Selected Topics in Signal Processing. Special Issue on Performance Limits of Ultra-Wideband Systems (1 relecture).
– Automatica (3 relectures).
– Revue traitement du signal (1 relecture).
– *IEEE* Transactions on Industrial Electronics (1 relecture).

Evaluations de projets de recherches : total : 2 relectures
– Israel Science Foundation (ISF)(2 relectures)

Livres : total : 1 relecture

– Revue Instrumentation, Mesure et Métrologie (Revue des Systèmes, Hermes-Lavoisier)

Conférences internationales et nationales : total : 29 relectures
J'ai été relecteur pour 21 conférences :
IEEE Workshop on Statistical Signal Processing (SSP-2005), Bordeaux, France. European Signal Processing Conference (EUSIPCO-2006), Firenze, Italy. *IEEE* International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC-2006), Helsinki, Finland. *IEEE* Workshop on sensor array and multi-channel processing (SAM-2006), Waltham, MA, USA. *IEEE* International Conference on Acoustics, Speech, and Signal Processing (ICASSP-2007), Honolulu, HI, USA. Fifteenth Annual Workshop on Adaptive Sensor Array Processing (ASAP-2007), Lexington, MA. International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS-2007), Xiamen, China. Neural Information Processing Systems Conference (NIPS-2007), Vancouver, B.C., Canada. European Signal Processing Conference (EUSIPCO-2007), Poznan, Poland. *IEEE* Workshop on Computational Advances in Multi-Channel Sensor Array Processing (CAMSAP-2007), US Virgin Island, USA. European Signal Processing Conference (EUSIPCO-2008), Lausanne, Switzerland. IFAC Symposium on System Identification, (SYSID 2009), Saint-Malo, France. *IEEE* Vehicular Technology Conference (VTC-2009-Fall), Anchorage, AK, USA. European Signal Processing Conference (EUSIPCO-2009), Glasgow, Scotland. *IEEE* Workshop on Statistical Signal Processing (SSP-2009), Cardiff, Wales, UK. *IEEE* International Conferences on Advanced Technologies for Communications (ATC 2009), Hai Phong, Vietnam. *IEEE* Sensor Array Multichannel Workshop (SAM-2010), Kibutz Ma'ale Hahamisha, Israel. *IEEE* Workshop on Statistical Signal Processing (SSP-2011), Nice, France. European Signal Processing Conference (EUSIPCO-2011), Barcelona, Spain. 23ème Colloque GRETSI 2011, Bordeaux, France. Irish Signals and Systems Conference (ISSC-2011), Dublin, Ireland.

1.11 Activités d'enseignements

Mes enseignements concernent (ou ont concerné) le traitement du signal, le codage de sources, les probabilités, l'analyse, l'algébre, les signaux et systèmes linéaires, l'automatique, l'algorithmique et language C dans des filières L3 Information Systèmes et Technologies, M1 Information Systèmes et Technologies, M2-R Automatique et Traitement du Signal, M2-P Réseaux et Télécom de l'université Paris-Sud 11 ainsi qu'en première et deuxième années d'école d'ingénieur (Polytech Paris-Sud, formation d'ingénieur de l'université Paris-Sud 11).

En particulier, je suis ou ai été responsable des modules :
– Signaux et images (M1 IST 452).
– Communications numériques et codage canal (M2P Réseaux et Télécom).
– Automatique (M1 IST 422).

En outre, dans le cadre du pôle universitaire Français (PUF) au Vietnam, les modules du M1-IST de l'université Paris-Sud 11 sont dupliqués à l'université nationale du Vietnam d'Hanoi (mais dispensés en Anglais). Je suis responsable du module 452 : Signaux et images (environ 50 heures de cours/TD/TP hors service). Pour cela je me suis rendu trois années de suite [1] à Hanoi et j'ai rédigé un polycopié d'une centaine de pages dans sa forme actuelle. J'ai également commencé une collaboration scientifique avec Linh-Trung Nguyen (assistant professor, Vietnam National University) sur le thème de la modélisation de bruit par des processus hétéroskedastiques. A ce sujet, nous avons encadré conjointement le stage de M2-R de Tran Hoang Tung.

1. Ce cours n'a pas ouvert pour l'année scolaire 2010-2011 du à un manque d'effectif

J'ai également été moniteur durant 3 ans (2003-2006) à l'IUT de Cachan (Université Paris-Sud 11) au département Génie Électrique et Informatique Industrielle (GEII2) et, durant ma dernière année (DEA 2002-2003) à l'École Normale Supérieure de Cachan, j'ai été vacataire à l'IUT de Vitry (Université Paris 12) au département Génie des Télécommunications et Réseaux (GTR). Durant ces quatre années, mes enseignements ont concerné l'électronique, le traitement du signal, les télécommunications et l'informatique industrielle. C'est ainsi que j'ai enseigné en première année IUT GTR, Licence IUP GEII, Licence professionnelle SEITR, Maîtrise IUP GEII, première année d'école d'ingénieur IFIPS. J'ai également rédigé un polycopié de cours (60 pages) avec exercices et mis au point un TP d'initiation à Matlab.

Le bilan quantitatif est donc le suivant :

2011-2012 Maître de conférences Université Paris-Sud 11 Service provisoire.
L3 - Information Systèmes et Technologie (IST)
– Algébre linéaire (30 heures equiv. TD de cours/TD)
– Probabilités et statistiques (24 heures de cours, 16 heures de TD et 16 heures de TP)
M1 - Information Systèmes et Technologie (IST)
– Signaux et images (13 heures de cours et 9 heures de TD)
– Automatique (22 heures de cours)
M2R - Automatique et Traitement du Signal (ATS)
– Remise à niveau traitement statistique du signal (12 heures de cours)
M2P - Réseaux et Télécom
– Communications numériques et codage canal (16 heures de cours)
IFIPS (désormais Polytech Paris-Sud)
– Traitement du signal (10H30 heures de TD) IFIPS Dpt Électronique 2ème année
Université de Hanoi - Vietnam
– Signaux et images (21 heures de cours/TD et 9 heures de TP). En collaboration avec un correspondant Vietnamien Linh-Trung Nguyen.

2010-2011 Maître de conférences Université Paris-Sud 11
Obtention (au titre du CNU) d'un demi-CRCT (6 mois) pour l'année scolaire 2010/2011.
L3 - Information Systèmes et Technologie (IST)
– Analyse fonctionnelle (30 heures de TD)
M1 - Information Systèmes et Technologie (IST)
– Signaux et images (9 heures de TD)
M2R - Automatique et Traitement du Signal (ATS)
– Remise à niveau traitement statistique du signal (12 heures de cours)
M2P - Réseaux et Télécom
– Communications numériques et codage canal (16 heures de cours)
IFIPS (désormais Polytech Paris-Sud)
– Traitement du signal (10H30 heures de TD) IFIPS Dpt Électronique 2ème année

2009-2010 Maître de conférences Université Paris-Sud 11
L3 - Information Systèmes et Technologie (IST)
– Probabilités et statistiques (16 heures de TD et 8 heures de TP)
M1 - Information Systèmes et Technologie (IST)
– Signaux et images (13H00 de cours, 9 heures de TD, et 8 heures de TP)
– Codage de sources (4 heures de TD et 12 heures de TP)
M2R - Automatique et Traitement du Signal (ATS)

– Remise à niveau traitement statistique du signal (12 heures de cours)

M2P - SESIS

– Traitement du signal et application sur DSP (16 heures de TP)

M2P - Réseaux et Télécom

– Communications numériques et codage canal (16 heures de cours)

Université de Hanoi - Vietnam

– Signaux et images (21 heures de cours/TD et 9 heures de TP). En collaboration avec un correspondant Vietnamien Linh-Trung Nguyen.

IFIPS (désormais Polytech Paris-Sud)

– Traitement du signal (10H30 heures de TD) IFIPS Dpt Électronique 2ème année
– Outils mathématiques pour le traitement du signal (24 heures de TD et 8 heures de TP) IFIPS Dpt Électronique 1ère année

2008-2009 Maître de conférences Université Paris-Sud 11

L3 - Information Systèmes et Technologie (IST)

– Informatique générale (36 heures de TP)
– Probabilités et statistiques (16 heures de TD et 16 heures de TP)
– Signaux et systèmes linéaires (12 heures de TD)

M1 - Information Systèmes et Technologie (IST)

– Signaux et images (7 H45 de cours, 9 heures de TD, et 24 heures de TP)
– Codage de sources (12 heures de TP)

M2R - Automatique et Traitement du Signal (ATS)

– Remise à niveau traitement statistique du signal (12 heures de cours)

Université de Hanoi - Vietnam

– Signaux et images (42 heures de cours/TD et 9 heures de TP).

IFIPS (désormais Polytech Paris-Sud)

– Traitement du signal (10H30 heures de TD) IFIPS Dpt Électronique 2ème année
– Fonction de l'électroniques (8 heures de TP) IFIPS Dpt Optronique 1ère année
– Outils mathématiques pour le traitement du signal (24 heures de TD et 8 heures de TP) IFIPS Dpt Électronique 1ère année

2007-2008 Maître de conférences Université Paris-Sud 11

L3 - Information Systèmes et Technologie (IST)

– Informatique générale (36 heures de TP)
– Probabilités et statistiques (20 heures de TD et 8 heures de TP)

M1 - Information Systèmes et Technologie (IST) et M1 - Informatique (pour la partie codage de sources)

– Signaux et images (12 heures de TP)
– Théorie de l'information et codage de source (4 heures de TD et 8 heures de TP)

M2R - Automatique et Traitement du Signal (ATS) / Réseaux et Télécommunications (RT)

– Codage de sources (8 heures de TP)
– Remise à niveau traitement du signal (4 heures de TD et 4 heures de TP)

IFIPS Dpt Électronique 1ère année

– Outils mathématiques pour le traitement du signal (24 heures de TD et 8 heures de TP)

IFIPS Dpt Électronique 2ème année

– Traitement du signal (25 heures de TD)

IFIPS Dpt Optronique 2ème année

– Fonction de l'électroniques (40 heures de TP)

Université de Hanoi - Vietnam

– Signaux et images (45 heures de cours/TD et 9 heures de TP).

2005-2006 Moniteur Université Paris-Sud 11 (IUT de Cachan)
– Traitement du signal, IFIPS 1ère année. 20 heures équivalent TD de cours/TD/TP.
– Traitement du signal, Licence professionnel Systèmes embarqués et informatique temps réel (SEITR). 32 heures équivalent TD de cours/TD/TP.

2004-2005 Moniteur Université Paris-Sud 11 (IUT de Cachan)
Électronique, Maîtrise IUP Génie électrique et informatique industrielle (GEII). 88 heures équivalent TD de cours/TD/TP.

2003-2004 Moniteur Université Paris-Sud 11 (IUT de Cachan)
Électronique numérique, Licence IUP Génie électrique et informatique industrielle (GEII). 64 heures équivalent TD de TD/TP.

2002-2003 Vacataire Université Paris 12 (IUT de Vitry)
Télécommunications, 1ère année d'IUT Génie des télécommunications et réseaux (GTR). 60 heures de TP.

1.12 Divers

– Membre du comité de sélection pour le poste MCF-069 (section 61/63) CNAM Paris 2011.
– Obtention (au titre du CNU) d'un demi-CRCT (6 mois) pour l'année scolaire 2010/2011.
– Membre de la CCSU 60-61-62 Université Paris-Sud 11 Collège B 2010.
– Titulaire de la Prime d'Excellence Scientifique (ex PEDR) depuis octobre 2009.
– 2009 Représentant du L2S pour la journée de présentation des laboratoires et master recherches au département EEA de l'École Normale Supérieure de Cachan (journée du 17-12-2009).
– Représentant de la division Signaux du L2S pour la mise en place du nouveau site web du laboratoire (réunions débutées en janvier 2010). (mise en place du cahier des charges, faisabilité technique, etc.)
– Membre du comité de sélection pour le poste MCF-1374 (section 61) Université Paris X Nanterre 2009.
– Qualifié aux fonctions de maître de conférences en section CNU 61 (génie informatique, automatique et traitement du signal) 2006.
– *IEEE* member 2008-présent ; *IEEE* Student member 2006-2007.
– Membre du conseil de laboratoire SATIE (représentant des doctorants) 2005-2006.
– Membre du European Network of Excellence in Wireless COMmunications (NEWCOM) Department 1 (Analysis and Design of Algorithms for Signal Processing at Large in Wireless Systems) and in Departement 2 (MIMO Radio Channel Modelling for Design Optimisation and Performance Assessment of Next Generation Communication Systems) 2004-2006.

Chapitre 2

Synthèse des activités de recherche

Je joins en annexe les articles [J5], [J6], [J7], [J8], [J10], [J12], [J13], [J17], [J19], [CI8], et [CI13] qui me semblent être les contributions les plus pertinentes concernant les activités de recherches détaillées dans ce chapitre. J'ai choisi de détailler uniquement les travaux réalisés avec les thèsards que j'ai encadrés. De ce fait, certaines parties de ce chapitre issues de collaborations plus ponctuelles sont volontairement résumées au strict minimum et le lecteur est invité à lire l'annexe correspondante.

2.1 Contexte et premiers résultats

Je décris dans ce chapitre mon activité de recherche depuis mon recrutement en tant que Maître de conférences à l'Université Paris-Sud (octobre 2007) jusqu'à aujourd'hui. Mais, avant de rentrer dans les détails qui ont motivé cette recherche, un rapide voyage vers mes travaux antérieurs (DEA, thèse) s'impose.

Dans le cadre de mon DEA puis de mon doctorat, j'ai étudié le problème de l'estimation de paramètres physiques de signaux multiples captés par un réseau de capteurs. Bien que la motivation initiale ait été d'ordre militaire avec le développement du radar et du sonar pendant la seconde guerre mondiale, les applications civiles sont aujourd'hui nombreuses : positionnement par satellite, téléphonie mobile, etc. Les paramètres d'intérêt peuvent être les directions d'arrivées (localisation de sources), l'amplitude des signaux, le nombre de signaux et plus généralement le canal de propagation. La théorie de l'estimation offre un cadre solide pour résoudre ces problèmes et a fourni une large classe d'algorithmes pour estimer ces différents paramètres d'intérêt. Je me suis focalisé sur les méthodes dites paramétriques qui exploitent un modèle de signaux reçus. Plus particulièrement, en estimation dite paramétrique, on se fixe un modèle d'observation dépendant des paramètres d'intérêt et, à partir d'une règle pré-établie basée sur ce modèle, on estime les paramètres d'intérêt (figure 2.1).

Le modèle d'observation est basé sur les connaissances de l'utilisateur concernant le processus physique considéré. Il dépend généralement d'une fonction non-linéaire des paramètres et incorpore un modèle statistique. La règle d'estimation constitue la pierre angulaire du traitement du signal moderne. Elle conditionne les performances d'estimation, c'est-à-dire le fait que l'estimée soit "suffisamment proche" de la vraie valeur du paramètre et la charge de calcul à mettre en œuvre lors de l'estimation. L'expérience montre qu'un estimateur présentant des performances proches de l'optimalité devient gourmand en charge de calcul alors qu'un estimateur sous-optimal demande une charge de calcul moindre ce qui présente un intérêt en terme d'implémentation pratique.

L'estimation d'un paramètre s'effectue généralement par la recherche d'un maximum global

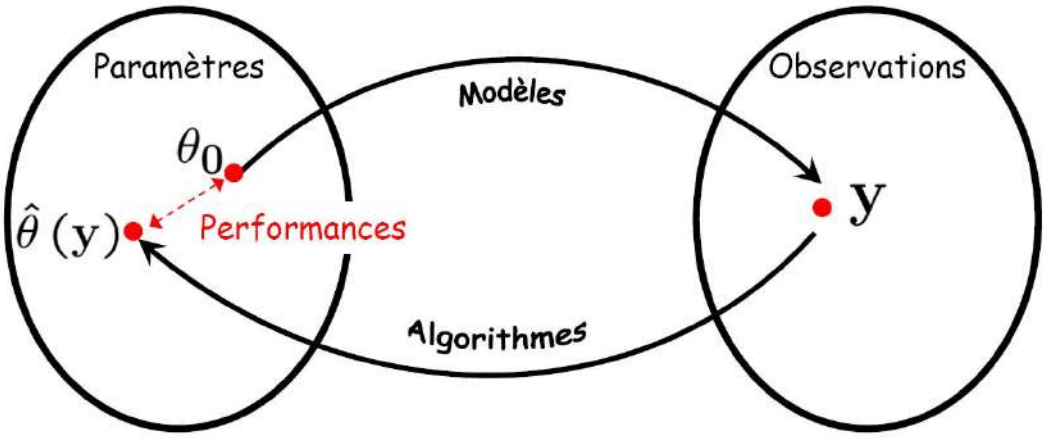


FIGURE 2.1 – Principe de l'estimation paramétrique.

d'un critère dépendant de la règle d'estimation. Lorsque la recherche du paramètre s'effectue sur un support fini, comme par exemple lors de l'estimation d'un angle ou d'une fréquence normalisée, l'Erreur Quadratique Moyenne (EQM) de l'estimateur qui est l'indicateur de performance le plus répandu s'articule autour de trois régions [Van68] [RB74]. Lorsque le Rapport Signal sur Bruit (RSB) ou le nombre d'observations est élevé, la région est dite asymptotique et l'erreur d'estimation est généralement faible. Lorsque le RSB ou le nombre d'observations décroît, il apparaît un accroissement rapide de l'EQM, donc de l'erreur, dû à l'apparition d'outliers dans le critère. On appelle cette région la zone de décrochement. Enfin, lorsque le RSB ou le nombre d'observations est très faible, le signal observé se réduit principalement à la composante de bruit, ce qui conduit à une distribution de l'estimée quasi-uniforme sur le support du paramètre. L'EQM exhibe alors un comportement plat : c'est la zone de non-information.

La figure 2.2 donne l'évolution de l'EQM de l'estimateur au sens du maximum de vraisemblance de la fréquence d'une cisoïde bruitée (gaussien), pour les trois régions susmentionnées (10 observations et 10000 réalisations).

La zone de décrochement d'un estimateur délimite sa zone de fonctionnement optimal, c'est-à-dire la zone où l'on peut faire abstraction des outliers dans l'étude des performances. Avec le durcissement des cahiers des charges (temps d'acquisitions réduits, environnements électromagnétiques perturbés, etc...), les estimateurs sont désormais amenés à travailler dans une région de plus en plus proche de cette limite, voire en dessous. En conséquence, la connaissance de la valeur du RSB ou du nombre d'observations, pour un scénario donné, pour laquelle cette rupture brutale des performances apparaît est fondamentale dans l'étude des performances d'estimation et doit être considérée au même titre que l'étude du biais ou de la variance.

La borne de Cramér-Rao (BCR) [Fis22] [Dug37] [Fre43] [Dar45] [Rao45] [Cra46] est également reportée sur la figure 2.2. Elle donne la variance minimale qu'un estimateur non biaisé peut espérer atteindre et est l'outil le plus utilisé par la communauté du traitement du signal pour porter un jugement sur les performances d'un estimateur. Un estimateur qui atteint la BCR est dit efficace. On pourra noter sur cet exemple l'efficacité asymptotique de l'estimateur du maximum de vraisemblance. Néanmoins, dans les zones non-asymptotiques, c'est-à-dire la zone de décrochement et la zone de non-information, la BCR n'est plus pertinente puisqu'elle repose sur un développement limité à l'ordre 1 de l'erreur alors que la présence d'outliers est un phénomène global.

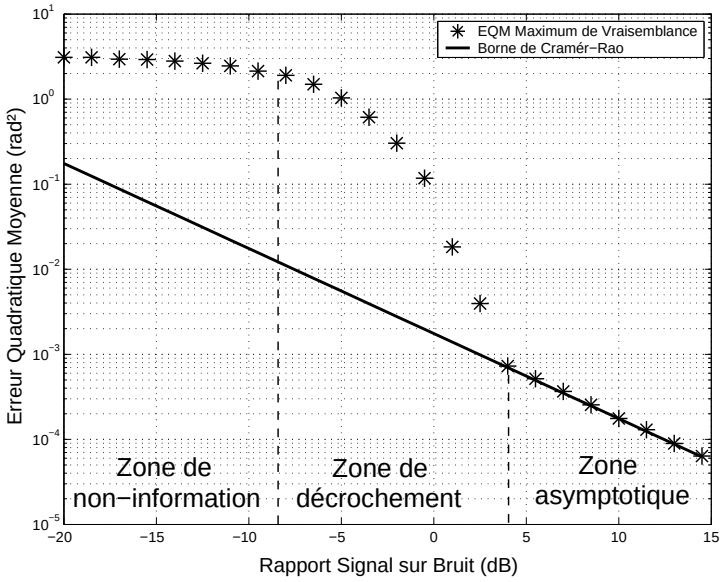


FIGURE 2.2 – Comportement de l'EQM d'un estimateur du maximum de vraisemblance.

Dans le contexte du traitement d'antenne, et en particulier de la localisation des sources, les méthodes paramétriques sont souvent dites à haute résolution puisqu'elles permettent, sous certaines conditions, d'avoir un pouvoir de résolution infini (l'aptitude à séparer deux sources infiniment proches). On peut classer ces techniques en deux catégories : les méthodes de sous-espaces et la méthode du maximum de vraisemblance. Les méthodes de sous-espaces s'appuient sur une décomposition de l'espace des observations, grâce à la matrice de covariance des observations, en deux sous-espaces : le sous-espace signal et le sous-espace bruit, et exploitent la propriété d'orthogonalité entre ceux-ci. On citera par exemple la méthode MUSIC (MUltiple Signal Classification) [Sch81] et ses nombreuses variantes : l'algorithme Min-Norm [KT83], ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) [PRK86], WSF (Weighted Subspace Fitting) [VOK91], etc. Les méthodes de sous-espaces requièrent généralement une simple optimisation mono-dimensionnelle. La méthode du maximum de vraisemblance s'appuie quant à elle sur une exploitation des connaissances statistiques des observations par l'intermédiaire de la fonction de vraisemblance. La méthode du maximum de vraisemblance conduit à l'optimisation d'un critère multi-dimensionnel. Elle est souvent qualifiée "d'optimale" pour ses bonnes propriétés statistiques en termes de consistance et d'efficacité. Le revers de la médaille est bien sûr une charge de calcul élevée pour la méthode du maximum de vraisemblance, mais l'augmentation de la puissance des calculateurs a conduit la communauté du traitement du signal à s'y intéresser.

Dans ce cadre, mes activités de recherche au sein du laboratoire SATIE (DEA et doctorat) se sont articulées selon deux axes.

Premièrement, nous avons caractérisé le comportement des estimateurs de directions d'arrivées dans la zone asymptotique en RSB et pour un nombre d'observations fini. En effet, si les performances asymptotiques en terme de nombre d'observations ont été largement étudié dans la littérature [SN90b] [SN90a] [VO91] les performances asymptotiques en terme de RSB (pour un nombre d'observations fini) ont été très peu traitées (par exemple par le biais de simulations dans [Ath05]).

Dans le contexte d'estimation de directions d'arrivées, il existe deux méthodes du maximum

de vraisemblance dépendant du modèle des signaux sources. Lorsque les signaux sources sont modélisés par un processus aléatoire gaussien, on parle de la méthode du Maximum de Vraisemblance Stochastique (MVS) [Böh86]. Ce modèle trouve son utilité dans des applications où le théorème de la limite centrale est applicable, par exemple, l'estimation d'échos Doppler en météorologie (étude d'orages) ou encore la localisation de cibles avec un large nombre de points brillants. Ce modèle est intéressant dans le sens où, en plus des directions d'arrivées et de la puissance du bruit, il reste à estimer le matrice de covariance finie des signaux reçus. D'un autre côté, lorsque les signaux sources sont supposés déterministes, on parle de la méthode du Maximum de Vraisemblance Déterministe (MVD) [Böh84]. Ce modèle trouve sont utilité dans des applications où le modèle gaussien pour les signaux est inapplicable, par exemple pour des signaux de communications numériques. Ici il faut non seulement estimer les directions d'arrivées et la puissance du bruit, mais également l'amplitude des signaux pour chaque observation ce qui devient problématique pour un grand nombre d'observations.

- Dans le cadre du MVS, c'est-à-dire lorsque les signaux sources sont modélisés par un processus aléatoire gaussien, nous avons démontré que l'estimateur du MVS pour les directions d'arrivées est non-gaussien et non-efficace (il n'atteint pas la borne de Cramér-Rao stochastique) lorsque le nombre d'observations T est fini et le RSB infini [RFBL04] [RFBL] [FBLR03]. Ce résultat est valide quel que soit le nombre de sources et quelle que soit la géométrie de l'antenne. Dans le cas d'un scenario à une seule source, nous avons démontré que la loi de l'estimateur est une loi de student à $2T$ degrés de liberté et que la variance de l'estimée était égale à $T/(T-1)$ fois la borne de Cramér-Rao stochastique. Dans le cas d'un scenario à deux sources nous avons établi une forme analytique de la matrice de covariance des estimées [RFB04] [RFBL]. Ces résultats sont en opposition avec le cas où le nombre d'observations tend vers l'infini où l'estimateur du MVS est connu pour être gaussien et efficace [SN90a].
- Lorsque les signaux sources sont supposés déterministes, nous avons établi la gaussianité et l'efficacité asymptotique en RSB de l'estimateur du MVD quel que soit le nombre de sources et quelle que soit la géométrie de l'antenne [RFCL06]. Ce travail a été réalisé en collaboration avec Eric Chaumette, à l'époque à Thales Air Defence et désormais à l'ONERA. Ces résultats sont à opposer au cas où le nombre d'observations tend vers l'infini où l'estimateur du MVD est connu pour être gaussien mais non-efficace (puisque le nombre de paramètres à estimer tend vers l'infini) [SN90b].

La figure 2.3 résume les deux résultats principaux que nous avons obtenus sur les performances des méthodes du Maximum de Vraisemblance en traitement d'antenne.

La seconde partie de notre étude a été dédiée à la prédiction du décrochement susmentionné et à l'étude des performances des estimateurs dans les trois zones de fonctionnement. Dans ce contexte, nous avons d'abord étudié et établi des bornes inférieures de l'EQM plus précises que la borne de Cramér-Rao. En effet, celle-ci devient trop optimiste dans la zone non-asymptotique et ne rend pas compte du phénomène de décrochement. L'étude des bornes inférieures de l'erreur quadratique moyenne constituent la part la plus importante de ma thèse et de tous les travaux que j'ai effectués depuis lors.

Les bornes inférieures de l'EQM fournissent des indications sur les performances ultimes qu'un estimateur peut espérer atteindre pour un modèle d'observation donné. En conséquence, elles sont utilisées comme "benchmarks" afin de jauger les performances d'un estimateur et de savoir si, à partir d'un cahier des charges donné, une amélioration est possible. Il existe une pléthore de bornes inférieures de l'EQM qui ont été dérivées depuis plus de soixante ans à l'aide de différentes

	MV stochastique	MV déterministe
Observations $\rightarrow \infty$ (RSB fini) [SN90]	Gaussien Efficace	Gaussien Non-efficace
RSB $\rightarrow \infty$ (Observations finies) Renaux et al.	Non-gaussien Non-efficace	Gaussien Efficace

FIGURE 2.3 – Performances asymptotiques des méthodes du Maximum de Vraisemblance en traitement d'antenne

inégalités mathématiques telles que : l'inégalité de Cauchy-Schwartz, l'inégalité de Kotelnikov, l'inégalité de Hölder, l'inégalité de Ibragimov-Hasminskii, l'inégalité de Bhattacharyya, l'inégalité de Kiefer, etc.

Les bornes inférieures de l'EQM se classent en deux catégories distinctes : les bornes déterministes lorsque les paramètres $\boldsymbol{\theta}_0$ sont supposés déterministes et les bornes bayésiennes lorsque les paramètres $\boldsymbol{\theta}$ sont supposés aléatoires avec une densité de probabilité *a priori* $p(\boldsymbol{\theta})$. Les bornes déterministes sont : la borne de Cramér-Rao [Fis22] [Dug37] [Fre43] [Dar45] [Rao45] [Cra46], la borne de Bhattacharyya [Bha46] [FG52], la borne de Chapman-Robbins [Ham50] [CR51] [Kie52], la borne de Barankin [Bar49] [MS69], et la borne d'Abel [Abe90] [Abe93]. Les bornes bayésiennes se sub-divisent en deux catégories : les bornes de la famille Ziv-Zakaï qui dérivent d'un problème de test d'hypothèse binaire (et plus généralement d'un test d'hypothèses M-aires) et les bornes de la famille Weiss-Weinstein qui dérivent, comme les bornes déterministes, d'un principe d'inégalité de covariance. La famille Ziv-Zakaï est composées de la borne de Ziv-Zakaï [ZZ69], de la borne de Bellini-Tartara [BT74], de la borne de Chazan-Zakaï-Ziv [CZZ75], de la borne de Weinstein [Wei88], de la borne de Bell-Steinberg-Ephraim-VanTrees [BSET97] et de la borne de Bell [Bel95]. La famille Weiss-Weinstein est composée de la borne de Cramér-Rao bayésienne [Van68], de la borne de Bobrovsky-MayerWolf-Zakaï [BMWZ87], de la borne de Bhattacharyya bayésienne [Van68], de la borne de Bobrovsky-Zakaï [BZ76], de la borne de Reuven-Messer [RM97] et de la borne de Weiss-Weinstein [WW85].

Il faut garder présent à l'esprit que les bornes déterministes sont des minorants de l'EQM locale en $\boldsymbol{\theta}_0$ c'est-à-dire de

$$\mathbf{EQM}_{locale}(\boldsymbol{\theta}_0) = \int_{\Omega} \left(\hat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta}_0\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta}_0\right)^H p(\mathbf{y}|\boldsymbol{\theta}_0)\, d\mathbf{y}, \tag{2.1}$$

où $\mathbf{y} \in \Omega$ est un vecteur aléatoire d'observations, $p(\mathbf{y}|\boldsymbol{\theta}_0)$ représente la vraisemblance des observations paramétrée par la vraie valeur du paramètre $\boldsymbol{\theta}_0$ dont $\hat{\boldsymbol{\theta}}(\mathbf{y})$ est un estimateur.

D'un autre côté, les bornes bayésiennes sont des minorants de l'EQM globale, c'est-à-dire de

$$\mathbf{EQM}_{globale} = \int_{\Theta}\int_{\Omega} \left(\hat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta}\right)^H p(\mathbf{y}, \boldsymbol{\theta})\, d\mathbf{y} d\boldsymbol{\theta}, \tag{2.2}$$

où $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ est le vecteur de paramètres aléatoires muni d'une densité de probabilité *a priori* $p(\boldsymbol{\theta}) = \frac{p(\mathbf{y},\boldsymbol{\theta})}{p(\mathbf{y}|\boldsymbol{\theta})}$ où $p(\mathbf{y}, \boldsymbol{\theta})$ est la densité de probabilité jointe des observations et des paramètres.

Cette profusion désordonnée a nuit à l'utilisation et à la dissémination de ces bornes puisqu'il devenait difficile pour le chercheur non spécialiste de discerner les différents concepts théoriques sous-jacents à leur l'établissement permettant de les retranscrire dans le contexte d'un problème d'estimation spécifique.

Les travaux théoriques sur les bornes minimales que nous avons menés sont basés sur les résultats concernant l'unification des bornes déterministes [Gla72]. En effet, l'ensemble des bornes déterministes peut être exprimé comme la solution d'un problème de minimisation sous contraintes. Le choix spécifique des contraintes conduit aux différentes bornes déterministes.

- En adoptant la même demarche, nous avons unifié les bornes Bayésiennes de la famille Weiss-Weinstein [RFL05a] [RFL05b] [RFL$^+$08]. Nous avons démontré que cette famille est également la solution d'un problème d'optimisation sous contraintes. Particulièrement, nous avons établi un pont entre les bornes susmentionnées et la "meilleure" borne bayésienne : l'EQM de l'estimateur de la moyenne conditionnelle. Ceci nous a permis d'établir une relation d'ordre entre les différentes bornes souvent inaccessible dans les développements traditionnels exploitant une inégalité de covariance. Ce point sera approfondi dans la suite de ce chapitre et le lecteur pourra également se reporter à l'**annexe C.**
- De plus, ces contraintes sont sensiblement similaires aux contraintes intervenant dans l'unification des bornes déterministes. A la lumière d'une telle analogie, il est apparu que toutes les bornes déterministes ont une version Bayésienne dans la famille Weiss-Weinstein et vice versa. Ceci nous a conduit à proposer deux nouvelles bornes minimales : la version Bayésienne de la borne d'Abel [RFLR06] en collaboration avec Christ Richmond du MIT Lincoln Laboratory (USA) et la version déterministe de la borne de Weiss-Weinstein. Cette dernière apparaît comme un outil performant dans l'étude de la prédiction du décrochement et est encore actuellement à l'étude (voir l'**annexe K** pour les avancées récentes de ce travail).

Afin de tirer partie des résultats théoriques obtenus, nous avons appliqué ces bornes à différents problèmes d'estimation.

Dans ce contexte, j'ai participé au réseau d'excellence Européen NEWCOM (Network of Excellence in Wireless COMmunications) dont le but est de renforcer les liens entre les différents groupes de recherche leader dans le domaine des "systèmes de communications numériques sans fils après la 3G". L'un des problèmes majeurs est la caractérisation des performances d'estimation traitée au sein du Département 2 (MIMO Radio Channel Modelling for Design Optimisation and Performance Assessment of Next Generation Communication Systems).

- Nous avons proposé une forme simplifiée des bornes d'Abel (déterministe et bayésienne). Cette forme dépend explicitement de la borne de Cramér-Rao (bayésienne) et de la borne de Chapman-Robbins (Bobrovsky-Zakaï dans le contexte bayésien). Elle est plus pertinente que ces deux dernières sans nécessiter une charge de calcul supplémentaire [RAFL07].
- Pour un problème de communications numériques (estimation de fréquences porteuses avec symboles pilotes) et d'analyse spectrale, nous avons proposé différentes formes analytiques pour l'ensemble des bornes déterministes et bayésiennes de la famille Weiss-Weinstein [Ren07].
- La plupart des bornes étudiées requièrent une maximisation par rapport à des paramètres appelés points tests. Nous avons étudié le comportement de ces points tests et nous en avons déduit des bornes moins gourmandes en charge de calcul [RAFL07].
- Nous avons également montré comment l'évolution de la valeur optimale des points tests permet de prédire de manière automatique les valeurs de RSB ou du nombre d'observations pour lesquelles le décrochement des estimateurs apparaît [RAFL07]. Les bornes de Weiss-Weinstein (déterministe et bayésienne) se sont révélées être un outil intéressant pour prédire

le décrochement des estimateurs.

Dans la suite de ce chapitre je m'attacherais à détailler les travaux de recherche que j'ai effectué lors de mon post-doctorat et mon integration à l'Université Paris-Sud. Il sera découpé en deux parties où j'ai choisi de présenter les travaux que je juge théoriques puis les travaux qui consistent en une application des bornes inférieures de l'EQM à des problèmes de traitement du signal.

2.2 Aspects théoriques

Dans cette section, je présente les aspects théoriques de mes travaux. Il s'agit de comprendre les bornes inférieures de l'EQM pour elles mêmes et sans forcement chercher une finalité applicative.

2.2.1 Bornes déterministes

J'ai travaillé, en collaboration avec Eric Chaumette, ONERA, sur différents aspect permettant d'unifier et d'étendre les bornes déterministes (lorsque les paramètres sont supposés déterministes) :

- Une grande variété de traitement incorpore un test de detection binaire afin de réduire le jeu des observations et d'estimer les paramètres d'intéret. Ce conditionnement n'était pas pris en compte lors de l'etablissement des bornes inférieures de l'EQM. Nous avons réecrit ces bornes dans ce contexte et les avons appliquées aux traitements de données radar. Voir [CGV$^+$07] pour les détails.
- Les bornes déterministes sont un outil particulièrement utile pour approximer la valeur du RSB pour laquelle l'EQM d'un estimateur décroche. L'une des bornes capable de prédire ce décrochement est la borne de Barankin et, plus précisement, les différentes approximations de cette borne. Ces approximations résultent d'une transformation linéaire de la contrainte de biais nulle uniforme introduite par Barankin. Néanmoins, il existe également des transformations non-linéaires qui peuvent étre utilisées pour certaine classe de densité de probabilité des observations. Nous avons mélangé ces transformations (linéaires et non-linéaires) afin d'obtenir de nouvelles bornes plus précises pour la prédiction du décrochement. Voir **annexe J** et **annexe K** pour les détails.

2.2.2 Bornes bayésiennes

Il s'agit en fait d'un travail commencé durant ma thèse mais dont les démonstrations ainsi qu'une présentation plus claire ont été continuées durant mon séjour post-doctoral et ont été publiées dans [RFL$^+$08] (joint en **annexe C**). La famille Weiss-Weinstein repose sur un principe d'inégalité de covariance et de ce fait semble plus naturelle que les bornes de la famille Ziv-Zakaï pour lesquelles une réécriture astucieuse de l'EQM est nécessaire. Nous avons présenté une unification des bornes de la famille Weiss-Weinstein basée sur l'unification précédement proposée pour les bornes deterministes dans [Gla72]. Il est bon de noter qu'une première approche pour l'unification des bornes de la famille Weiss-Weinstein a déjà été proposée par Weiss et Weinstein dans [WW88]. Cette approche se fonde sur l'inégalité suivante démontrée dans [WW88]

$$EQM_{globale} \geq \frac{\mathbb{E}^2_{\mathbf{y},\theta}\left[\theta\psi\left(\mathbf{y},\theta\right)\right]}{\mathbb{E}_{\mathbf{y},\theta}\left[\psi^2\left(\mathbf{y},\theta\right)\right]}, \tag{2.3}$$

où l'espérance mathématique porte sur la loi jointe $p(\mathbf{y},\theta)$ et où la fonction $\psi(\mathbf{y},\theta)$ doit vérifier

$$\int_{\Theta} \psi(\mathbf{y},\theta)\, p(\mathbf{y},\theta)\, d\theta = 0. \tag{2.4}$$

Weiss et Weinstein donnent plusieurs fonctions $\psi(\mathbf{y},\theta)$ vérifiant (2.4) pour lesquelles on retrouve la borne de Cramér-Rao bayésienne, la borne de Bhattacharyya bayésienne, la borne de Bobrovsky-Zakaï et la borne de Weiss-Weinstein. Malheureusement, il n'existe pas de règle générale pour trouver $\psi(\mathbf{y},\theta)$ et cette étude ne permet pas de déduire systématiquement une relation d'ordre entre les différentes bornes. Ici, l'unification de la famille Weiss-Weinstein s'appuie sur la meilleure borne bayésienne, à savoir l'EQM de l'estimateur de la moyenne conditionnelle. En réécrivant cette EQM nous unifions la borne de Cramér-Rao bayésienne, la borne de Bhattacharyya bayésienne, la borne de Bobrovsky-Zakaï, la borne de Reuven-Messer, la borne de Bobrovsky-MayerWolf-Zakaï et la borne de Weiss-Weinstein. L'intérêt de cette approche provient du fait qu'elle nous permet d'obtenir des relations d'ordre entre les différentes bornes proposées puisque nous les comparons directement à la borne ultime.

Dans le contexte bayésien, l'EQM globale minimale est obtenue par l'estimateur de la moyenne conditionnelle $\hat{\theta}(\mathbf{y}) = \int_{\Theta} \theta p(\theta|\mathbf{y})\, d\theta$. L'obtention d'une forme analytique de cette EQM est un problème généralement non-trivial. L'estimateur de la moyenne conditionnelle est solution du problème de minimisation suivant

$$\min_{\hat{\theta}(\mathbf{y})} \int_{\Omega}\int_{\Theta} \left(\hat{\theta}(\mathbf{y}) - \theta\right)^2 p(\mathbf{y},\theta)\, d\theta d\mathbf{y}. \tag{2.5}$$

On peut réécrire ce problème sous la forme d'un problème d'optimisation sous contraintes

$$\begin{cases} \min\limits_{v} \int\int_{\Omega\Theta} v^2(\mathbf{y},\theta)\, p(\mathbf{y},\theta)\, d\theta d\mathbf{y} \\ \text{sujet à } v(\mathbf{y},\theta) = \hat{\theta}(\mathbf{y}) - \theta \end{cases} \tag{2.6}$$

D'un autre côté, en posant $\mathbf{x} = \left[\begin{array}{cc} \mathbf{y}^T & \theta \end{array}\right]^T$ et $v(\mathbf{y},\theta) = \frac{u(\mathbf{y},\theta)}{\sqrt{p(\mathbf{y},\theta)}}$, la solution du problème d'optimisation sous contraintes suivant

$$\begin{cases} \min\limits_{v} \int\int_{\Omega\Theta} v^2(\mathbf{y},\theta)\, p(\mathbf{y},\theta)\, d\theta d\mathbf{y} \\ \text{sujet à} \int\int_{\Omega\Theta} v(\mathbf{y},\theta)\, g_k(\mathbf{y},\theta) \sqrt{p(\mathbf{y},\theta)} d\theta d\mathbf{y} = c_k \quad k = 0,...,K, \end{cases} \tag{2.7}$$

est donnée par le résultat suivant :

Soit $\mathbf{x} \in \mathcal{M}_{N\times 1}(\mathbb{R})$ un vecteur réel. Soient $f(\mathbf{x})$ et $g(\mathbf{x})$ deux fonctions de $\mathbb{R}^N \to \mathbb{R}$. Soit

$$\langle f(\mathbf{x}), g(\mathbf{x})\rangle = \int_{\mathbb{R}^N} f(\mathbf{x})\, g(\mathbf{x})\, d\mathbf{x}, \tag{2.8}$$

un produit scalaire de ces deux fonctions et sa norme associée $\|f(\mathbf{x})\|^2 = \langle f(\mathbf{x}), f(\mathbf{x})\rangle$. Soit $u(\mathbf{x})$, $g_0(\mathbf{x})$,..., et $g_K(\mathbf{x})$, un jeu de fonctions de $\mathbb{R}^N \to \mathbb{R}$ et c_0, c_1,..., et c_K, $K+1$ nombres réels. Alors, le minimum de $\|u(\mathbf{x})\|^2$ avec les $K+1$ contraintes

$$\langle u(\mathbf{x}), g_k(\mathbf{x})\rangle = c_k \quad \text{pour } k = 0,...,K, \tag{2.9}$$

est donné par

$$\begin{cases} \min\limits_{u} \|u(\mathbf{x})\|^2 = \mathbf{c}^T \mathbf{G}^{-1} \mathbf{c}, \\ \text{sujet à (2.9)}, \end{cases} \tag{2.10}$$

avec

$$\mathbf{c} = \begin{bmatrix} c_0 & c_1 & \cdots & c_K \end{bmatrix}^T, \tag{2.11}$$

et

$$G_{m,n} = \langle g_m(\mathbf{x}), g_n(\mathbf{x}) \rangle . \tag{2.12}$$

□

Nous avons établi l'équivalence entre l'unique contrainte du problème d'optimisation (2.6) et un continuum de contraintes particulières du problème d'optimisation (2.7). Ce continuum n'est pas utilisable en l'état dans le théorème précédent, puisque par définition, il correspond à une infinité de contraintes qui conduirait à un vecteur $\mathbf{c}$ et une matrice $\mathbf{G}$ de dimensions infinies. Néanmoins, en réduisant ce continuum, c'est-à-dire en choisissant des valeurs particulières pour $f(\mathbf{y}, \theta)$, h, et s en nombre fini, on obtiendra, par le théorème précédent, une EQM minimale globale inférieure à l'EQM de l'estimateur de la moyenne conditionnelle, donc, par définition, à des bornes minimales d'estimation. Le choix des paramètres $f(\mathbf{y}, \theta)$, h, et s donne une appréciation de la "distance" à l'EQM de l'estimateur de la moyenne conditionnelle, c'est-à-dire que plus on choisira de valeurs, plus on se rapprochera du continuum et donc de la meilleure borne minimale. Les choix des paramètres $f(\mathbf{y}, \theta)$, h, et s conduisant aux différentes bornes de la famille Weiss-Weinstein ainsi que les preuves et détails calculatoires sont donnés en **annexe C**.

2.2.3 Bornes hybrides

Nous avons parlé dans les sections précédentes de la borne de Cramér-Rao déterministe (BCR) et de la borne de Cramér-Rao bayésienne (BCRB). Ces deux bornes sont parmi les bornes les plus utilisées dans les problèmes classiques de traitement du signal. Cependant certains scénarios nécessitent l'estimation conjointe de paramètres déterministes et de paramètres aléatoires. Dans de tels cas, ni la BCR, ni la BCRB ne permettent d'évaluer les performances de l'estimation réalisée et il n'est pas toujours cohérent de contourner le problème grâce à une marginalisation ou grâce à une borne de Cramér-Rao modifiée. Le problème a été partiellement résolu par l'introduction d'une borne de Cramér-Rao dite hybride permettant de considérer un vecteur paramètres d'intérêt composé de paramètres déterministes et aléatoires [RS87a, RS87b, Mes06]. Cette borne hybride a été démontrée avec des hypothèses restrictives pour répondre à des scénarios particuliers qui avaient la particularité de présenter des paramètres aléatoires $\boldsymbol{\mu_r}$ et déterministes $\boldsymbol{\mu_d}$ (avec $\boldsymbol{\theta}^T = \left[\boldsymbol{\mu_r^T}, \boldsymbol{\mu_d^T}\right]^T$) indépendants c'est à dire tels que :

$$p(\boldsymbol{\mu_r} | \boldsymbol{\mu_d}) = p(\boldsymbol{\mu_r}) . \tag{2.13}$$

Les formules alors utilisées, en particulier dans [Mes06], ne sont donc pas valables dans le cas général.

Nous avons généralisé le travail mené par [RS87a, RS87b] dans le but d'avoir une formulation de la borne de Cramér-Rao générale et ce, quelle que soit la nature des paramètres d'intérêt (déterministes et/ou aléatoires). Ce travail a été effectué lors de la thése de doctorat de Stéphanie Bay et les détails sont donnés en annexe B.

Nous avons défini la matrice d'information hybride (MIH) de la manière suivante

$$\mathbf{H}\left(\boldsymbol{\mu_d}\right) \triangleq \mathbb{E}_{\mathbf{y},\boldsymbol{\mu_r}|\boldsymbol{\mu_d}}\left[-\Delta_{\boldsymbol{\mu}}^{\boldsymbol{\mu}} \ln\, p\left(\mathbf{y},\boldsymbol{\mu_r}|\boldsymbol{\mu_d}\right)\right] \tag{2.14}$$

$$= \mathbb{E}_{\mathbf{y},\boldsymbol{\mu_r}|\boldsymbol{\mu_d}}\left[\left(\nabla_{\boldsymbol{\mu}} \ln\, p\left(\mathbf{y},\boldsymbol{\mu_r}|\boldsymbol{\mu_d}\right)\right)\left(\nabla_{\boldsymbol{\mu}} \ln\, p\left(\mathbf{y},\boldsymbol{\mu_r}|\boldsymbol{\mu_d}\right)\right)^T\right], \tag{2.15}$$

En développant la log-vraisemblance :

$$\ln\, p\left(\mathbf{y},\boldsymbol{\mu_r}|\boldsymbol{\mu_d}\right) = \ln\, p\left(\mathbf{y}|\boldsymbol{\mu_r},\boldsymbol{\mu_d}\right) + \ln\, p\left(\boldsymbol{\mu_r}|\boldsymbol{\mu_d}\right),$$

nous obtenons :

$$\mathbf{H}\left(\boldsymbol{\mu_d}\right) = \mathbb{E}_{\boldsymbol{\mu_r}|\boldsymbol{\mu_d}}\left[\mathbf{F}(\boldsymbol{\theta})\right] + E_{\boldsymbol{\mu_r}|\boldsymbol{\mu_d}}\left[\Delta_{\boldsymbol{\mu}}^{\boldsymbol{\mu}} \ln\, p\left(\boldsymbol{\mu_r}|\boldsymbol{\mu_d}\right)\right],$$

où $\mathbf{F}(\boldsymbol{\theta})$ est la matrice d'information de Fisher

$$\mathbf{F}(\boldsymbol{\theta}) = \mathbb{E}_{\mathbf{y}|\boldsymbol{\theta}}\left[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \ln\, p(\mathbf{y}|\boldsymbol{\theta})\right]. \tag{2.16}$$

Nous avons ensuite démontré que l'inverse de la matrice d'information hybride est un minorant de l'EQM du vecteur paramètre d'intérêt $\boldsymbol{\theta}$. La démonstration est similaire à celle menée dans [Mes06] mais avec certaines hypothèses moins restrictives ce qui constitue une généralisation de l'inégalité de Cramér-Rao :

Si $\hat{\boldsymbol{\theta}}$ est un estimateur non biaisé du vecteur paramètre $\boldsymbol{\theta} = \left[\boldsymbol{\mu_r^T} \quad \boldsymbol{\mu_d^T}\right]^T$, alors

$$\mathbf{C}_{H\hat{\boldsymbol{\theta}}}\left(\boldsymbol{\mu_d}\right) \geq \left(\mathbf{H}\left(\boldsymbol{\mu_d}\right)\right)^{-1} = \left(\mathbb{E}_{\mathbf{y},\boldsymbol{\mu_r}|\boldsymbol{\mu_d}}\left[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \ln\, p\left(\mathbf{y},\boldsymbol{\mu_r}|\boldsymbol{\mu_d}\right)\right]\right)^{-1}, \tag{2.17}$$

où $\mathbf{C}_{H\hat{\boldsymbol{\theta}}}\left(\boldsymbol{\mu_d}\right)$ est la matrice de covariance, et $\mathbf{F}(\boldsymbol{\theta})$ la matrice d'information de Fisher.

Pour illustrer ce résultat, nous avons appliqué cette borne dans le cadre de l'estimation d'un offset de phase et de sa dérive linéaire dans le cas d'une transmission MPD2. Différentes contributions ont déjà partiellement traité ce problème. En particulier, dans leur article [BC07], Barbieri et Colavolpe ont calculé la borne de Cramér-Rao associée à l'estimation d'un offset de phase en considérant le bruit de phase comme un paramètre de nuisance dans une transmission data-aided. Nous avons de plus montré que la borne hybride présentée dans l'article de Messer [Mes06] conduit à une impasse en aboutissant à une matrice d'information non inversible.

On souhaite transmettre une séquence binaire non codée à travers un canal perturbé par un bruit additif blanc et gaussien. Après échantillonnage, les observations $\mathbf{y} = [y_1 \cdots y_K]^T$ s'écrivent en bande de base :

$$y_k = a_k e^{j\theta_k} + n_k \quad \text{avec } k = 1 \ldots K, \tag{2.18}$$

où a_k représente le $k^{\text{ème}}$ bit transmis inconnu. n_k représente toujours un bruit additif, circulaire, blanc, gaussien de moyenne nulle et de variance σ_n^2. Les symboles transmis sont considérés indépendants et identiquement distribués. Différentes sources de bruit viennent perturber la récupération de la phase. Nous avons travaillé sur un modèle général qui prend en compte la dérive linéaire présente dans chaque système de transmission ainsi qu'un bruit de phase de variance inconnue :

$$\theta_k = \theta_{k-1} + \xi + w_k, \tag{2.19}$$

où θ_k représente la phase que nous souhaitons estimer, ξ représente la dérive linéaire qui est un paramètre déterministe mais inconnu. w_k est un bruit additif blanc gaussien indépendant et identiquement distribué, de moyenne nulle et de variance σ_w^2 inconnue. Ce modèle est couramment utilisé pour décrire le comportement d'oscillateurs réels [McN94, DMR00]. Un des points intéressants à remarquer est que les observations sont indépendantes mais non identiquement distribuées. La log-vraisemblance s'écrit dans ce cas

$$\ln p\left(\mathbf{y}|\boldsymbol{\theta}_K, \boldsymbol{\mu_d}\right) = \sum_{k=1}^{K} \ln \left(\sum_{a\in\{-1,+1\}} p(y_k|a_k = a, \theta_k, \sigma_n)\, p(a_k = a) \right)$$
$$= \sum_{k=1}^{K} \left(\ln\left(-\pi\sigma_n^2\right) - \frac{1+|y_k|^2}{\sigma_n^2} + \ln\left(\cosh\left(\frac{2}{\sigma_n^2}\Re\left(y_k e^{-j\theta_k}\right)\right)\right)\right), \qquad (2.20)$$

où $\boldsymbol{\mu_d} \triangleq [\,\xi\ \sigma_n\ \sigma_w\,]^T$, et $\boldsymbol{\theta}_K = [\theta_1 \cdots \theta_K]^T$. Les symboles a_k sont considérés comme des paramètres de nuisance. $\mathbf{y}$ représente le vecteur d'observation et le vecteur paramètre d'intérêt est la concaténation de deux vecteurs, l'un représentant les paramètres inconnus déterministes $\boldsymbol{\mu_d}$ et l'autres les paramètres inconnus aléatoires :

$$\boldsymbol{\mu} \triangleq \left[\boldsymbol{\mu_r}^T\, \boldsymbol{\mu_d}^T\right]^T \quad \text{où} \quad \begin{array}{ll} \boldsymbol{\mu_r} & \triangleq \boldsymbol{\theta}_K = [\theta_1 \cdots \theta_K]^T, \\ \mu_d & \triangleq [\,\xi\ \sigma_n\ \sigma_w\,]^T. \end{array} \qquad (2.21)$$

Le lecteur trouvera en **annexe B** l'expression analytique de la borne de Cramér-Rao hybride ainsi que quelques résulats de simulations.

2.3 Aspects applicatifs

Je vais maintenant détailler les travaux qui concernent l'application des bornes inférieures de l'EQM à des problèmes de traitement du signal. Depuis mon post-doctorat (2006-2007) et mon recrutement en temps que Maître de Conférence en octobre 2007, mon activité de recherche à principalement consisté à montrer à la communauté scientifique l'utilité des bornes inférieures de l'erreur quadratique moyenne pour le traitement du signal. En particulier, j'ai souhaité accentuer mes recherches vers le fait que ces bornes ne sont pas de simple outils théoriques mais qu'elles peuvent être utiles en pratique dans les systèmes multi-capteurs du futur. Nous avons travaillé sur les bornes de Cramér-Rao qui restent une bonne source d'information concernant les performances asymptotiques des estimateurs, mais aussi sur la borne de Barankin et ces approximations ainsi que sur la borne de Weiss-Weinstein.

2.3.1 Borne de Cramér-Rao

2.3.1.1 Applications directes

2.3.1.1.1 Localisation de source en champ proche Dans le cadre de la thèse de Mohammed Nabil El Korso, nous nous sommes intéressés à la localisation de source en champ proche. La localisation passive de sources à l'aide d'une antenne composée d'un réseau de capteurs est un sujet d'une importance croissante avec plusieurs applications à la clef : radar, sismologie, communication numériques, etc. Le cas de la localisation de sources en champ lointain a été largement traité et une pléthore d'algorithmes d'estimation a été proposée dans la littérature [Kie52, Cap69, LC03b, KT83, LVT89, Van95, CM97, KV96, Mar98, Van02, SM05b]. Dans cette configuration, on peut faire l'hypothèse que les fronts d'ondes sont plans. Cependant, si les sources sont localisées dans le champ proche, la courbure des fronts d'ondes incidents sur les capteurs ne peut plus être négligée. Par conséquent, chaque source doit être caractérisée par son azimut et sa distance (la distance entre la source et un point de référence sur l'antenne considérée). Il existe différents algorithmes d'estimation adaptés à ce problème [HB91, YF98, GAMH05, ZC07, EBBM09], mais il existe très peu de travaux étudiant les performances optimales associées à ce modèle.

En traitement d'antenne on peut distinguer deux types de modèles concernant les signaux issus de sources [Van68, SN89, Kay93, OVSN93] : 1) le modèle déterministe (ou conditionnel), c'est-à-dire, où l'on suppose que les signaux sont déterministes mais inconnus, et 2) le modèle stochastique (ou inconditionnel), c'est-à-dire, où l'on suppose que les signaux émis suivent une loi Gaussienne complexe circulaire de moyenne nulle et de matrice de covariance $\mathbf{\Sigma}$ inconnue. La validité du modèle dépend de l'application en question. En effet, l'hypothèse du modèle stochastique n'est pas valable pour des applications telles que le radar avec des formes d'onde connues à l'émission [Van01, BT06, NS09] ou la communication radio [LC03a] par exemple. Dans ce cas, le choix légitime serait alors de considérer un modèle déterministe. Cependant, d'autres applications sont mieux décrites par le modèle stochastique, comme, le traitement sismique [Van02] ou la tomographie [Hay85].

Pour caractériser les performances des estimateurs, la borne de Cramér-Rao (BCR) est un outil mathématique très utilisé en traitement du signal. Cette dernière exprime une borne inférieure de la matrice de covariance d'erreur de tout estimateur non biaisé [Cra46]. Bien entendu, la BCR dépend du modèle considéré. De ce fait, dans le cas de l'estimation de direction d'arrivée d'une source située en champ proche nous avons étudié les deux BCRs ; la BCR pour le modèle déterministe, dénommée la BCR déterministe (BCRD) et la BCR pour le modèle aléatoire, dénommée la BCR stochastique (BCRS). Il faut noter qu'asymptotiquement, la BCRS est atteinte en terme de nombre d'observations par l'estimateur du maximum de vraisemblance stochastique (MVS) [SN89, OVSN93]. La BCRD est, quant à elle, atteinte asymptotiquement en terme de rapport signal à bruit (RSB) (ou en nombre de capteurs) par l'estimateur du maximum de vraisemblance déterministe (MVD) [RFCL06].

La plupart des résultats existants dans la littérature sur la BCRS et la BCRD concerne la localisation de sources en champ lointain. De plus, il est important de dire que la majeure partie de ces résultats donnent seulement des expressions matricielles de la matrice d'information de Fisher (MIF) qui est l'inverse de la BCR. De ce fait, le coût calculatoire associé à la BCR est très important pour un grand nombre des observations (en ce qui concerne la BCRD) ou un grand nombre de capteurs (en ce qui concerne la BCRS), d'où la nécessité d'avoir des expressions non matricielles pour le calcul de la BCR. Dans [SN90a], la BCRS a été indirectement calculée dans le cas asymptotique (en terme de nombre des observations) comme étant équivalente à la matrice de covariance du MVS. Dix ans après, Stoica *et al.* [SLG01], Pesavento et Gershman [PG01] et Gershman *et al.* [GSPL02] ont recalculé la BCRS sous forme matricielle (avant inversion de la MIF qui est simplement donnée par la formule de Slepian Bang [SM05b]) dans le cas d'un bruit blanc, coloré et de matrice de covariance inconnue. D'autre part, la BCRD pour le contexte champ lointain a été calculée par Stoica *et al.* dans [SN89].

Contrairement au champ lointain, la BCR pour les problèmes de localisation de source en champ proche a été peu étudiée. On peut trouver des expressions matricielles pour la BCRS dans [WF93]. Récemment, Grosicki *et al.* [GAMH05] ont étendu les formules matricielles de la BCRS dans le contexte du champ lointain au champ proche. A notre connaissance, aucune expression analytique non matricielle de la BCRS ou de la BCRD n'était disponible dans la littérature. Nous avons calculé et analysé des expressions non matricielles compactes de la BCRD et de la BCRS dans le cas d'une source à bande étroite située en champ proche.

Nous avons considéré une Antenne Linéaire Uniforme (ALU) composée de N ($N > 1$) capteurs avec une distance entre capteurs notée d. L'ALU reçoit un signal émis par une source à bande étroite située dans le champ proche. Par conséquent, le modèle d'observation peut s'écrire comme suit

$$y_n(t) = s(t)e^{j\tau_n} + v_n(t), \tag{2.22}$$

avec $t = 1, \ldots, T$ et $n = 0, \ldots, N-1$. $y_n(t)$ et $s(t)$ représentent le signal observé à la sortie

du $n^{ème}$ capteur et le signal source, respectivement. Le processus aléatoire $v_n(t)$ est un bruit additif et T est le nombre d'observations. Puisque la source est supposée dans le champ proche de l'antenne, le retard temporel τ_n qui représente le temps de propagation du signal de la source au $n^{ème}$ capteur est donné par [GAMH05]

$$\tau_n = \frac{2\pi r}{\lambda}\left(\sqrt{1+\frac{n^2d^2}{r^2}-\frac{2nd\sin\theta}{r}}-1\right), \tag{2.23}$$

où λ est la longueur d'onde et r, $\theta \in [0, \pi/2[$ représentent la distance et l'azimut du signal source, respectivement. Si la distance appartient à la région de Fresnel [HB91], c'est-à-dire, si

$$0.62(d^3(N-1)^3/\lambda)^{1/2} < r < 2d^2(N-1)^2/\lambda, \tag{2.24}$$

alors le temps de propagation τ_n peut être approximé par

$$\tau_n = \omega n + \phi n^2 + O\left(\frac{d^2}{r^2}\right), \tag{2.25}$$

où ω et ϕ sont généralement appelées les angles électriques. Ils s'expriment en fonction des paramètres physiques du problème à l'aide de $\omega = -2\pi\frac{d}{\lambda}\sin(\theta)$ et de $\phi = \pi\frac{d^2}{\lambda r}\cos^2(\theta)$.
Par conséquent, en utilisant (2.25), le modèle des observations peut s'écrire comme suit

$$y_n(t) = s(t)e^{j\left(\omega n+\phi n^2\right)} + v_n(t). \tag{2.26}$$

De ce fait, le vecteur d'observation peut être exprimé comme suit

$$\mathbf{y}(t) = [y_1(t)\dots y_N(t)]^T = \mathbf{a}(\omega,\phi)s(t) + \mathbf{v}(t), \tag{2.27}$$

où $\mathbf{v}(t) = [v_1(t)\dots v_N(t)]^T$ et où le $n^{ème}$ élément du vecteur directionnel $\mathbf{a}(\omega,\phi)$ est donné par $[\mathbf{a}(\omega,\phi)]_n = e^{j(\omega n+\phi n^2)}$ (notons que pour le champ lointant ϕ est supposé égal à zero).

Les hypothèses suivantes seront utilisées :

– On admet que le bruit est un processus complexe circulaire blanc Gaussien aléatoire avec une moyenne nulle et une variance inconnue σ^2,
– Le bruit est supposé décorélé temporellement et spatialement,

La fonction de densité de probabilité conjointe des observations $\boldsymbol{\chi} = [\mathbf{y}^T(1)\dots\mathbf{y}^T(T)]^T$ pour un vecteur de paramètres inconnus $\boldsymbol{\xi}$ donné peut s'écrire comme suit :

$$p(\boldsymbol{\chi}|\,\boldsymbol{\xi}) = \frac{1}{\pi^{NT}\det(\boldsymbol{\Sigma})}e^{-(\boldsymbol{\chi}-\boldsymbol{\mu})^H\boldsymbol{\Sigma}^{-1}(\boldsymbol{\chi}-\boldsymbol{\mu})}, \tag{2.28}$$

où $\boldsymbol{\Sigma}$ et $\boldsymbol{\mu}$ représentent la matrice de covariance et la moyenne de $\boldsymbol{\chi}$, respectivement.

Après calcul et inversion de la matrice d'information de Fisher (voir [EBRM10]), nous avons obtenu les résultats suivants :

– BCR pour le modèle déterministe

Premièrement, pour le modèle déterministe avec $s(t) = \alpha(t)e^{j(2\pi f_0 t+\psi(t))}$ qui représente le signal émis pour une fréquence porteuse valant f_0 et $\alpha(t)$, $\psi(t)$ représentent l'amplitude réelle et la phase de la source, respectivement. Notons $\boldsymbol{\psi} = [\psi(1)\dots\psi(T)]^T$ et $\boldsymbol{\alpha} = [\alpha(1)\dots\alpha(T)]^T$ où les vecteurs de paramètres inconnus sont $\boldsymbol{\xi} = [\omega\ \phi\ \boldsymbol{\psi}^T\ \boldsymbol{\alpha}^T\ \sigma^2]^T$ (si on s'intéresse aux angles électriques) ou $\boldsymbol{\kappa} = [\theta\ r\ \boldsymbol{\psi}^T\ \boldsymbol{\alpha}^T\ \sigma^2]^T$ (si on s'intéresse aux paramètres physiques). Ainsi on

aura $\mathbf{\Sigma} = \sigma^2\mathbf{I}_{NT}$ et $\boldsymbol{\mu} = [s(1)\mathbf{a}^T(\omega,\phi)\dots s(T)\mathbf{a}^T(\omega,\phi)]^T$. Les expressions non matricielles de la matrice $\mathbf{BCRD}(\boldsymbol{\xi})$ correspondant aux angles électriques, pour $N \geq 3$, sont données par

$$\begin{aligned}
\mathrm{BCRD}(\omega) &= \frac{6(2N-1)(8N-11)}{\mathrm{D}_{RSB}(N^2-1)N(N^2-4)}, && (2.29)\\
\mathrm{BCRD}(\phi) &= \frac{90}{\mathrm{D}_{RSB}(N^2-1)N(N^2-4)}, && (2.30)\\
\mathrm{BCRD}\left(\psi(t)\right) &= \frac{1}{2\alpha^2(t)\mathrm{D}_{RSB}}\frac{N^4-31N^3+48N^2-26N+2}{N^2(N+1)(N^2-4)}, \quad \forall t && (2.31)\\
\mathrm{BCRD}\left(\alpha(t)\right) &= \frac{\sigma^2}{2N} \quad \forall t, && (2.32)
\end{aligned}$$

et

$$\mathrm{BCRD}\left(\sigma^2\right) = \frac{\sigma^4}{NT}. \qquad (2.33)$$

Et les termes croisés sont donnés par $\mathrm{BCRD}\left(\omega,\phi\right) = \mathrm{BCRD}\left(\phi,\omega\right) = -\frac{90}{\mathrm{D}_{RSB}N(N^2-4)(N+1)}$.

Les expressions non matricielles de la BCRD pour une source située dans le champ proche, avec $N \geq 3$ et $\theta \neq \frac{\pi}{2}$, sont données par

$$\begin{aligned}
\mathrm{BCRD}(\theta) &= \frac{3\lambda^2}{2\mathrm{D}_{RSB}d^2\pi^2\cos^2(\theta)}\frac{(8N-11)(2N-1)}{(N^2-1)N(N^2-4)}, && (2.34)\\
\mathrm{BCRD}(r) &= \frac{6r^2\lambda^2}{\mathrm{D}_{RSB}\pi^2d^4} \times \frac{15r^2+30drp_1(N)\sin(\theta)+d^2p_2(N)\sin^2(\theta)}{p_3(N)\cos^4(\theta)}, &&
\end{aligned}$$

où $p_1\left(N\right) = N-1$, $p_2(N) = (8N-11)(2N-1)$ et $p_3(N) = N(N^2-1)(N^2-4)$.

Notons que, bien sûr, $\mathbf{BCRD}\left(\psi\right)$, $\mathbf{BCRD}\left(\alpha\right)$ et $\mathrm{BCRD}\left(\sigma^2\right)$ restent inchangées. Et les termes croisés entre θ et r s'expriment désormais comme suit

$$\mathrm{BCRD}\left(\theta,r\right) = \mathrm{BCRD}\left(r,\theta\right) = -\frac{3\lambda^2 r}{\mathrm{D}_{RSB}\pi^2d^3}\frac{15rp_1(N)+dp_2(N)\sin(\theta)}{p_3(N)\cos^3(\theta)}. \qquad (2.35)$$

– BCR pour le modèle aléatoire

Concernant le calcul de la BCRS pour des signaux sources supposés complexes circulaires Gaussiens (de moyenne nulle et de variance σ_s^2) indépendants du bruit, les vecteurs de paramètres inconnus sont donnés par $\boldsymbol{\rho} = [\omega\ \phi\ \sigma_s^2\ \sigma^2]^T$ (si on s'intéresse aux angles électriques) ou $\boldsymbol{\vartheta} = [\theta\ r\ \sigma_s^2\ \sigma^2]^T$ (si on s'intéresse aux paramètres physiques). De ce fait, $\boldsymbol{y}(t)|\boldsymbol{\rho} \sim \mathcal{CN}(\mathbf{0},\mathbf{\Sigma})\ \forall t = 1,\dots,T$, où la matrice de covariance est donnée par $\mathbf{\Sigma} = \sigma_s^2\mathbf{a}(\omega,\phi)\mathbf{a}^H(\omega,\phi) + \sigma^2\mathbf{I}_N$. Les expressions non matricielles de $\mathbf{BCRS}(\boldsymbol{\rho})$ correspondant aux angles électriques, pour $N \geq 3$, sont données par

$$\begin{aligned}
\mathrm{BCRS}(\omega) &= \left(1+\frac{1}{\mathrm{S}_{RSB}\ N}\right)\frac{6(2N-1)(8N-11)}{\mathrm{S}_{RSB}\ T(N^2-1)N(N^2-4)}, && (2.36)\\
\mathrm{BCRS}(\phi) &= \left(1+\frac{1}{\mathrm{S}_{RSB}\ N}\right)\frac{90}{\mathrm{S}_{RSB}\ T(N^2-1)N(N^2-4)}. && (2.37)
\end{aligned}$$

De plus, les termes croisés sont donnés par

$$[\mathbf{BCRS}(\boldsymbol{\rho})]_{1,2} = [\mathbf{BCRS}(\boldsymbol{\rho})]_{2,1} = -\left(1+\frac{1}{\mathrm{S}_{RSB}\ N}\right)\frac{90}{\mathrm{S}_{RSB}\ TN(N^2-4)(N+1)}. \qquad (2.38)$$

Les expressions non matricielles de $\mathbf{BCRS}(\boldsymbol{\vartheta})$ pour une source située dans le champ proche, avec $N \geq 3$ et $\theta \neq \frac{\pi}{2}$, sont données par

$$\mathrm{BCRS}(\theta) = \left(1 + \frac{1}{\mathrm{S}_{RSB}\ N}\right) \frac{3\lambda^2}{2\mathrm{S}_{RSB}\ Td^2\pi^2\cos^2(\theta)} \frac{(8N-11)(2N-1)}{N(N^2-1)(N^2-4)}, \tag{2.39}$$

$$\mathrm{BCRS}(r) = \left(1 + \frac{1}{\mathrm{S}_{RSB}\ N}\right) \frac{6r^2\lambda^2}{\mathrm{S}_{RSB}\ T\pi^2 d^4} \frac{15r^2 + 30dr(N-1)\sin(\theta) + d^2(8N-11)(2N-1)\sin^2(\theta)}{N^2(N^2-1)(N^2-4)\cos^4(\theta)}. \tag{2.40}$$

Les termes croisés entre θ et r sont donnés par

$$\begin{aligned}[\mathbf{BCRS}(\boldsymbol{\vartheta})]_{1,2} &= [\mathbf{BCRS}(\boldsymbol{\vartheta})]_{2,1} \\ &= -\left(1 + \frac{1}{\mathrm{S}_{RSB}\ N}\right) \frac{3\lambda^2 r}{\mathrm{S}_{RSB}\ T\pi^2 d^3} \frac{15r(N-1) + d(8N-11)(2N-1)\sin(\theta)}{N(N^2-1)(N^2-4)\cos^3(\theta)}.\end{aligned} \tag{2.41}$$

Ces expressions analytiques nous ont permis de remarquer que :
- La BCRD et la BCRS sont invariantes par rapport à la phase du signal source.
- Comme dans le cas du champ lointain, BCRD(θ) et BCRS(θ) ne dépendent que de l'azimut selon $1/\cos^2(\theta)$. De ce fait, l'ALU n'est pas une antenne isotrope pour une source située en champ proche.
- BCRD(r) et BCRS(r) dépendent à la fois de l'azimut et de la distance. Pour $\lambda, r \propto d$, la dépendance par rapport à la distance est quant à elle de $O(r^2)$, ce qui veut dire que l'estimation s'améliore quand la source se rapproche de l'antenne (cependant il faut respecter la contrainte de la région de Fresnel 2.24). La dépendance en la distance par rapport à l'azimut est $1/\cos^4(\theta)$. Si θ est proche de $\pi/2$, BCR(r) tend vers l'infini et cette convergence est plus rapide que celle de BCR(θ).
- Pour un nombre suffisant de capteurs, BCRD(θ), BCRD(r), BCRS(θ) et BCRS(r) sont de l'ordre de $O(1/N^3)$.
- Pour $\lambda \propto d$, BCRD(θ) et BCRS(θ) sont indépendantes de la fréquence porteuse f_0. Ce qui n'est pas le cas de BCRD(r) et de BCRS(r).
- Pour un grand nombre de capteurs et une distance entre capteurs fixée, BCRD(θ) et BCRS(θ) exprimées dans le champ proche tendent vers les BCR exprimées dans le champ lointain [SN90b]. Cette dernière remarque est en adéquation avec l'intuition car, du fait de la contrainte (2.24) un grand nombre de capteurs implique une grande distance entre la source et l'antenne (tout en restant dans le région de Fresnel.)
- Les expressions de la BCRD montrent que les paramètres physiques d'intérêt sont fortement couplés vu que BCRD (θ, r) est $O(1/N^3)$ comme BCRD(θ) et BCRD(r). Les mêmes remarques s'appliquent pour la BCRS.

2.3.1.1.2 Processus non-gaussien Dans le contexte de la detection radar, l'estimation de la matrice de covariance du clutter est la pierre angulaire permettant de réaliser des detecteurs optimaux. Bien que le cas d'un clutter gaussien ait été intensivement étudié, les nouvelles avancées technologiques ont montré que les modèles de clutter non-gaussien doivent être pris en compte. Parmi ces modèles, les vecteurs aléatoires sphériquement invariants sont particulièrement interessants puisqu'ils incluent le modèle de clutter K-distribué connu pour être à même de bien représenter les données réelles. Plus précisement, l'estimateur au sens du maximum de vraisemblance de la matrice de covariance d'un clutter K-distribué a déjà été proposé dans la littérature. Avec Fréderic Pascal, SONDRA/Supelec, nous avons proposé une analyse statistique

de cet estimateur au travers de sa consistence et son caractère non-biaisé pour un nombre fini d'échantillons. Afin, de montrer l'efficacité de cet estimateur, nous avons également calculé une forme analytique de la borne ce Cramér-Rao. Voir l'**annexe D** pour les détails.

2.3.1.2 Applications indirectes

2.3.1.2.1 Convergence de l'algorithme EM Avec Cédric Herzet, INRIA, nous avons étudié la vitesse de convergence moyenne de l'algorithme EM. Nous avons montré que le comportement asymptotique (en termes de nombre d'observations) de cet algorithme peut être décrit en fonction des bornes de Cramér-Rao associées aux jeu de données complétes et incomplétes et que la borne de Cramér-Rao associée aux données complétes n'est autre que la borne de Cramér-Rao modifiée. Les détails sont donnés en **annexe A**.

2.3.1.2.2 Seuils de résolution limites Le seuil statistique de résolution limite (SRL), aussi nommé pouvoir séparateur, c'est-à-dire, la distance minimale entre deux signaux permettant une correcte séparation/estimation des paramètres d'intérêt, est un sujet d'une importance croissante et qui vise diverses applications comme le radar, le sonar, le traitement d'images, l'analyse spectrale, etc. En traitement d'antenne, pour la formation de voies, le seuil de résolution ne dépend que de l'ouverture de l'antenne [Ste76]. Ainsi, on peut citer la résolution de Fourier et de Rayleigh données respectivement par la largeur du lobe principal de la fonction de directivité et sa largeur à 3 dB [Mar98, Abe06], respectivement. Avec la venue des méthodes à haute résolution, les seuils de la résolution définis par Fourier et Rayleigh ont été repoussés, d'où la nécessité d'introduire de nouveaux critères du seuil statistique de résolution limite.

Dans la littérature on définit/calcule le SRL principalement selon trois familles de critères [Cox73, KB86, Lee92, Lee94, SD95, Dil98, Smi98, SM05a, Smi05, DA06, LN07, FLV08, AW08, AD08, KG09, KBRM11b, KBRM11a, VEB$^+$11, EBRM11a, EBRM11b] :

1. La première famille est basée sur le pseudo-spectre des algorithmes d'estimation (voir figure 2.4.) Si nous supposons, par exemple, que deux signaux sont paramétrés par les directions d'arrivée (DDA) θ_1 et θ_2, alors, le critère de Cox [Cox73] stipule que *les deux signaux sont résolus si les moyennes des valeurs du pseudo-spectre aux points θ_1 et θ_2 sont inférieures à la moyenne de la valeur du pseudo-spectre au point $\frac{\theta_1+\theta_2}{2}$*. Un second critère, basé lui aussi sur le pseudo-spectre, a été proposé par Sharman et Durrani [SD95] et stipule que *les deux signaux sont résolus si la dérivée seconde de la moyenne du pseudo-spectre au point $\frac{\theta_1+\theta_2}{2}$ est négative.*

   On remarque bien que cette famille de critères est spécifique aux algorithmes utilisés. Pour plus d'exemples, le lecteur pourra se référer aux références suivantes [Cox73, KB86, SD95, AD08]. Dans la suite, nous présenterons deux autres familles de critères valables pour tout type d'algorithme.
2. La seconde famille est basée sur un test d'hypothèses binaires [SM05a, LN07, AW08]. L'idée principale consiste à utiliser ces tests d'hypothèses pour décider si un ou deux signaux sont présents. Le but est alors de relier la distance minimale (entre les deux signaux pour un rapport signal à bruit donné) à la probabilité de fausse alarme P_{fa} et/ou à la probabilité de détection P_d. Ainsi, dans [SM05a], Sharman et Milanfar ont considéré le problème du calcul du SRL en analyse spectrale en utilisant le test du rapport de vraisemblance (TRV). Par conséquent, les auteurs ont calculé l'expression du SRL pour une P_{fa} et une P_d données. Dans [LN07], Liu et Nehorai ont défini le seuil de résolution limite angulaire (c'est-à-dire, le SRL par rapport aux DDA) en utilisant le TRV dans sa forme asymptotique en terme de nombre d'observations. A cet effet, les auteurs ont relié l'expression du SRL à la borne

de Cramér-Rao. Enfin, on peut trouver dans [AW08] le calcul du SRL par rapport aux fréquences pour des sinusoïdes complexes en utilisant, cette fois ci, l'approche Bayésienne.

3. La troisième famille est basée sur la précision d'estimation des paramètres en terme de variance [Lee92,LL93,Lee94,Smi98,Dil98,Smi05,EBRM10,KBRM11b,KBRM11a]. Puisque la BCR est une borne inférieure de l'erreur quadratique moyenne, elle exprime ainsi les performances ultimes en terme d'estimation paramétrique. Par conséquent, elle peut être utilisée pour définir/obtenir le SRL. De ce fait, on distingue deux critères *intuitifs* du SRL basés sur la BCR :

   i) Le premier critère a été introduit par Lee en 1992 [Lee92]. Dans un contexte de traitement d'antenne, il stipule que *deux signaux sont correctement résolus par rapport aux DDA, si l'écart type maximal est inférieur à au moins deux fois la différence entre* θ_1 *et* θ_2. Ainsi, et sous certaines conditions de régularité [LC03b], les écarts types $\sigma_{\hat{\theta}_1}$ et $\sigma_{\hat{\theta}_2}$, d'un estimateur non biaisé $\hat{\boldsymbol{\theta}} = [\hat{\theta}_1\ \hat{\theta}_2]^T$ peuvent être approximés par $\sqrt{\mathrm{BCR}(\theta_1)}$ et $\sqrt{\mathrm{BCR}(\theta_2)}$, respectivement. Par conséquent, le SRL est, selon le critère de Lee, égal à $2\max\left\{\sqrt{\mathrm{BCR}(\theta_1)}, \sqrt{\mathrm{BCR}(\theta_2)}\right\}$. Pour des applications se basant sur le critère de Lee, le lecteur pourra se référer à [Lee92, Lee94, Dil98].

   ii) On peut noter que le couplage entre les paramètres est ignoré par ce dernier critère (c'est-à-dire, absence du terme croisé dans la matrice de Cramèr-Rao $\mathrm{BCR}(\theta_1, \theta_2)$). C'est pour cela que Smith [Smi98, Smi05] a introduit un critère qui tient compte de ce couplage entre les paramètres d'intérêt. Ce critère est donné comme suit : *deux signaux sont résolus par rapport aux DDA si la différence entre les DDA,* δ*, est plus grande que l'écart type de la différence de ces DDA*. L'écart type peut être approximé par la BCR (sous certaines conditions de régularité.) Par conséquent, le SRL au sens de Smith peut être défini comme étant $\delta = |\theta_1 - \theta_2|$ pour lequel l'inégalité suivante

   $$\delta < \sqrt{\mathrm{BCR}\,(\delta)} \tag{2.42}$$

   est atteinte. Par conséquent, le SRL est donné par la solution de l'équation suivante

   $$\delta^2 = \mathrm{BCR}\,(\delta)\,. \tag{2.43}$$

   Dans [Smi98, Smi05], Smith a calculé le SRL (en terme des DDA) pour deux signaux modélisés par des pôles complexes. Dans [DA06], Delmas et Abeida ont calculé le SRL en terme de DDA suivant le critère de Smith pour des signaux sources discrets modulés en BPSK (*binary phase-shift keying*), QPSK (*quadrature phase-shift keying*) et MSK (*minimum-shift keying.*)

On notera que le SRL basé sur la précision de l'estimation (c'est-à-dire, basé sur la BCR) est un concept intuitif. Dans [LN07], les auteurs ont relié le SRL basé sur un test d'hypothèses à celui basé sur la BCR (plus particulièrement celui basé sur le critère de Smith). Ainsi, et pour un grand nombre d'observations, les auteurs ont montré que le SRL basé sur un test d'hypothèses binaire peut être écrit comme la solution de l'équation suivante [LN07, eq. 9]

$$\delta^2 = \lambda\ \mathrm{BCR}\,(\delta)\,, \tag{2.44}$$

où λ est un facteur de translation exprimé en fonction de P_{fa} et P_d.

A noter qu'il existe deux autres critères moins utilisés dans la littérature. Le premier est basé sur le critère d'information d'Akaike (CIA). Dans [SSS95] les auteurs ont relié la fonction

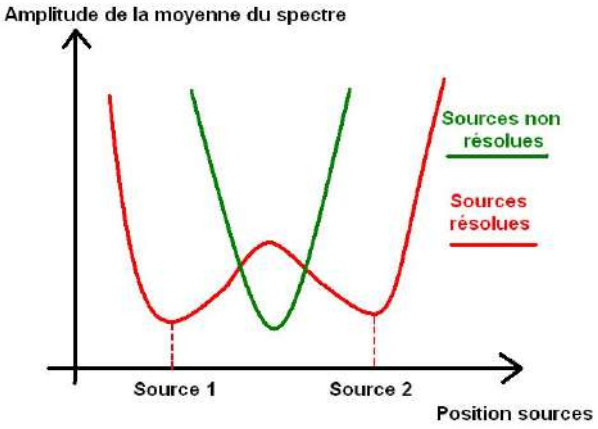


FIGURE 2.4 – Illustration du cas *sources résolues* et *sources non résolues* à l'aide d'un pseudo-spectre d'algorithme d'estimation.

relative au CIA à la séparation δ. Un autre critère est basé sur l'application du lemme de Stein permettant ainsi de relier la distance de Kullback-Leibler à la séparation δ [VEB+11]

Dans le cadre de la thèse de Mohammed Nabil El Korso, nous avons caractérisé le SRL pour la localisation de sources polarisées et pour la localisation de sources en présence des interférences qui sont des applications pour lesquelles aucun résultat n'était disponible dans la littérature. Lors du calcul du SRL dans le contexte champ proche (c'est-à-dire, deux paramètres d'intérêt par source qui sont la distance et la DDA), nous avons constaté que tous les critères susmentionnés ont été introduits dans le cas d'un seul paramètre d'intérêt par signal. Toutefois, dans de nombreux problèmes d'estimation, les signaux sont paramétrés par plus d'un paramètre d'intérêt par signal. On peut citer les problèmes liés au cas Multiple-Input Multiple-Output (MIMO) [God97, THL+01, GS05, STWT06, NS09] ou dans le cadre général des problèmes d'harmoniques multidimensionnelles [JStB01, MSPM04], etc. A cet effet nous avons proposé un critère du SRL pour des signaux multidimensionnels, nommé SRLM (voir **annexe F**). Ainsi que le calcul du SRLM pour les modèles suivants : modèle harmonique multidimensionnel, le radar MIMO mais aussi la localisation de sources en champ proche.

Nous avons étudié deux exemples applicatifs du calcul du SRL pour des signaux monodimensionels, c'est-à-dire, des signaux comportant un seul paramètre d'intérêt par source.

– Seuil de résolution limite pour la localisation de sources polarisées

Considérons une antenne linéaire uniforme de type COLD (cocentered orthogonal loop and dipole array) notée ALU-COLD composée de N capteurs ; où chaque capteur est formé d'une boucle et d'un dipôle. Cette ALU-COLD reçoit deux signaux provenant de deux sources émettrices. Le signal observé par le $\ell^{ème}$ capteur à la $t^{ème}$ observation est donné par [LC91, LSZ96]

$$\boldsymbol{y}_\ell(t) = \begin{bmatrix} y_{\rm b}(t) & y_{\rm d}(t) \end{bmatrix}^T = \sum_{m=1}^{2} \alpha_m(t)\mathbf{u}_m e^{j\ell\omega_m} + \boldsymbol{v}_\ell(t), \tag{2.45}$$

où $\ell = 0 \ldots N-1$ et $t = 1 \ldots T$ avec T représentant le nombre d'observations. Notons $\omega_m = \frac{2\pi}{\lambda} d \sin(\theta_m)$ l'angle électrique où θ_m, d et λ représentent l'azimut de la $m^{ème}$ source, l'espacement inter-capteur et la longueur d'onde, respectivement. Les deux signaux sources sont modélisés par $\alpha_m(t) = a_m e^{j(2\pi f_0 t + \phi_m(t))}$ où a_m est l'amplitude du signal (non nulle), $\phi_m(t)$ la phase du signal et f_0 la fréquence porteuse du signal. Le bruit additif sera noté $\boldsymbol{v}_\ell(t) = \begin{bmatrix} v_{\rm b}(t) & v_{\rm d}(t) \end{bmatrix}^T$. Le vecteur

de polarisation $\mathbf{u}_m$ est donné par

$$\mathbf{u}_m = \begin{bmatrix} \frac{2j\pi A_{sl}}{\lambda}\cos(\rho_m) \\ -L_{sd}\sin(\rho_m)e^{j\psi_m} \end{bmatrix}, \tag{2.46}$$

où $\rho_m \in [0, \pi/2]$ et $\psi_m \in [-\pi, \pi]$ sont les paramètres de polarisation.

Par conséquent, la forme vectorielle du modèle des observations à la $t^{ème}$ observation est donnée par :

$$\boldsymbol{y}(t) = \begin{bmatrix} \boldsymbol{y}_0^T(t) & \dots & \boldsymbol{y}_{N-1}^T(t) \end{bmatrix}^T = \sum_{m=1}^{2} \mathbf{A}_m(t)\mathbf{a}_m + \begin{bmatrix} \boldsymbol{v}_0^T(t) & \dots & \boldsymbol{v}_{N-1}^T(t) \end{bmatrix}^T, \tag{2.47}$$

où $\mathbf{A}_m(t) = \mathbf{I}_L \otimes (\alpha_m(t)\mathbf{u}_m)$. Le vecteur directionnel est défini par $\mathbf{a}_m = \begin{bmatrix} 1 & e^{j\omega_m} & \dots & e^{j(N-1)\omega_m} \end{bmatrix}^T$. Dans la suite, nous ferons les hypothèses suivantes :

- Le bruit est supposé complexe circulaire blanc Gaussien de moyenne nulle et de variance inconnue σ^2.
- Les signaux sources sont supposés connus et déterministes [LC93, LHSV95, CM97]. Le vecteur de paramètres inconnus est alors donné par $\boldsymbol{\xi} = [\omega_1\ \omega_2\ \sigma^2]^T$.
- Afin de simplifier les calculs, et sans perte de généralité, on supposera $L_{sd} = \frac{2\pi A_{sl}}{\lambda} = 1$ [LSZ96] et $w_1 > w_2$.

Après calcul puis inversion de la MIF, on trouve :

$$\mathrm{BCR}(\omega_1) = \frac{\sigma^2}{2N}\frac{a_2^2\alpha}{a_1^2a_2^2\alpha^2 - \Re^2\{r_T\mathbf{u}_1^H\mathbf{u}_2\eta\}}, \tag{2.48}$$

$$\mathrm{BCR}(\omega_2) = \frac{\sigma^2}{2N}\frac{a_1^2\alpha}{a_1^2a_2^2\alpha^2 - \Re^2\{r_T\mathbf{u}_1^H\mathbf{u}_2\eta\}}, \tag{2.49}$$

$$\mathrm{BCR}(\omega_1,\omega_2) = -\frac{\sigma^2}{2N}\frac{\Re\{r_T\mathbf{u}_1^H\mathbf{u}_2\eta\}}{a_1^2a_2^2\alpha^2 - \Re^2\{r_T\mathbf{u}_1^H\mathbf{u}_2\eta\}}. \tag{2.50}$$

où $\alpha = \frac{1}{6}(N-1)N(2N-1)$, $r_T = \frac{1}{T}\sum_{t=1}^{T}\alpha_1^*(t)\alpha_2(t)$ et

$$\eta = \sum_{\ell=0}^{L-1}\ell^2 e^{-j(\omega_1-\omega_2)\ell} = \sum_{\ell=0}^{L-1}\ell^2 e^{-j\mathrm{sgn}(\omega_1-\omega_2)\delta_\omega^{(\mathrm{COLD})}\ell}, \tag{2.51}$$

avec $\delta_\omega^{(\mathrm{COLD})} = \omega_1 - \omega_2$.

Notons $\delta_\omega^{(\mathrm{COLD})}$ le SRL associé au modèle (2.47) qui est donné par le critère de Smith [Smi05] comme suit :

$$\delta_\omega^{(\mathrm{COLD})} = \sqrt{\mathrm{BCR}(\delta_\omega^{(\mathrm{COLD})})} \Longleftrightarrow f(\delta_\omega^{(\mathrm{COLD})}) = \left(a_1^2 + a_2^2\right)\alpha, \tag{2.52}$$

où dans notre cas $f(\delta_\omega^{(\mathrm{COLD})}) = \frac{2}{\sigma^2T^2}\left(a_1^2a_2^2\alpha^2 - \Re^2\{r_T\mathbf{u}_1^H\mathbf{u}_2\eta\}\right)\left(\left(\delta_\omega^{(\mathrm{COLD})}\right)^2 + 2\mathrm{BCR}(\omega_1,\omega_2)\right)$ et

$$\mathrm{BCR}(\delta_\omega^{(\mathrm{COLD})}) = \mathrm{BCR}(\omega_1) + \mathrm{BCR}(\omega_2) - 2\mathrm{BCR}(\omega_1,\omega_2). \tag{2.53}$$

Par conséquent, en utilisant (2.48-2.50) et après calcul on obtient :

i) Le SRL, solution de l'équation implicite (2.52), pour des signaux sources orthogonaux (c'est-à-dire, $r_T = \frac{1}{T}\sum_{t=1}^{T}\alpha_1^*(t)\alpha_2(t) = 0$ [LC93]), est donné par

$$\delta_\omega^{(\mathrm{COLD-O})} = \frac{\sigma}{\sqrt{2T\alpha}}\sqrt{\frac{a_1^2 + a_2^2}{a_1^2a_2^2}}. \tag{2.54}$$

Enfin, dans le cas de deux signaux sources ayant la même puissance (c'est-à-dire, $a_1 = a_2 = a$), on obtient

$$\delta_\omega^{(\text{COLD}-\text{O})} = \frac{1}{\sqrt{T\alpha \text{RSB}}}, \tag{2.55}$$

où $\text{RSB} = a^2/\sigma^2$. Ce résultat est qualitativement équivalent à celui trouvé dans [DA06, AW08] pour des sources non polarisées.

ii) Le SRL, solution implicite de l'équation (2.52) pour des signaux sources non-orthogonaux (c'est-à-dire, $r_T \neq 0$), est donné par :

$$\delta_\omega^{(\text{COLD})} = \alpha \sqrt{\frac{a_1^2 a_2^2 - \Re^2\{r_T \mathbf{u}_1^H \mathbf{u}_2\}}{2\beta \Im\{r_T \mathbf{u}_1^H \mathbf{u}_2\}} \left(1 - \sqrt{1 - \frac{2\sigma^2 \beta \Im\{r_T \mathbf{u}_1^H \mathbf{u}_2\}}{\alpha T} \frac{\left(\left(a_1^2 + a_2^2\right) + 2\Re\{r_T \mathbf{u}_1^H \mathbf{u}_2\}\right)}{(\Re^2\{r_T \mathbf{u}_1^H \mathbf{u}_2\} - a_1^2 a_2^2)^2}}\right)}, \tag{2.56}$$

où $\beta = \sum_{\ell=0}^{N-1} \ell^3 = \frac{1}{4}(N-1)^2 N^2$. Pour un grand nombre d'observations (voir, Fig. 2.5), on peut exprimer le SRL par

$$\delta_\omega^{(\text{COLD})} = \frac{\sigma}{\sqrt{2T\alpha}} \sqrt{\frac{a_1^2 + a_2^2 + 2\Re\{r_T \mathbf{u}_1^H \mathbf{u}_2\}}{a_1^2 a_2^2 - \Re^2\{r_T \mathbf{u}_1^H \mathbf{u}_2\}}}. \tag{2.57}$$

Remarquons, que dans ce cas, le SRL est fonction des paramètres de polarisation. De plus, dans le cas de signaux sources ayant la même puissance, on obtient :

$$\delta_\omega^{(\text{COLD})} = \frac{1}{\sqrt{T\alpha \text{RSB}}} \sqrt{\frac{1 + \Re\{\tilde{r}_T \mathbf{u}_1^H \mathbf{u}_2\}}{1 - \Re^2\{\tilde{r}_T \mathbf{u}_1^H \mathbf{u}_2\}}}, \tag{2.58}$$

où $\tilde{r}_T = \frac{1}{T} \sum_{t=1}^{T} e^{j(\phi_2(t) - \phi_1(t))}$.

Nous avons comparé le SRL pour des sources polarisées à celui des sources non-polarisées. et obtenu le SRL pour des sources non-polarisées :

$$\delta_\omega^{(\text{ALU})} = \frac{\sigma}{\sqrt{2T\alpha}} \sqrt{\frac{a_1^2 + a_2^2 + 2\Re\{r_T\}}{a_1^2 a_2^2 - \Re^2\{r_T\}}}. \tag{2.59}$$

i) On remarque d'après (2.54) et (2.59) que le SRL pour des signaux sources orthogonaux (c'est-à-dire $r_T = 0$ [LC93]) et polarisées est égale au SRL pour des signaux sources orthogonaux non-polarisées. Par conséquent, la polarisation n'apporte aucune amélioration du SRL pour des signaux sources orthogonaux.

ii) Intéressons nous maintenant aux signaux sources non orthogonaux. De (2.57) et (2.59), on peut vérifier que

$$\delta_\omega^{(\text{COLD})} \leq \delta_\omega^{(\text{ALU})} \quad \text{ssi} \quad \Re\{r_T\} \geq \Re\{r_T \mathbf{u}_1^H \mathbf{u}_2\}. \tag{2.60}$$

De plus, comme $\Re\{r_T \mathbf{u}_1^H \mathbf{u}_2\} = \Re\{r_T\}\Re\{\mathbf{u}_1^H \mathbf{u}_2\} - \Im\{r_T\}\Im\{\mathbf{u}_1^H \mathbf{u}_2\}$ et $\Re\{\mathbf{u}_1^H \mathbf{u}_2\} \leq 1$, la condition (2.60) est satisfaite pour $\Im\{r_T\} = 0$ et/ou pour $\Im\{\mathbf{u}_1^H \mathbf{u}_2\} = 0$. Par conséquent, $\delta_\omega^{(\text{COLD})} < \delta_\omega^{(\text{ALU})}$ pour les trois cas suivants :

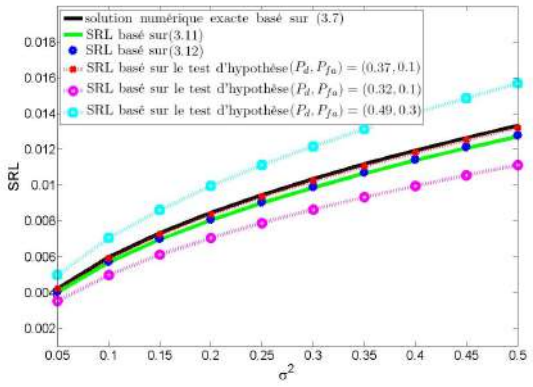


FIGURE 2.5 – Le SRL en fonction de σ^2 pour $T = 100$ observations : Le SRL basé sur l'équation (2.56) et (2.57) est sensiblement égal à la solution numérique exacte basée sur (2.52). Ceci valide nos expressions du SRL. De plus, on note que, par exemple, pour $P_d = 0.37$ et $P_{fa} = 0.1$, le SRL basé sur le critère de Smith est sensiblement égal au SRL calculé en utilisant le test d'hypothèse (voir **Annexe B.1**). Les courbes correspondant à $(P_d, P_{fa}) = (0.49, 0.3)$ et $(P_d, P_{fa}) = (0.32, 0.1)$, nous montrent l'influence du facteur de translation sur le SRL.

C1. si les signaux sources sont réels et positifs, c'est-à-dire, $\Im\{r_T\} = 0$ ou à phase commune, c'est-à-dire, $\phi_1(t) = \phi_2(t), \forall t$.
C2. si $\psi_1 = \psi_2$, c'est-à-dire, $\Im\{\mathbf{u}_1^H \mathbf{u}_2\} = 0$.
C3. si $\rho_1 = 0$ ou $\rho_2 = 0$, c'est-à-dire, $\Im\{\mathbf{u}_1^H \mathbf{u}_2\} = 0$.

Les conditions **C1.**, **C2.** et **C3.** sont des conditions suffisantes pour avoir $\delta_\omega^{(\mathrm{COLD})} < \delta_\omega^{(\mathrm{ALU})}$. Afin d'étudier les autres cas, on a tracé dans la Fig. 2.6 la variable $D(r_L, \mathbf{u}_1^H \mathbf{u}_2) = \Re\{r_T\} - \Re\{r_T \mathbf{u}_1^H \mathbf{u}_2\}$ en fonction des paramètres de polarisation ρ et ψ. On constate que si $D > 0$ alors $\delta_\omega^{(\mathrm{COLD})} < \delta_\omega^{(\mathrm{ALU})}$. La Fig. 2.6 montre qu'en général $\delta_\omega^{(\mathrm{COLD})} < \delta_\omega^{(\mathrm{ALU})}$ et que $\delta_\omega^{(\mathrm{COLD})} > \delta_\omega^{(\mathrm{ALU})}$ seulement pour une petite région (celle qui correspond à la partie inférieure délimitée par le plan horizontal.) Cela signifie, que généralement, le SRL est amélioré grâce aux paramètres de polarisation pour des signaux sources non-orthogonaux.

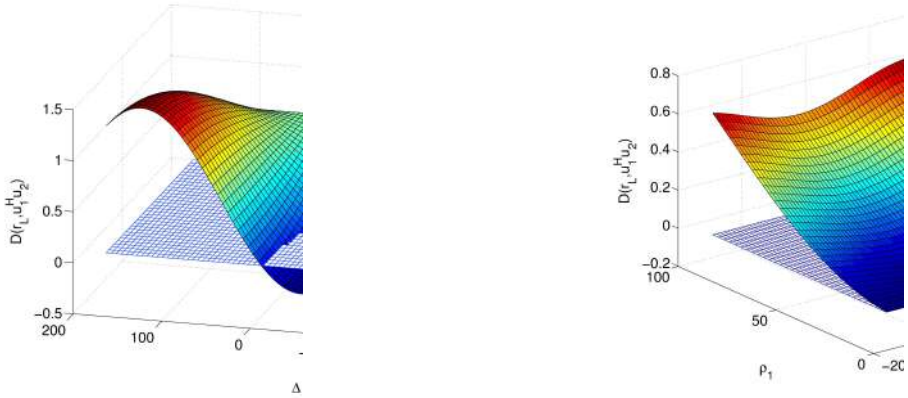


FIGURE 2.6 – $D(r_L, \mathbf{u}_1^H \mathbf{u}_2)$ en fonction des paramètres de polarisation ρ et ψ ; $a_1 = 2$, $a_2 = 3$, $r_T = \frac{1+i}{20}$ avec $N = 20$. (à gauche) $\rho_2 = 85$ deg et (à droite) $\rho_2 = 5$ deg.

– Seuil de résolution limite pour des sources en présence d'interférences

Contrairement aux précédents travaux sur le SRL [SM04, SM05a, LN07, AW08], dans le contexte de la localisation de source en champs lointain, nous considérons un modèle plus riche incluant la présence de sources interférences. En effet, nous considérons le problème du calcul du SRL pour deux sources d'intérêt (notées $\boldsymbol{s}_1$ et $\boldsymbol{s}_2$) noyées dans un sous-espace des interférences engendré par des sources de nuisance (voir, Fig. 2.7).

Considérons une antenne linéaire non uniforme (ALNU) avec N capteurs, qui reçoit un signal émis par M sources se situant dans le champ lointain $\{s_1(t), \ldots, s_M(t)\}$ (M étant supposé connu ou précédemment estimé [Van68]). Le signal observé à la $t^{ème}$ observation par rapport au $n^{ème}$ capteur est donné par [KV96]

$$y_n(t) = \sum_{m=1}^{M} s_m(t) \exp(j\omega_m d_n) + v_n(t), \quad t = 1, \ldots, T, \quad n = 0, \ldots, N-1, \tag{2.61}$$

où T représente le nombre d'observations, $\omega_m = -2\pi \sin(\theta_m)/\nu$ le paramètre d'intérêt de la $m^{ème}$ source avec θ_m et ν qui représentent l'angle d'arrivée et la longueur d'onde. Dans la suite, on notera d_n la distance entre le premier capteur et le $n^{ème}$ capteur. Le bruit additif $v_n(t)$ est supposé complexe circulaire Gaussien de moyenne nulle et de variance σ^2. La forme vectorielle des observations à la $t^{ème}$ observation est donnée par

$$\boldsymbol{y}(t) = \begin{bmatrix} y_0(t) & \ldots & y_{N-1}(t) \end{bmatrix}^T = \begin{bmatrix} \boldsymbol{a}_1 & \ldots & \boldsymbol{a}_M \end{bmatrix} \breve{\boldsymbol{s}}(t) + \boldsymbol{v}(t), \tag{2.62}$$

où $\boldsymbol{v}(t) = [v_0(t) \ldots v_{N-1}(t)]^T$, $\breve{\boldsymbol{s}}(t) = [s_1(t) \ \ldots \ s_M(t)]^T$ et $[\boldsymbol{a}_m]_{n+1} = \exp(j\omega_m d_n)$, $m = 1, 2, \ldots, M$. Par conséquent, le vecteur des observations complet est donné par

$$\boldsymbol{y} \stackrel{\Delta}{=} \begin{bmatrix} \boldsymbol{y}^T(1) & \boldsymbol{y}^T(2) & \ldots & \boldsymbol{y}^T(T) \end{bmatrix}^T. \tag{2.63}$$

Dans la suite nous avons calculé le SRL, δ, dans le contexte de la localisation de sources en présence des interférences (voir, Fig. 2.7). A cet effet, nous supposons que :

– Les deux sources d'intérêt (SDI) sont notées $\boldsymbol{s}_1$ et $\boldsymbol{s}_2$ (avec $\boldsymbol{s}_1 \neq \boldsymbol{s}_2$). Par conséquent, le SRL (c'est-à-dire, la séparation) est défini par $\delta \stackrel{\Delta}{=} \omega_2 - \omega_1$.
– Le sous-espace des interférences (SI) [BS94] est représenté par les $M-2$ sources restantes $\{\boldsymbol{s}_3, \ldots, \boldsymbol{s}_M\}$. Chaque paire de sources est considérée comme largement espacée. Une condition suffisante est que la séparation minimale sur l'ensemble de la combinaison des paires de sources des interférences, notée Δ_ω, doit vérifier $\Delta_\omega > \delta$.

Pour calculer le SRL nous utilisons le critère basé sur le test d'hypothèses. L'hypothèse $\mathcal{H}_0$ représente le cas où les deux SDI sont combinées en un seul signal, alors que l'hypothèse $\mathcal{H}_1$ incarne la situation où les deux SDI sont résolues.

$$\begin{cases} \mathcal{H}_0 : & \delta = 0, \\ \mathcal{H}_1 : & \delta \neq 0. \end{cases} \tag{2.64}$$

La séparation δ est un paramètre inconnu, donc, il est impossible de concevoir un test de détection du type Neyman-Pearson. L'alternative la plus utilisée en traitement du signal est alors l'utilisation du test du rapport de vraisemblance généralisé (TRV) [Kay98] dont la statistique est donnée par :

$$G(\boldsymbol{y}) = \frac{\max_{\delta, \boldsymbol{\rho}_1} p(\boldsymbol{y}|\delta, \boldsymbol{\rho}_1, \mathcal{H}_1)}{\max_{\boldsymbol{\rho}_0} p(\boldsymbol{y}|\boldsymbol{\rho}_0, \mathcal{H}_0)} = \frac{p(\boldsymbol{y}|\hat{\delta}, \hat{\boldsymbol{\rho}}_1, \mathcal{H}_1)}{p(\boldsymbol{y}|\hat{\boldsymbol{\rho}}_0, \mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta', \tag{2.65}$$

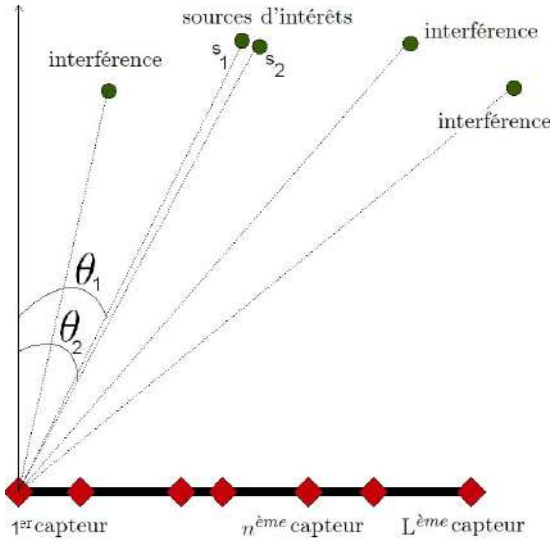


FIGURE 2.7 – Deux SDI proches noyées dans les interférences formées par 3 sources de nuisance.

où $p(\boldsymbol{y}|\boldsymbol{\rho}_0, \mathcal{H}_0)$ et $p(\boldsymbol{y}|\delta, \boldsymbol{\rho}_1, \mathcal{H}_1)$ représentent la densité de probabilité des observations sous $\mathcal{H}_0$ et $\mathcal{H}_1$, respectivement, et où η', $\hat{\delta}$ et $\hat{\boldsymbol{\rho}}_i$ sont le seuil de détection, l'estimation du maximum de vraisemblance de δ sous $\mathcal{H}_1$ et l'estimation du maximum de vraisemblance du vecteur $\boldsymbol{\rho}_i$ (qui contient tous les paramètres de nuisance) sous $\mathcal{H}_i, i = 0, 1$. Malheureusement, la solution analytique relative au test (2.65) par rapport à δ n'existe pas à cause de la non-linéarité du modèle des observations [SM04, Van68, Van02, OVSN93]. De ce fait, nous proposons d'approximer le modèle d'observation en nous basant sur l'hypothèse que δ est proche de zéro (cette hypothèse est soutenue par le fait que les algorithmes à haute résolution ont, asymptotiquement, un pouvoir de résolution infini [Van02].) Par conséquent, le modèle approché (à l'aide d'un développement limité à l'ordre 1 en $\delta = 0$, tel que $\omega_1 = \omega_c - \frac{\delta}{2}$ et $\omega_2 = \omega_c + \frac{\delta}{2}$) s'écrit :

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{s}_+ + \delta \boldsymbol{B}\boldsymbol{s}_- + \boldsymbol{e} + \boldsymbol{v}, \tag{2.66}$$

où $\boldsymbol{e} = \boldsymbol{C}\boldsymbol{s}$. La matrice $\boldsymbol{C}$ supposée connue ou précédemment estimée [Beh90] est donnée par $\boldsymbol{C} = [\boldsymbol{A}_3 \ \ldots \ \boldsymbol{A}_M]$, $\boldsymbol{s} = [\boldsymbol{s}_3^T \ \ldots \ \boldsymbol{s}_M^T]^T$, et

$$\boldsymbol{s}_+ = \boldsymbol{s}_1 + \boldsymbol{s}_2, \tag{2.67}$$

$$\boldsymbol{s}_- = \boldsymbol{s}_2 - \boldsymbol{s}_1, \tag{2.68}$$

avec $\boldsymbol{s}_i = [s_i(1) \ \ldots \ s_i(T)]^T$. $\boldsymbol{d} = [d_0 \ d_1 \ \ldots \ d_{N-1}]^T$ et $\boldsymbol{a}$ représente le vecteur directionnel par rapport au paramètre $\omega_c = \frac{\omega_1+\omega_2}{2}$ supposé connu [LN07] ou précédemment estimé [SM05a] (c'est-à-dire, $[\boldsymbol{a}]_{n+1} \triangleq \exp(j\omega_c d_n)$, $n = 0, \ldots, N-1$). On définit aussi

$$\boldsymbol{A} \triangleq \begin{bmatrix} \boldsymbol{a} & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{a} \end{bmatrix}_{(NT)\times T} = \boldsymbol{I}_T \otimes \boldsymbol{a}, \tag{2.69}$$

$$\boldsymbol{B} \triangleq \frac{j}{2}\boldsymbol{I}_T \otimes \dot{\boldsymbol{a}}, \quad \text{avec} \quad \dot{\boldsymbol{a}} \triangleq \boldsymbol{a} \odot \boldsymbol{d}, \tag{2.70}$$

$$\boldsymbol{A}_m \triangleq \boldsymbol{I}_T \otimes \boldsymbol{a}_m, \quad \text{pour } m = 3, \ldots, M. \tag{2.71}$$

Le modèle (2.66) étant linéaire en δ, nous pouvons alors calculer le SRL.

La relation entre le SRL et le RSBI-RSB requis pour résoudre deux SDI pour différents cas (SDI connue ou inconnue, SI connu ou inconnu avec variance du bruit connue ou inconnue) est représentée dans le tableau 2.1 où on a défini $\boldsymbol{w} = \boldsymbol{B}\boldsymbol{s}_-$ et $\boldsymbol{D} = [\boldsymbol{A}\ \boldsymbol{C}]$ avec le rapport signal à bruit et le rapport signal à bruit plus les interférences défini par RSB $\stackrel{\Delta}{=} \frac{\sum_{m=1}^{2}||\boldsymbol{s}_m||^2}{\sigma^2}$ et RSBI $\stackrel{\Delta}{=} \frac{\sum_{m=1}^{2}||\boldsymbol{s}_m||^2}{\|\boldsymbol{s}\|^2+\sigma^2}$, respectivement. δ_i et λ_i, $i = 1, \ldots, 4$, représentent le SRL et le facteur de translation pour le $i^{ème}$ cas, respectivement. La loi du chi2 centrée avec i degrés de liberté est désignée par χ_i^2 et la loi du F centrée avec i_1 et i_2 degrés de liberté est notée F_{i_1,i_2}. Le paramètre de translation λ_1 est estimé numériquement comme solution de $Q^{-1}_{\chi_1^2}(P_{fa}) = Q^{-1}_{\chi_1^2(\lambda_1)}(P_d)$, où $Q^{-1}_{\chi_1^2}(.)$ est la fonction inverse de $Q_{\chi_1^2}(.)$ qui désigne la surface sous la queue de distribution à droite de la loi χ_1^2. De même on définit λ_2, λ_3 et λ_4 comme solution de $Q^{-1}_{\chi_1^2}(P_{fa}) = Q^{-1}_{\chi_1^2(\lambda_2)}(P_d)$, $Q^{-1}_{\chi^2_{2T}}(P_{fa}) = Q^{-1}_{\chi^2_{2T}(\lambda_3)}(P_d)$, et $Q^{-1}_{F_{2L,2(N-M)T}}(P_{fa}) = Q^{-1}_{2L,2(N-M)T(\lambda_4)}(P_d)$, respectivement.

	SDI	SI	Variance du bruit	RSBI pour $M \geq 2$	RSB pour $M = 2$
Cas 1	connue	connu	connue	$\frac{\|\boldsymbol{s}_1\|^2+\|\boldsymbol{s}_2\|^2}{\|\boldsymbol{s}\|^2+\frac{2}{\lambda_1}\delta_1^2\|\boldsymbol{w}\|^2}$	$\lambda_1\frac{\|\boldsymbol{s}_1\|^2+\|\boldsymbol{s}_2\|^2}{2\delta_1^2\|\boldsymbol{w}\|^2}$
Cas 2	connue	inconnu	connue	$\frac{\|\boldsymbol{s}_1\|^2+\|\boldsymbol{s}_2\|^2}{\|\boldsymbol{s}\|^2+\frac{2}{\lambda_2}\delta_2^2\|\|\boldsymbol{P}^{\perp}_{\boldsymbol{C}}\boldsymbol{w}\|^2}$	$\lambda_2\frac{\|\boldsymbol{s}_1\|^2+\|\boldsymbol{s}_2\|^2}{2\delta_2^2\|\boldsymbol{w}\|^2}$
Cas 3	inconnue	inconnu	connue	$\frac{\|\boldsymbol{s}_1\|^2+\|\boldsymbol{s}_2\|^2}{\|\boldsymbol{s}\|^2+\frac{2}{\lambda_3}\delta_3^2\|\|\boldsymbol{P}^{\perp}_{\boldsymbol{D}}\boldsymbol{w}\|^2}$	$\lambda_3\frac{\|\boldsymbol{s}_1\|^2+\|\boldsymbol{s}_2\|^2}{2\delta_3^2\|\|\boldsymbol{P}^{\perp}_{\boldsymbol{A}}\boldsymbol{w}\|^2}$
Cas 4	inconnue	inconnu	inconnue	$\frac{\|\boldsymbol{s}_1\|^2+\|\boldsymbol{s}_2\|^2}{\|\boldsymbol{s}\|^2+\frac{2}{\lambda_4}\delta_4^2\|\|\boldsymbol{P}^{\perp}_{\boldsymbol{D}}\boldsymbol{w}\|^2}$	$\lambda_4\frac{\|\boldsymbol{s}_1\|^2+\|\boldsymbol{s}_2\|^2}{2\delta_4^2\|\|\boldsymbol{P}^{\perp}_{\boldsymbol{A}}\boldsymbol{w}\|^2}$

TABLE 2.1 – La relation entre le SRL et le RSBI/RSB requis pour résoudre deux SDI.

La Fig. 2.8 représente le RSBI en fonction du SRL pour les différents cas (avec $T = 100$ observations, $\nu = 0.5m$ et $(P_{fa}, P_d) = (0.01, 0.99)$.) Nous remarquons que la différence entre le cas 1 et le cas 2 (10 dB) est due au projecteur orthogonal sur l'espace des interférences $\boldsymbol{P}^{\perp}_{\boldsymbol{C}}$. De même la différence entre le cas 2 et le cas 3 (25 dB) est due principalement au projecteur orthogonal sur l'espace engendré par la direction centrale ω_c et les interférences présentes dans $\boldsymbol{P}^{\perp}_{\boldsymbol{D}}$. Enfin, la différence entre le cas 3 et le cas 4 est minime (0.5 dB) et est seulement due au paramètre de translation λ_4. En conclusion, les différences des RSBIs et donc des SRLs sont principalement dues

- à l'effet des sous espaces des interférences ($\langle\boldsymbol{C}\rangle$ et $\langle\boldsymbol{D}\rangle$),
- et, avec un degré moindre, au facteur de translation λ_i.

Passons à une analyse numérique du RBSI en fonction du SRL. Le nombre d'observations est fixé à $T = 100$ avec $\nu = 0.5m$ et $(P_{fa}, P_d) = (0.01, 0.99)$. Nous constatons que le RSBI (ou par équivalence le SRL) est affecté par :

- *la connaissance a priori des signaux sources :* en effet, on constate que la connaissance a priori des sources a un fort impact sur le SRL évalué approximativement à 40 dB (voir, Fig. 2.9*(à gauche)*),
- *les sources des interférences :* de la Fig. 2.9*(à droite)* nous constatons que les sources des interférences additionnelles n'ont aucun effet si elles sont bien espacées, c'est-à-dire, si $\Delta_\omega \gg \delta$. Par contre si Δ_ω est de l'ordre de δ alors la dégradation du SRL est évaluée à 30 dB.

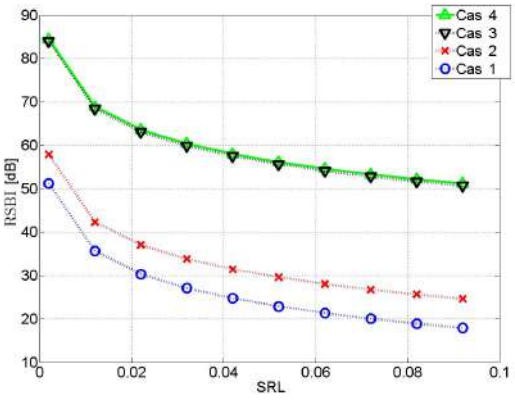


FIGURE 2.8 – RSBI en fonction du SRL pour des sources en champ lointain en présence d'interférences.

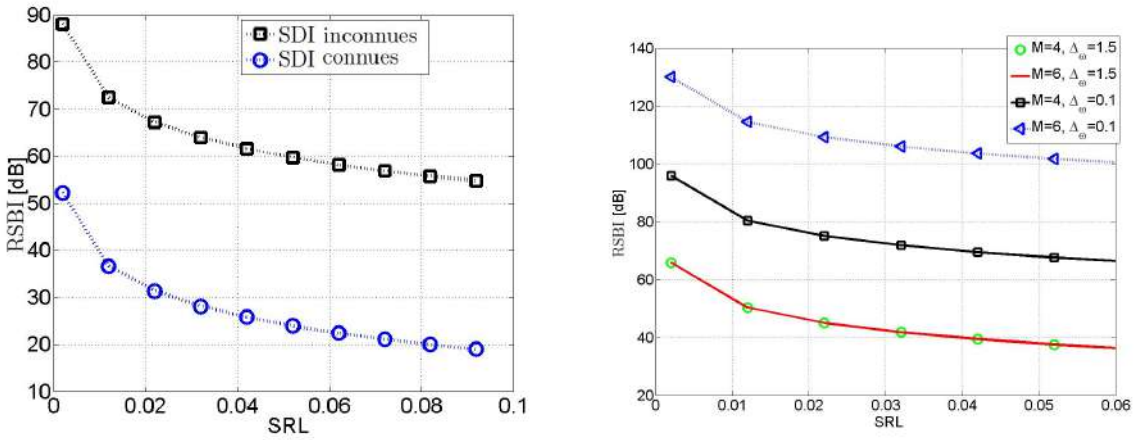


FIGURE 2.9 – *(à gauche)* Le RSBI requis pour résoudre deux SDI connues/inconnues pour une ALU avec $N = 10$ capteurs, $d = \frac{\nu}{2}$ et $M = 4$ avec $\Delta_\omega = 0.75$. *(à droite)* Le RSBI requis pour résoudre deux SDI inconnues pour une ALU avec $N = 10$ capteurs, $d = \frac{\nu}{2}$ et pour différentes valeurs de M et de Δ_ω.

- *L'orthogonalité des sources* améliore aussi le SRL. De la Fig. 2.10 nous constatons que le gain apporté par l'orthogonalité est approximativement égal à 3 dB,
- *la géométrie d'antenne :* dans le tableau 2.2, nous avons présenté certaines géométries d'antenne linéaire pour $N = 4$ capteurs. D'après les simulations de la Fig. 2.11, nous remarquons que l'ouverture d'antenne (c'est-à-dire, l'ajout d'une distance d) produit un gain de 2 dB par rapport au SRL. D'un autre côté, nous constatons que le SRL est sensiblement le même pour différentes géométries d'antennes avec le même nombre de capteur et la même ouverture d'antenne (une différence de seulement 1 dB).
- Seuil de résolution limite pour des signaux multidimensionnels

Après avoir traité le SRL dans le cas monodimensionnel, nous nous sommes intéressés au seuil statistique de résolution limite pour des signaux multidimensionnels (SRLM). Nous rappelons que le seuil de résolution limite basé sur l'approche par la borne de Cramér-Rao n'était introduit que dans le cas monodimensionnel. Pour ce faire, nous avons tout d'abord introduit un critère du SRLM basé sur l'extension du critère de Smith. En deuxième lieu, nous avons montré que

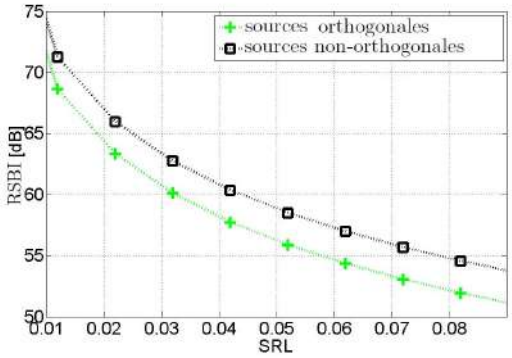


FIGURE 2.10 – Le RSBI requis pour résoudre deux sources du type BPSK inconnues orthogonales/non-orthogonales pour une ALU avec $N = 10$ capteurs, $d = \frac{\nu}{2}$ et $M = 4$.

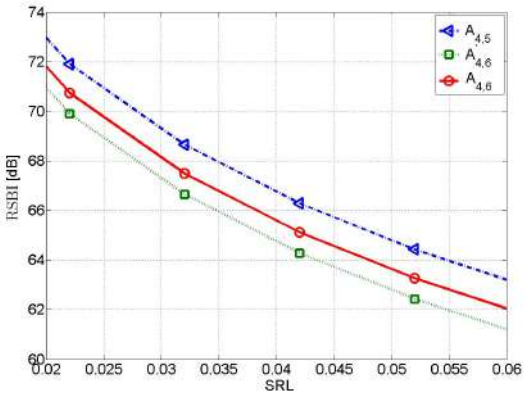


FIGURE 2.11 – Le RSBI requis pour résoudre deux sources connues à l'aide d'une antenne parfaite $A_{4,6}$, une antenne quelconque $A'_{4,6}$ et une antenne à minimum de redondance $A_{4,5}$ décrites au tableau 2.2.

Configuration	Position des capteurs	ouverture	distance redondante	Écart manquant
Antenne parfaite $A_{4,6}$	$[0, 1, 4, 6]$	$6d$	$R = \{\}$	$G = \{\}$
Antenne quelconque $A'_{4,6}$	$[0, 1, 2, 6]$	$6d$	$R = \{1\}$	$G = \{3\}$
Antenne à minimum de redondance $A_{4,5}$	$[0, 1, 2, 5]$	$5d$	$R = \{1\}$	$G = \{\}$

TABLE 2.2 – Caractéristique des différentes géométries d'antennes avec le même nombre de capteurs et des ouvertures différentes. L'antenne parfaite ne contient pas de redondance et aucun écart n'est manquant. La position des capteurs traduit leur coordonnées sur l'axe des abscisses. L'unité, d, correspond à l'écart minimal entre deux capteurs successifs. L'écart entre deux capteurs est donc un multiple de d qui doit être compris entre d et $(L-1)d$. Une distance redondante se traduit par la répétition d'un écart [VH86]. Un écart manquant est dû à l'absence d'un écart entre d et $(L-1)d$ [Mof68]. Une antenne est dite parfaite, si aucun écart n'est manquant et si aucune distance n'est redondante [AGGS96, ASG99, MD01].

le critère proposé est asymptotiquement équivalent (à un facteur de translation près) à un test UPP (uniformément le plus puissant). Enfin, nous avons donné quelques exemples applicatifs (modèle harmonique multidimensionnel, radar MIMO mais aussi le SRLM dans le contexte de la localisation de sources en champ proche).

Dans ce qui suit, nous ne citerons que les principaux résultats. Le lecteur trouvera le détail calculatoire, ainsi qu'une analyse théorique et numérique plus approfondie dans l'**annexe F**.

Dans ce qui suit nous allons utiliser un modèle d'observation généré par deux signaux sources avec P paramètres d'intérêt par signal. Le modèle d'observation est structuré sous forme vectorielle comme suit :

$$\mathbf{y} = \boldsymbol{f}(\boldsymbol{\xi}_1) + \boldsymbol{f}(\boldsymbol{\xi}_2) + \mathbf{v}, \tag{2.72}$$

où $\mathbf{y} \in \mathbb{R}^N$ et $\mathbf{v} \in \mathbb{R}^N$ représentent respectivement le vecteur d'observation [1] et le vecteur bruit avec une densité de probabilité connue. Le vecteur de paramètres inconnus est donné par $\boldsymbol{\xi}_m \subseteq \mathbb{R}^{P+q_m}$, $m = 1, 2$ avec $q_1 + q_2 = Q$ où q_m représente la dimension des paramètres de nuisance par signal. Nous supposerons aussi que le modèle (2.72) est identifiable et que la matrice d'information de Fisher existe et est inversible. Nous pouvons rassembler tous les paramètres dans un vecteur $\tilde{\boldsymbol{\xi}} = \left[\boldsymbol{\xi}_1^T \boldsymbol{\xi}_2^T\right]^T \in \mathbb{R}^{2P+Q}$. En réarrangeant les élements du vecteur $\tilde{\boldsymbol{\xi}}$, nous obtenons $\boldsymbol{\xi} = \left[\boldsymbol{\omega}^T \ \boldsymbol{\rho}^T\right]^T$ où $\boldsymbol{\omega} \in \mathbb{R}^{2P}$ et $\boldsymbol{\rho} \in \mathbb{R}^Q$ représentent, respectivement, le vecteur des paramètres d'intérêt, et le vecteur des paramètres de nuisance.

– Extension du seuil de résolution limite au cas multidimensionnel

Dans ce qui suit nous allons considérer les hypothèses suivantes :

– **A1.** Les éléments du vecteur $\boldsymbol{\omega}$ sont de la même nature, c'est-à-dire, les paramètres d'intérêt ont la même unité de mesure (par exemple des fréquences).

– **A2.** Chaque $p^{ème}$ paramètre, noté $\omega_1^{(p)}$, correspondant à la première source, peut être aussi proche que possible du $p^{ème}$ paramètre, $\omega_2^{(p)}$, correspondant à la deuxième source, mais jamais égal. Cette hypothèse est fréquemment utilisée pour les signaux multidimensionnels, car l'événement $\omega_1^{(p)} = \omega_2^{(p)}$ est considéré comme ayant une probabilité quasi-nulle [GS05, p74].

Sous ces hypothèses nous proposons le critère du SRLM comme suit

Le SRLM, noté δ, pour le modèle (2.72) est donné comme la solution implicite de l'équation suivante

$$\delta^2 = BCR\,(\delta) \tag{2.73}$$

avec

$$\delta = \sum_{p=1}^{P} \delta_p, \tag{2.74}$$

où δ_p représente le SRL dit "local" qui est donné par $\delta_p \triangleq \left|\omega_2^{(p)} - \omega_1^{(p)}\right|$.

Après calcul, nous obtenons le résultat suivant :

Le SRLM pour le modèle (2.72) à P paramètres d'intérêt par signal est donné par δ, qui représente la solution implicite de l'équation suivante :

$$\delta^2 - A_{\text{direct}} - A_{\text{croisé}} = 0,$$

1. Si les observations sont complexes, alors le vecteur d'observation (2.72) sera formé par la concaténation de la partie réelle et de la partie imaginaire des observations complexes. De ce fait, l'étude proposée dans la suite de la section reste valable.

où A_{direct} représente la contribution des termes directs (c'est-à-dire, par rapport au même paramètre p)

$$A_{\text{direct}} = \sum_{p=1}^{P} \left[\text{BCR}(\omega_1^{(p)}) + \text{BCR}(\omega_2^{(p)}) - 2\text{BCR}(\omega_1^{(p)}, \omega_2^{(p)}) \right], \tag{2.75}$$

et où $A_{\text{croisé}}$ représente la contribution des termes croisés (entre le $p^{éme}$ paramètre et le $p'^{éme}$ paramètre) qui est donnée par

$$A_{\text{croisé}} = \sum_{p=1}^{P} \sum_{\substack{p'=1 \\ p' \neq p}}^{P} g_p g_{p'} \left(\text{BCR}(\omega_1^{(p)}, \omega_1^{(p')}) + \text{BCR}(\omega_2^{(p)}, \omega_2^{(p')}) - 2\text{BCR}(\omega_1^{(p)}, \omega_2^{(p')}) \right), \tag{2.76}$$

avec $g_p = \text{sgn}\left(\omega_1^{(p)} - \omega_2^{(p)}\right)$.

Le critère précédemment introduit est un critère intuitif car il est basé sur le critère de Smith. Dans la section suivante, nous allons l'analyser et prouver que ce dernier est asymptotiquement équivalent (à un facteur de translation près) à un test d'hypothèses UPP (uniformément le plus puissant.)

Pour analyser le SRLM nous nous replaçons dans le cadre d'un test d'hypothèses [SM05a, LN07, AW08]. Plus précisément, l'hypothèse $\mathcal{H}_0$ représente le cas où les deux sources d'intérêt (SDI) sont combinées en un seul signal (c'est-à-dire, $\forall p \in [1 \dots P]$, $\omega_1^{(p)} = \omega_2^{(p)}$), alors que l'hypothèse $\mathcal{H}_1$ incarne la situation où les deux SDI sont résolues (c'est-à-dire, $\exists p \in [1 \dots P]$, $\omega_1^{(p)} \neq \omega_2^{(p)}$) :

$$\begin{cases} \mathcal{H}_0 : & \delta_{\text{detection}} = 0, \\ \mathcal{H}_1 : & \delta_{\text{detection}} > 0, \end{cases} \tag{2.77}$$

où $\delta_{\text{detection}}$ représente la distance entre C_1 et C_2 avec $C_q = \left\{ \omega_q^{(1)}, \omega_q^{(2)}, \dots, \omega_q^{(P)} \right\}$, $q = 1, 2$. Par conséquent, la mesure naturelle est celle de Minkowski à l'ordre 1 qui est donnée comme suit :

$$\delta_{\text{detection}} \triangleq \sum_{p=1}^{P} \left| \omega_2^{(p)} - \omega_1^{(p)} \right|. \tag{2.78}$$

Comme dans le cas monodimensionnel, la distance $\delta_{\text{detection}}$ est un paramètre inconnu, donc, il est impossible de concevoir un test de détection du type Neyman-Pearson. L'alternative la plus utilisée en traitement du signal est alors l'utilisation du test du rapport de vraisemblance (TRV) [Kay98] dont la statistique est donnée par :

$$\begin{aligned} L_G(\mathbf{y}) &= \frac{\max_{\delta_{\text{detection}}, \boldsymbol{\rho}_1} \; p(\mathbf{y}|\delta_{\text{detection}}, \boldsymbol{\rho}_1, \mathcal{H}_1)}{\max_{\boldsymbol{\rho}_0} \; p(\mathbf{y}|\boldsymbol{\rho}_0, \mathcal{H}_0)} \\ &= \frac{p(\mathbf{y}|\hat{\delta}_{\text{detection}}, \hat{\boldsymbol{\rho}}_1, \mathcal{H}_1)}{p(\mathbf{y}|\hat{\boldsymbol{\rho}}_0, \mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \varsigma', \end{aligned} \tag{2.79}$$

où $p(\mathbf{y}|\boldsymbol{\rho}_0, \mathcal{H}_0)$ et $p(\mathbf{y}|\delta, \boldsymbol{\rho}_1, \mathcal{H}_1)$ représentent la densité de probabilité des observations sous $\mathcal{H}_0$ et $\mathcal{H}_1$, respectivement, et où ς', $\hat{\delta}_{\text{detection}}$ et $\hat{\boldsymbol{\rho}}_i$ sont le seuil de détection, l'estimation du maximum de vraisemblance de $\delta_{\text{detection}}$ sous $\mathcal{H}_1$ et l'estimation du maximum de vraisemblance du vecteur

$\boldsymbol{\rho}_i$ (qui contient tous les paramètres de nuisance) sous $\mathcal{H}_i, i = 0, 1$. Pour simplifier les calculs, nous considérons la statistique équivalente à (2.79) :

$$T_G(\mathbf{y}) = \text{Ln } L_G(\mathbf{y}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \varsigma = \text{Ln}\varsigma'. \tag{2.80}$$

Malheureusement, la solution analytique relative au test (2.80) n'existe généralement pas [SM04, Van68, Van02, OVSN93]. Par conséquent, nous considérerons le cas asymptotique (en terme d'observations [LN07]). Dans [Kay98, eq (6C.1)], l'auteur a démontré (pour un grand nombre d'observations) que la statistique $T_G(\mathbf{y})$ suivait les lois de probabilité suivantes :

$$T_G(\mathbf{y}) \sim \begin{cases} \chi_1^2 & \text{sous } \mathcal{H}_0 \\ \chi_1^2(\kappa'\left(P_{fa}, P_d\right)) & \text{sous } \mathcal{H}_1 \end{cases} \tag{2.81}$$

où P_{fa} et P_d représentent, respectivement, la probabilité de fausse alarme et la probabilité de détection par rapport au test (2.77). Supposons que $\text{BCR}(\delta_{\text{detection}})$ existe (voir les hypothèses **A.1** et **A.2**), le paramètre de décentrage $\kappa'\left(P_{fa}, P_d\right)$ [Kay98, p.239] est alors donné par

$$\kappa'\left(P_{fa}, P_d\right) = \delta^2_{\text{detection}} \left(\text{BCR}(\delta_{\text{detection}})\right)^{-1}. \tag{2.82}$$

D'un autre côté, le paramètre de décentrage peut être évalué à l'aide de P_{fa} et P_d [Sch91, LN07] comme solution de l'équation suivante :

$$Q^{-1}_{\chi_1^2}(P_{fa}) = Q^{-1}_{\chi_1^2(\kappa'\left(P_{fa}, P_d\right))}(P_d), \tag{2.83}$$

où $Q^{-1}_{\chi_1^2}(P_{fa})$ et $Q^{-1}_{\chi_1^2(\kappa'\left(P_{fa}, P_d\right))}(P_d)$ sont, respectivement, les fonctions inverses de $Q_{\chi_1^2}(.)$ et $Q_{\chi_1^2(\kappa'\left(P_{fa}, P_d\right))}(.)$ qui désignent la surface sous la queue de distribution à droite des lois χ_1^2 et $\chi_1^2(\kappa'(P_{fa}, P_d))$.

En combinant, (2.82) et (2.83) nous obtenons

$$\delta_{\text{detection}} = \kappa(P_{fa}, P_d)\sqrt{\text{BCR}(\delta_{\text{detection}})}, \tag{2.84}$$

où le facteur de translation est donné par $\kappa(P_{fa}, P_d) = \sqrt{\kappa'(P_{fa}, P_d)}$ (voir, Fig. 2.12).

Il est intéressant de noter que le test d'hypothèses (2.77) est un test binaire unilatéral et que l'estimateur du MV utilisé est sans contrainte. Ainsi, on peut en déduire que le TRV, utilisé pour déterminer le SRLM, est [LC03b, Sch91, Kay98] : i) asymptotiquement UPP, et ii) a un taux de fausse alarme asymptotiquement constant.

Par conséquent, de (2.73) et (2.84), le SRLM, donné en (2.73), basé sur l'extension du critère de Smith est asymptotiquement équivalent (à un facteur de translation près) à un test d'hypothèses UPP (2.77).

Enfin, de (2.84) on note que le SRLM basé sur l'extension de Smith est exactement égal au SRLM basé sur le test d'hypothèses (2.77) pour toute valeur de P_{fa} et P_d vérifiant $\kappa(P_{fa}, P_d) = 1$ (voir, Fig. 2.12)

– Applications aux modèles harmoniques multidimensionnels

Nous avons appliqué le critère énoncé dans la section précédente au calcul du SRLM pour le modèle multidimensionnel harmonique, à deux sources et à P paramètres d'intérêt par source. Ce modèle très général peut être ainsi utilisé dans plusieurs applications, par exemple, la localisation de sources sous marines acoustiques [WZ97], le sondage de canal sans fil [MSPM04, STWT06], la localisation de sources en champ proche [EBRM09], la localisation des cibles multiples dans

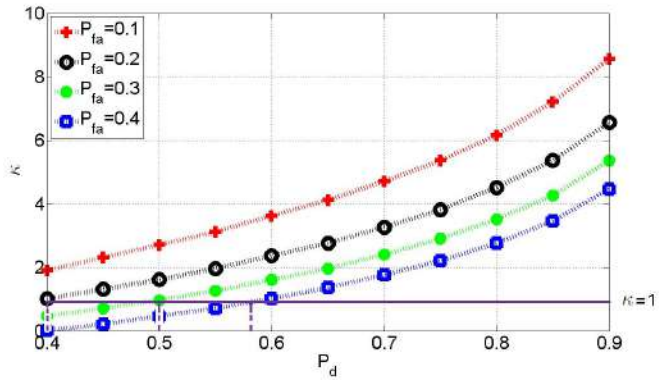


FIGURE 2.12 – Le facteur de translation κ en fonction de la probabilité de fausse alarme P_{fa} et la probabilité de détection P_d. On peut remarquer qu'augmenter P_d ou diminuer P_{fa} a pour effet d'augmenter la valeur du facteur de translation κ (ce qui est normal, puisque ceci correspond à un test d'hypothèses plus sélectif [Sch91, Kay98].

un système radar MIMO [NS09] etc. Le modèle multidimensionnel harmonique est donné par [HN98, PMB04, GS05, RHG07, Boy08, NS10] :

$$[\boldsymbol{\mathcal{Y}}(t)]_{n_1,\ldots,n_P} = [\boldsymbol{\mathcal{X}}(t)]_{n_1,\ldots,n_P} + [\boldsymbol{\mathcal{V}}(t)]_{n_1,\ldots,n_P}, \quad t = 1,\ldots,T, \quad \text{et} \quad n_p = 0,\ldots,N_p-1, \quad (2.85)$$

où les tenseurs $\boldsymbol{\mathcal{Y}}(t)$, $\boldsymbol{\mathcal{X}}(t)$ et $\boldsymbol{\mathcal{V}}(t)$ représentent les observations bruitées, les observations non bruitées et le bruit additif. Le nombre d'observations et le nombre de capteurs dans chaque vecteur sont notés T et $(N_1,\ldots,N_P)$, respectivement. Plus précisément, les observations pour le modèle multidimensionnel harmonique non bruitées sont données par [PMB04, HN98, RHG07, Boy08] :

$$[\boldsymbol{\mathcal{X}}(t)]_{n_1,\ldots,n_P} = \sum_{m=1}^{2} s_m(t) \prod_{p=1}^{P} e^{j\omega_m^{(p)} n_p}, \quad (2.86)$$

où $\omega_m^{(p)}$ et $s_m(t)$ sont la $m^{ème}$ fréquence le long de la $p^{ème}$ dimension et le $m^{ème}$ signal source, respectivement. Le signal source est supposé de la forme $s_m(t) = \alpha_m(t)e^{j\phi_m(t)}$ où $\alpha_m(t)$ et $\phi_m(t)$ représentent l'amplitude réelle et la phase du $m^{ème}$ signal source à la $t^{ème}$ observation, respectivement. Afin de simplifier les expressions, nous supposerons que le bruit est un processus aléatoire blanc, complexe circulaire, Gaussien de moyenne nulle et de variance inconnue σ^2. De plus, les signaux sources sont supposés connus et orthogonaux [LC93, NS09].

Le vecteur de paramètres inconnus est alors donné par

$$\boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{\omega}^T & \sigma^2 \end{bmatrix}^T, \quad (2.87)$$

où

$$\boldsymbol{\omega} = \begin{bmatrix} \left(\boldsymbol{\omega}^{(1)}\right)^T & \ldots & \left(\boldsymbol{\omega}^{(P)}\right)^T \end{bmatrix}^T, \quad (2.88)$$

avec

$$\boldsymbol{\omega}^{(p)} = \begin{bmatrix} \omega_1^{(p)} & \omega_2^{(p)} \end{bmatrix}^T. \quad (2.89)$$

Afin d'appliquer (2.84) nous commençons par calculer la BCR. Rappelons, qu'à notre connaissance, aucune expression analytique non matricielle de la BCR pour le modèle (2.85) n'est disponible dans la littérature. Après calculs, nous obtenons que la BCR pour le modèle harmonique

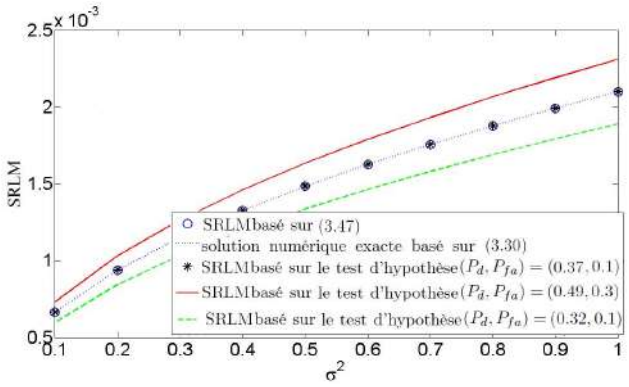


FIGURE 2.13 – SRLM en fonction de σ^2 pour $T = 100$.

multidimensionnel à P paramètres d'intérêt par source, sous l'hypothèse d'orthogonalité des signaux sources, pour le paramètre $\omega_m^{(p)}$ est donné par

$$\mathrm{BCR}(\omega_m^{(p)}) = \frac{6}{TN\mathrm{RSB}_m} C_p, \quad m \in \{1, 2\}, \tag{2.90}$$

où $N = \prod_{p=1}^{P} N_p$, $\mathrm{RSB}_m = \frac{\|\boldsymbol{\alpha}_m\|^2}{\sigma^2}$ représente le rapport signal à bruit de la $m^{ème}$ source et où

$$C_p = \frac{N_p(1 - 3V_P) + 3V_P + 1}{(N_p + 1)(N_p^2 - 1)} \quad \text{avec} \quad V_P = \frac{1}{1 + 3\sum_{p=1}^{P} \frac{N_p - 1}{N_p + 1}}. \tag{2.91}$$

De plus, les termes croisés sont donnés par

$$\mathrm{BCR}(\omega_m^{(p)}, \omega_{m'}^{(p')}) = \begin{cases} 0 & \text{pour } m \neq m', \\ \frac{-6}{TN\mathrm{RSB}_m} \tilde{C}_{p,p'} & \text{pour } m = m' \text{ et } p \neq p', \end{cases} \tag{2.92}$$

où

$$\tilde{C}_{p,p'} = \frac{3V_P}{(N_p + 1)(N_{p'} + 1)}. \tag{2.93}$$

Nous obtenons le SRLM pour le modèle (2.85) :

Le SRLM pour le modèle harmonique multidimensionnel à P paramètres d'intérêt par source, sous l'hypothèse d'orthogonalité des signaux sources, est donné par

$$\delta = \sqrt{\frac{6}{TN\mathrm{RSB}_E} \left(\sum_{p=1}^{P} C_p - \sum_{\substack{p,\, p' = 1 \\ p \neq p'}}^{P} g_p g_{p'} \tilde{C}_{p,p'} \right)}, \tag{2.94}$$

où le RSB étendu est donné par $\mathrm{RSB}_E = \frac{\mathrm{RSB}_1 \mathrm{RSB}_2}{\mathrm{RSB}_1 + \mathrm{RSB}_2}$ avec $g_p = \mathrm{sgn}\left(\omega_1^{(p)} - \omega_2^{(p)}\right)$.

– Autre approche pour le calcul du SRLM

Nous avons également traité deux exemples assez connus en traitement d'antenne qui sont : la localisation de sources en champ proche et la localisation de cibles à l'aide d'un radar MIMO en présence d'interférences. Cette approche alternative diffère de la précédente dans le sens où

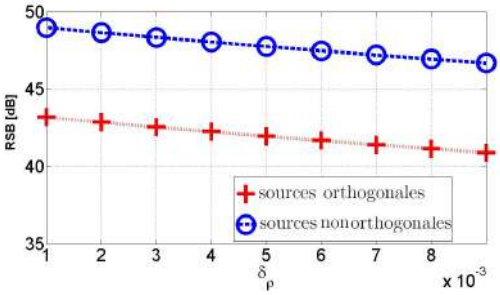


FIGURE 2.14 – Le RSB requis pour résoudre deux sources situées en champ proche en fonction de δ_ρ pour $\delta_\kappa = 0.003$ dans le cas de signaux sources orthogonaux et signaux sources non orthogonaux. On remarque le même comportement du RSB en fonction de δ_κ pour δ_ρ fixe.

le SRLM sera considéré comme un vecteur comportant tous les écarts des paramètres d'intérêt. Dés lors, le test d'hypothèses adéquat est le suivant :

$$\begin{cases} \mathcal{H}_0 : & \boldsymbol{\delta} \triangleq \left[\omega_1^{(1)} - \omega_2^{(1)} \quad \dots \quad \omega_1^{(P)} - \omega_2^{(P)}\right]^T = \mathbf{0}, \\ \mathcal{H}_1 : & \boldsymbol{\delta} \neq \mathbf{0}, \end{cases} \tag{2.95}$$

où l'hypothèse $\mathcal{H}_0$ représente le cas où les deux sources sont combinées en une seule source, alors que l'hypothèse $\mathcal{H}_1$ incarne la situation où les deux sources sont résolues.

Pour la localisation de sources en champ proche, nous utiliserons le modèle déterministe d'observations défini au paragraphe 2.3.1.1.1 avec deux sources émettrices. De plus nous supposerons que les signaux sources sont connus. Nous avons montré que pour un nombre d'observations très grand, la relation entre le SRLM et le rapport signal à bruit (RSB) requis pour résoudre deux sources connues se situant dans le champ proche est donnée par

$$RSB \triangleq \frac{\|\boldsymbol{s}_1\|^2 + \|\boldsymbol{s}_2\|^2}{\sigma^2} = \lambda(P_{fa}, P_d) \frac{\|\boldsymbol{s}_1\|^2 + \|\boldsymbol{s}_2\|^2}{2\|\boldsymbol{s}_-\|^2 \boldsymbol{\delta}^T \mathbf{F} \boldsymbol{\delta}} \tag{2.96}$$

où $\boldsymbol{\delta} = [\delta_\rho \ \delta_\kappa]^T = [\rho_2 - \rho_1 \ \kappa_2 - \kappa_1]^T$ avec $\rho_m = \frac{-2\pi}{\nu} \sin(\theta_m)$ et $\kappa_m = \frac{\pi}{\nu r_m} \cos^2(\theta_m)$. Avec

$$\mathbf{F} = \begin{bmatrix} f_2 & f_3 \\ f_3 & f_4 \end{bmatrix} \tag{2.97}$$

et $f_i = \sum_{n=0}^{N-1} (d_n)^i$ où d_n représente la distance entre le premier capteur et le $n^{ème}$ capteur où $\lambda(P_{fa}, P_d)$ est solution de $Q^{-1}_{(\lambda(P_{fa},P_d))}(P_d) = Q^{-1}(P_{fa})$.

Pour des sources orthogonales (c'est-à-dire, $\boldsymbol{s}_1^H \boldsymbol{s}_2 = 0$) nous obtenons le résultat suivant (voir, figure 2.14) :

La relation entre le SRLM et le rapport signal à bruit requis pour résoudre deux sources orthogonales connues, RSB_0, se situant dans le champ proche est donnée par

$$RSB_o = \frac{\lambda(P_{fa}, P_d)}{2\boldsymbol{\delta}^T \mathbf{F} \boldsymbol{\delta}}. \tag{2.98}$$

Nous avons également calculé le SRLM pour la localisation de sources à l'aide d'un radar MIMO. Les observations issues d'un radar MIMO (dans le cas d'antennes réceptrice et émettrice

espacées entre elles [JLL09]) qui reçoit un signal réfléchi sur M cibles sont données pour la $\ell^{ème}$ impulsion par

$$\boldsymbol{Y}_\ell = \sum_{m=1}^{M} \rho_m e^{2i\pi f_m \ell}\ \boldsymbol{a}_{\mathcal{R}}(\omega_m^{(\mathcal{R})})\ \boldsymbol{a}_{\mathcal{T}}(\omega_m^{(\mathcal{T})})^T\ \boldsymbol{S} + \boldsymbol{V}_\ell, \quad \ell \in [0 : L-1] \tag{2.99}$$

où L, ρ_m et f_m représentent le nombre d'échantillons par période d'impulsion, un coefficient proportionnel à la section efficace du radar et la fréquence Doppler normalisée de la $m^{ème}$ cible, respectivement. T, $N_{\mathcal{T}}$ et $N_{\mathcal{R}}$ sont, respectivement, le nombre d'observations, le nombre de capteurs émetteurs et le nombre de capteurs à la réception. Dans la suite, les symboles $\mathcal{T}$ et $\mathcal{R}$ représenteront la partie émettrice et la partie réceptrice du radar MIMO.

De plus, la matrice sources de taille $N_{\mathcal{T}} \times T$ est donnée par $\boldsymbol{S} = \begin{bmatrix} \boldsymbol{s}_0 & \dots & \boldsymbol{s}_{N_{\mathcal{T}}-1} \end{bmatrix}^T$ où $\boldsymbol{s}_{N_t} = \begin{bmatrix} s_{N_t}(1) & \dots & s_{N_t}(T) \end{bmatrix}^T$, et $\boldsymbol{V}_\ell$ (de taille $N_{\mathcal{R}} \times T$) représente la matrice du bruit. Les vecteurs directionnels de transmission et de réception sont donnés par $\boldsymbol{a}_{\mathcal{T}}(.)$ et $\boldsymbol{a}_{\mathcal{R}}(.)$. Le $i^{ème}$ élément de chaque vecteur directionnel est donné par $[\ \boldsymbol{a}_{\mathcal{T}}(\omega_m^{(\mathcal{T})})]_i = e^{j\omega_m^{(\mathcal{T})} d_i^{(\mathcal{T})}}$ et $[\ \boldsymbol{a}_R(\omega_m^{(\mathcal{R})})]_i = e^{j\omega_m^{(\mathcal{R})} d_i^{(\mathcal{R})}}$ où $\omega_m^{(\mathcal{T})} = \frac{2\pi}{\nu}\sin(\psi_m)$, $\omega_m^{(\mathcal{R})} = \frac{2\pi}{\nu}\sin(\theta_m)$ et ψ_m est l'angle de la cible vu de l'antenne émettrice, θ_m est l'angle de la cible vu de l'antenne réceptrice, ν est la longueur d'onde. La distance entre le premier capteur et le $i^{ème}$ capteur est notée $d_i^{(\mathcal{T})}$ et $d_i^{(\mathcal{R})}$ pour l'antenne de transmission et pour l'antenne de réception, respectivement. Avant de présenter les résultats, commençons par introduire les quantités suivantes : $\omega_c^{(\mathcal{T})} = \frac{\omega_1^{(\mathcal{T})}+\omega_2^{(\mathcal{T})}}{2}$ et $\omega_c^{(\mathcal{R})} = \frac{\omega_1^{(\mathcal{R})}+\omega_2^{(\mathcal{R})}}{2}$ qui représentent les paramètres centraux. $\dot{\boldsymbol{a}}_{\mathcal{T}}(.) \triangleq \boldsymbol{a}_{\mathcal{T}}(.) \odot \boldsymbol{d}_{\mathcal{T}}$, et $\dot{\boldsymbol{a}}_{\mathcal{R}}(.) \triangleq \boldsymbol{a}_{\mathcal{R}}(.) \odot \boldsymbol{d}_{\mathcal{R}}$ avec $\boldsymbol{d}_T = [d_0^{(\mathcal{T})}\ d_1^{(\mathcal{T})}\ \dots\ d_{N_{\mathcal{T}}-1}^{(\mathcal{T})}]^T$ et $\boldsymbol{d}_R = [d_0^{(\mathcal{R})}\ d_1^{(\mathcal{R})}\ \dots\ d_{N_{\mathcal{R}}-1}^{(\mathcal{R})}]^T$.

Nous appliquons un test d'hypothèses binaire

$$\begin{cases} \mathcal{H}_0 : & (\delta_{\mathcal{R}}, \delta_{\mathcal{T}}) = (0,0), \\ \mathcal{H}_1 : & (\delta_{\mathcal{R}}, \delta_{\mathcal{T}}) \neq (0,0), \end{cases} \tag{2.100}$$

où $\delta_{\mathcal{T}} \triangleq \omega_2^{(\mathcal{T})} - \omega_1^{(\mathcal{T})}$ et $\delta_{\mathcal{R}} \triangleq \omega_2^{(\mathcal{R})} - \omega_1^{(\mathcal{R})}$.

Le tableau 2.3 (voir aussi figure 2.15) résume le lien entre le rapport signal à bruit (défini comme RSB $\triangleq \frac{tr\{\boldsymbol{S}\boldsymbol{S}^H\}}{T\sigma^2}$) et le SRLM pour le modèle (2.99) sous les hypothèses suivantes :

	Avec interférences	Sans interférences $M=2$	Sans interférences et antennes symétriques
Variance connue	$\frac{N_{\mathcal{T}}\lambda_{\mathrm{K}}(P_{fa},P_d)}{2\boldsymbol{\zeta}^H\boldsymbol{G}^H\boldsymbol{P}_{\boldsymbol{P}_{\boldsymbol{D}}^{\perp}\boldsymbol{G}}\boldsymbol{G}\boldsymbol{\zeta}}$	$\frac{\lambda_{\mathrm{K}}(P_{fa},P_d)}{2L\boldsymbol{\zeta}^H\boldsymbol{K}\boldsymbol{\zeta}}$	$\frac{2N_{\mathcal{T}}\lambda_{\mathrm{K}}(P_{fa},P_d)}{L\left(\left(\delta_{\mathcal{R}}^2+\delta_{\mathcal{T}}^2\right)(\alpha_2-\alpha_1)^2+\delta_{\mathcal{R}}^2\delta_{\mathcal{T}}^2(\alpha_2+\alpha_1)^2\right)}$
Variance inconnue	$\frac{N_{\mathcal{T}}\lambda_{\mathrm{U}}(P_{fa},P_d)}{2\boldsymbol{\zeta}^H\boldsymbol{G}^H\boldsymbol{P}_{\boldsymbol{P}_{\boldsymbol{D}}^{\perp}\boldsymbol{G}}\boldsymbol{G}\boldsymbol{\zeta}}$	$\frac{\lambda_{\mathrm{U}}(P_{fa},P_d)}{2L\boldsymbol{\zeta}^H\boldsymbol{K}\boldsymbol{\zeta}}$	$\frac{2N_{\mathcal{T}}\lambda_{\mathrm{U}}(P_{fa},P_d)}{L\left(\left(\delta_{\mathcal{R}}^2+\delta_{\mathcal{T}}^2\right)(\alpha_2-\alpha_1)^2+\delta_{\mathcal{R}}^2\delta_{\mathcal{T}}^2(\alpha_2+\alpha_1)^2\right)}$

TABLE 2.3 – Le RSB requis pour résoudre deux cibles.

- Le sous-espace des interférences (engendré par les $M-2$ sources restantes) est connu,
- Les paramètres centraux, $\omega_c^{(\mathcal{T})}$ et $\omega_c^{(\mathcal{R})}$ sont supposés connus ou préalablement estimés.

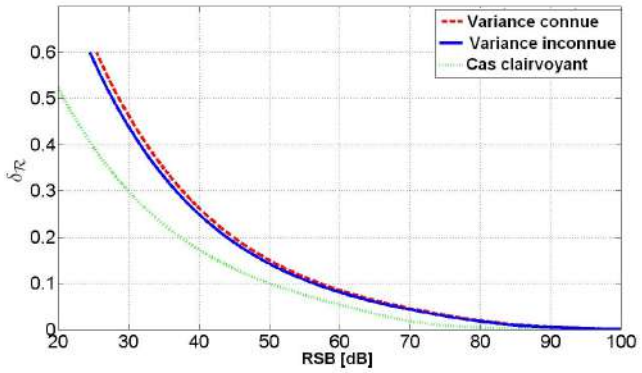


FIGURE 2.15 – $\delta_{\mathcal{R}}$ en fonction du RSB requis pour résoudre deux sources en présence d'une source interfèrente avec une ALU en émission et réception avec $N_{\mathcal{R}} = N_{\mathcal{T}} = 4$ capteurs, $L = 4$ et $T = 100$ observations. Le cas dit *clairvoyant* correspond au cas idéal où tous les paramètres sont connus y compris $\delta_{\mathcal{R}}$.

– Les α_m, $i = 1 \dots M$ sont considérés déterministes inconnus.
– le bruit est Gaussien complexe blanc circulaire de moyenne nulle et de variance σ^2.

Les différentes variables données dans le tableau 2.3 sont explicitées comme suit : $\boldsymbol{G} = \begin{bmatrix}\boldsymbol{\varrho}_2 & \boldsymbol{\varrho}_3 & \boldsymbol{\varrho}_4\end{bmatrix}$, la matrice des interférences $\boldsymbol{D} = \begin{bmatrix}\boldsymbol{\varrho}_1 & \boldsymbol{g}_3 & \dots & \boldsymbol{g}_M\end{bmatrix}$, $\boldsymbol{\alpha} = \begin{bmatrix}\alpha_1 + \alpha_2 & \alpha_3 & \dots & \alpha_M\end{bmatrix}^T$ avec

$$\boldsymbol{\zeta} = \frac{j}{2}\begin{bmatrix}\delta_{\mathcal{R}}(\alpha_2 - \alpha_1) \\ \delta_{\mathcal{T}}(\alpha_2 - \alpha_1) \\ \frac{j}{2}\delta_{\mathcal{R}}\delta_{\mathcal{T}}(\alpha_1 + \alpha_2)\end{bmatrix}, \tag{2.101}$$

et où $\boldsymbol{\varrho}_1 = \boldsymbol{c}(f) \otimes \boldsymbol{a}_{\mathcal{T}}(\omega_c^{(\mathcal{T})}) \otimes \boldsymbol{a}_{\mathcal{R}}(\omega_c^{(\mathcal{R})})$, $\boldsymbol{\varrho}_2 = \boldsymbol{c}(f) \otimes \boldsymbol{a}_{\mathcal{T}}(\omega_c^{(\mathcal{T})}) \otimes \dot{\boldsymbol{a}}_{\mathcal{R}}(\omega_c^{(\mathcal{R})})$, $\boldsymbol{\varrho}_3 = \boldsymbol{c}(f) \otimes \dot{\boldsymbol{a}}_{\mathcal{T}}(\omega_c^{(\mathcal{T})}) \otimes \boldsymbol{a}_{\mathcal{R}}(\omega_c^{(\mathcal{R})})$, $\boldsymbol{\varrho}_4 = \boldsymbol{c}(f) \otimes \dot{\boldsymbol{a}}_{\mathcal{T}}(\omega_c^{(\mathcal{T})}) \otimes \dot{\boldsymbol{a}}_{\mathcal{R}}(\omega_c^{(\mathcal{R})})$ et $\boldsymbol{c}(f) = [1 \dots e^{i2\pi f(L-1)}]^T$. De plus, $\boldsymbol{P}_{\boldsymbol{P}_{\boldsymbol{D}}^{\perp}\boldsymbol{G}} = \boldsymbol{P}_{\boldsymbol{D}}^{\perp} - \boldsymbol{P}_{[\boldsymbol{G}\boldsymbol{D}]}^{\perp}$. $\lambda_{\mathrm{K}}(P_{fa}, P_d)$ et $\lambda_{\mathrm{U}}(P_{fa}, P_d)$ sont solutions de $Q^{-1}_{\chi^2_{2r}}(P_{fa}) = Q^{-1}_{\chi^2_{2r}(\lambda_3(P_{fa},P_d))}(P_d)$ et $Q^{-1}_{F_{2r,2r'}}(P_{fa}) = Q^{-1}_{F_{2r,2r'}(\lambda_{\mathrm{U}}(P_{fa},P_d))}(P_d)$, respectivement, avec $r = T\mathcal{N}_T\mathcal{N}_R - M + 1$ et $r' = T\mathcal{N}_T\mathcal{N}_R - (M+2)$.

2.3.1.2.3 Géométrie d'antenne Les détails de cette partie sont donnés en **annexe G**. Dans le contexte de la localisation passive de sources utilisant un réseau de capteurs, les performances d'estimation de la direction d'arrivée (DDA) ne sont pas seulement liées au type d'estimateur considéré, mais aussi à la géométrie du réseau de capteurs, c'est-à-dire à la position des capteurs dans l'espace. Pour un réseau de capteurs, les "performances" peuvent être considérées de différents points de vue : les propriétés du faisceau de rayonnement, l'ambiguïté du réseau, l'isotropie, l'estimation en terme d'EQM, etc. Une pléthore de publications concernant l'étude de l'ambiguïté des réseaux (*e.g.*, [Man04, GC81, TGT96, LJ92, GW91]), du faisceau de rayonnement (*e.g.*, [Van02, SM05b]) et des propriétés isotropiques des réseaux (*e.g.*, [BM03]), sont disponibles dans la littérature.

Dans le cadre de la thèse de Dinh Thang Vu, nous nous sommes intéressés à la géométrie optimale des réseaux de capteurs qui mène aux meilleures performances en terme d'EQM. Plus particulièrement, nous nous sommes concentrés sur les géométries 3D des réseaux de capteurs qui ont été moins étudiées dans la littérature. En effet, malgré la richesse des résultats concernant les antennes planaires, les antennes 3D ont été peu étudiées. Par contre il existe des applications où les capteurs sont éparpillés dans l'espace ce qui fait que l'antenne est de forme arbitraire (réseau

de télescopes sur la surface de la terre, réseaux d'électrodes sur le crâne d'un patient, réseaux de capteurs dans une pièce ou dans un espace réduit pour des fonctions de robotique, réseau de bouées à la surface de la mer, etc). En outre, par rapport à l'antenne 2D, les antennes 3D ont certains avantages intuitifs, tels que lever l'ambiguïté de l'antenne 2D dans certains cas.

Les analyses fournies dans la littérature traitent de deux types de géométries : les géométries basées sur les réseaux circulaires [Van02] ou bien les réseaux sphériques [SSL68], et les géométries basées sur les réseaux linéaires tels que les réseaux linéaires uniformes, les réseaux en V, les réseaux en croix ou en rectangle. Plus particulièrement, nous nous sommes intéressés ici au deuxième type de géométrie.

La BCR a largement été utilisée dans la littérature pour décrire les propriétés fondamentales des réseaux. Par une forme simple des expressions de la BCR, [HS91] a montré l'impact de la position des capteurs sur la précision de l'estimation de la DDA dans le cas des réseaux de capteurs 2D. Concernant l'estimation de la DDA, dans [Nie94], [BM03], [MS91], des conditions sur la position des capteurs pour assurer la propriété d'isotropie pour laquelle, les réseaux de capteurs ont la même précision d'estimation sur l'ensemble des DDA, ont été calculées par l'étude des éléments de la BCR. Dans [YS05], [LS09], la BCR pour l'estimation de la position des sources basée sur la méthode TDOA est utilisée afin de montrer que la meilleure géométrie qui minimise la trace de la BCR contient des réseaux angulaires uniformes. L'application de la BCR Bayésienne pour le cas où les sources sont dans le même plan que le réseau et pour le contexte Bayésien où les paramètres (DDA) sont supposés aléatoires est disponible dans [OM05]. Dans [GM06], une étude approfondie de la BCR pour les réseaux de capteurs 2D a été proposée et a mené à des résultats intéressants concernant un type de réseau particulier appelé le réseau en V, en terme d'EQM et d'isotropie. Basé sur les résultats présentés dans [GM06], un nouveau type de géométrie des réseaux appelé antenne optimale sans ambiguïté basée sur antenne en V a été présenté dans [GAM09]. Finalement, dans [FC09], les auteurs ont montré que la BCR pour les modèles déterministe et stochastique, liée à la variance de l'estimation de la DDA obtenue par l'algorithme MUSIC, peut être exprimée avec la même expression qui est une fonction de la position des capteurs.

En traitement d'antenne, les signaux des sources sont généralement modélisés comme un processus aléatoire ou une séquence déterministe appelés modèle stochastique et modèle déterministe, respectivement [OVSN93]. Plus particulièrement, pour le modèle déterministe, la forme d'onde des signaux peut être supposée connue ou inconnue par le récepteur. Par conséquent, la charge de calcul varie en fonction de l'hypothèse sur la séquence des signaux utilisée. Si la séquence des signaux est supposée inconnue par le récepteur, ces signaux seront considérés comme des paramètres d'intérêt à estimer, ce qui augmente la dimension du vecteur des paramètres. Cependant, il y a des applications où l'amplitude des signaux est connue par le récepteur telles que dans les télécommunications. La connaissance du signal peut améliorer les performances d'estimation, et réduire également la complexité du problème. Les travaux concernant le contexte des signaux connus sont *e.g.*, [LC93, CM97, LHSV95, LV99, Cho04].

En fonction du modèle des signaux utilisé, il y a bien sûr deux types de BCR associés : la BCR stochastique et la BCR déterministe. Il a été montré que la BCR stochastique peut être atteinte pour un grand nombre des observations [SN90a], cependant, elle n'est pas atteinte pour un RSB élevé où le nombre des observations est faible [RFBL07]. D'autre part, la BCR déterministe est atteinte pour un RSB élevé [RFCL06] mais elle n'est pas atteinte pour un grand nombre des observations [SN90a]. De manière surprenante, tous les résultats proposés précédemment sont menés dans le cadre du modèle stochastique, c'est-à-dire que seule la BCR stochastique est utilisée. Nous avons montré que les résultats basés sur le modèle déterministe diffèrent considérablement de ceux basés sur le modèle stochastique.

Nous avons considéré les deux modèles déterministe et stochastique pour l'étude de la géométrie 3D.

- Tout d'abord, nous avons détaillé les expressions de la matrice d'information de Fisher (FIM) concernant l'estimation de l'azimut et de l'élévation dans le cas d'un réseau de capteurs 3D général.
- Deuxièmement, nous avons calculé les expressions analytiques de la BCR pour un réseau de capteurs quelconque constitué par un réseau planaire et une branche orthogonale. Ce modèle est la première étape pour analyser la contribution de la troisième dimension sur les performances d'estimation.
- Troisièmement, nous avons proposé plusieurs expressions analytiques de la BCR pour des réseaux de capteurs constitués par des branches ULA afin d'analyser l'impact de la géométrie d'antenne sur les performances d'estimation. Il faut noter que les types de géométries qui ont été étudiés dans le cas 2D et appelés antenne en V, antenne en L, ne sont que des cas particuliers dans notre approche. Dans [HSW91], les auteurs montrent que l'antenne en L permet d'améliorer de 37% la précision par rapport à l'antenne en croix. Dans [FT08], les auteurs ont introduit les conditions d'isotropie pour la position des capteurs, et pour l'angle d'ouverture entre les 2 branches de l'antenne en V uniforme/non-uniforme. Notre objectif était d'étendre ces géométries au cas 3D pour analyser l'impact de la branche 3D supplémentaire.
- Finalement, l'impact de la troisième dimension a été illustré à partir de comparaisons entre l'antenne 3D et l'antenne 2D, mais également entre l'antenne 3D et l'antenne circulaire uniforme (de même nombre de capteurs). Cependant, il faut noter que pour le même nombre de capteurs, afin de former la troisième dimension des réseaux, l'ouverture des réseaux sera diminuée. Par conséquent, la précision de l'estimation des réseaux de capteurs sera affectée.

On considère le problème classique où l'on cherche à localiser une source émettant un signal $s(t)$ déterministe (le cas stochastique est détaillé dans l'**annexe G**) et à bande étroite à l'aide d'un réseau de capteurs. Les capteurs sont supposés identiques et omni-directionnels. La source et le $i^{ème}$ capteur du réseau sont représentés dans l'espace à l'aide de leurs coordonnées sphériques, c'est-à-dire le couple (ϕ, φ) pour la source (supposée en champ lointain) et le triplet $(\rho_i, \varrho_i, \xi_i)$ pour le $i^{ème}$ capteur (voir figure 2.16(a)). Dans cette étude, nous considérons deux types de géométrie d'antenne. La première concerne une antenne planaire en V où les deux branches, séparées par un angle noté Δ, sont constituées par des antennes linéaires non obligatoirement uniformes. La deuxième géométrie considérée consiste simplement en une extension de l'antenne planaire où une branche (antenne linéaire non obligatoirement uniforme) orthogonale au plan est ajoutée (voir figure 1(b)). A partir des hypothèses susmentionnées, une simple analyse du retard inter-capteur conduit au modèle d'observation à la sortie du réseau suivant

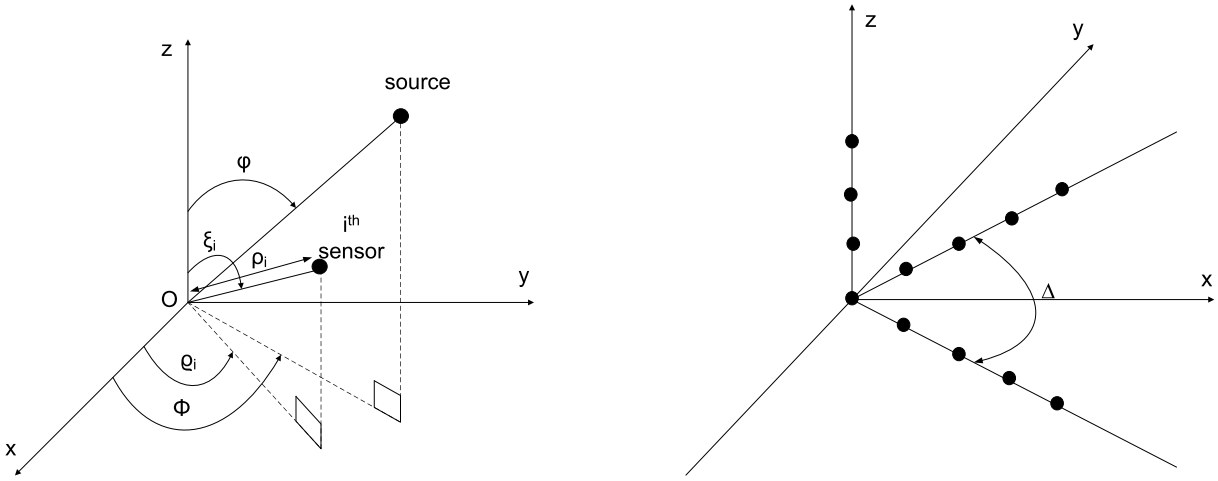
$$\mathbf{y}(t) = [y_1(t) \dots y_M(t)]^T = \mathbf{a}(\varphi,\phi)s(t) + \mathbf{b}(t) \tag{2.102}$$

où $t = 1, \dots, T$. T est le nombre d'observations, et le vecteur directionnel est donné par :

$$\mathbf{a}(\varphi,\phi) = \begin{bmatrix} e^{\frac{2j\pi\rho_1}{\lambda}(\sin\varphi\sin\xi_1\cos(\phi-\varrho_1)+\cos\xi_1\cos\varphi)} \\ \vdots \\ e^{\frac{2j\pi\rho_M}{\lambda}(\sin\varphi\sin\xi_M\cos(\phi-\varrho_M)+\cos\xi_M\cos\varphi)} \end{bmatrix}. \tag{2.103}$$

Le vecteur de bruit $\mathbf{b}(t) \in \mathbb{C}^M$ est supposé gaussien, circulaire, i.i.d., de moyenne nulle et de matrice de covariance $\sigma^2\mathbf{I}$. Le nombre de capteurs dans le plan est noté N_1 et le nombre de

capteurs pour la branche orthogonale dans le cas de l'antenne 3D est noté N_2. Le nombre total de capteurs $M = N_1 + N_2$ sera constant lors de la comparaison des deux géométries.



(a) Système de coordonnées pour la source et les capteurs du réseau

(b) Géométrie du réseau de capteurs 3D (Dans le cas du réseau 2D, seuls les capteurs dans le plan xOy sont présents)

FIGURE 2.16 – Géométrie du problème

– Bornes de Cramér-Rao

L'analyse des performances ultimes, en terme de variance, qu'un estimateur (non biaisé) peut espérer atteindre est généralement conduite à l'aide des BCR. Dans le cas du modèle d'observation (2.102), il est clair que $\mathbf{y}(t)$ est distribué selon une loi gaussienne multivariée de moyenne $\mathbf{a}(\varphi,\phi)s(t)$ et de variance $\sigma^2\mathbf{I}$. Les paramètres d'intérêt pour cette étude étant l'azimut et l'élévation, c'est-à-dire ϕ et φ (puisque la variance du bruit est découplée par rapport aux autres paramètres, cette dernière est omise du vecteur de paramètres). Seule la moyenne de $\mathbf{y}(t)$ est paramétrée. Dans ce cas, après concaténation de tous les vecteurs d'observation ($t = 1, \ldots, T$), la BCR, notée $\mathbf{C}$, est donnée par [SN89]

$$\mathbf{C} = \begin{bmatrix} C_{\varphi\varphi} & C_{\varphi\phi} \\ C_{\phi\varphi} & C_{\phi\phi} \end{bmatrix} = \frac{\sigma^2}{2\mathbf{s}^H\mathbf{s}} \begin{bmatrix} \mathrm{Re}\left(\frac{\partial \mathbf{a}^H(\varphi,\phi)}{\partial\varphi}\frac{\partial \mathbf{a}(\varphi,\phi)}{\partial\varphi}\right) & \mathrm{Re}\left(\frac{\partial \mathbf{a}^H(\varphi,\phi)}{\partial\varphi}\frac{\partial \mathbf{a}(\varphi,\phi)}{\partial\phi}\right) \\ \mathrm{Re}\left(\frac{\partial \mathbf{a}^H(\varphi,\phi)}{\partial\phi}\frac{\partial \mathbf{a}(\varphi,\phi)}{\partial\varphi}\right) & \mathrm{Re}\left(\frac{\partial \mathbf{a}^H(\varphi,\phi)}{\partial\phi}\frac{\partial \mathbf{a}(\varphi,\phi)}{\partial\phi}\right) \end{bmatrix}^{-1}. \quad (2.104)$$

où l'on définit $\mathbf{s} = [s(1) \ldots s(T)]^T$ et, où $C_{\varphi\varphi}$ et $C_{\phi\phi}$ représentent, respectivement, la BCR concernant l'élévation et la BCR concernant l'azimut. $C_{\varphi\phi}$ et $C_{\phi\varphi}$ représentent le couplage des paramètres φ et ϕ.

Grâce à la structure du vecteur directionnel donné par le modèle d'observation (2.102) et après quelques efforts calculatoires qui sont détaillés dans l'annexe, en posant : $\|\mathbf{s}\|^2 = \mathbf{s}^H\mathbf{s}$, $C_{RSB} = \frac{8\pi^2\|\mathbf{s}\|^2}{\sigma^2\lambda^2}$, $S_1 = \sum_{i=1}^{N_1}\rho_i^2$, et $S_2 = \sum_{i=N_1+1}^{N_1+N_2}\rho_i^2$, on obtient les expressions analytiques des BCR dans le cas de l'antenne 3D :

$$C_{\varphi\varphi}^{3D} = \frac{2}{C_{RSB}} \frac{1 - \cos\Delta\cos 2\phi}{S_1 \sin^2\Delta\cos^2\varphi + 2S_2\sin^2\varphi(1-\cos\Delta\cos 2\phi)}, \tag{2.105}$$

$$C_{\phi\phi}^{3D} = \frac{4}{C_{RSB}\sin^2\varphi} \frac{\frac{1}{2}S_1\cos^2\varphi(1+\cos\Delta\cos 2\phi) + S_2\sin^2\varphi}{S_1^2\sin^2\Delta\cos^2\varphi + 2S_1S_2\sin^2\varphi(1-\cos\Delta\cos 2\phi)}, \tag{2.106}$$

$$C_{\varphi\phi}^{3D} = \frac{1}{C_{RSB}\tan\varphi} \frac{S_1\cos\Delta\sin 2\phi}{S_1^2\sin^2\Delta\cos^2\varphi + 2S_1S_2\sin^2\varphi(1-\cos\Delta\cos 2\phi)}. \tag{2.107}$$

Puisque l'antenne 2D n'est qu'un cas particulier de l'antenne 3D ($N_2 = 0$), les BCR sont obtenues en posant $S_2 = 0$ dans les équations ci-dessus

$$C_{\varphi\varphi}^{2D} = \frac{2}{C_{RSB}} \frac{1 - \cos\Delta\cos 2\phi}{S_1\sin^2\Delta\cos^2\varphi}, \tag{2.108}$$

$$C_{\phi\phi}^{2D} = \frac{2}{C_{RSB}} \frac{1 + \cos\Delta\cos 2\phi}{S_1\sin^2\Delta\sin^2\varphi}, \tag{2.109}$$

$$C_{\varphi\phi}^{2D} = \frac{1}{C_{RSB}} \frac{\cos\Delta\sin 2\phi}{S_1\sin^2\Delta\cos\varphi\sin\varphi}. \tag{2.110}$$

De plus, dans le cas particulier où l'on travaille avec $\Delta = \frac{\pi}{2}$, c'est-à-dire lorsque l'antenne 2D et l'antenne 3D représentent, respectivement, la base canonique de $\mathbb{R}^2$ et de $\mathbb{R}^3$, on obtient des formules compactes :

$$C_{\varphi\varphi}^{3D\perp} = \frac{1}{C_{RSB}} \frac{2}{S_1\cos^2\varphi + 2S_2\sin^2\varphi}, \tag{2.111}$$

$$C_{\phi\phi}^{3D\perp} = \frac{2}{C_{RSB}S_1\sin^2\varphi}, \tag{2.112}$$

$$C_{\varphi\phi}^{3D\perp} = 0. \tag{2.113}$$

et

$$C_{\varphi\varphi}^{2D\perp} = \frac{2}{C_{RSB}S_1\cos^2\varphi}, \tag{2.114}$$

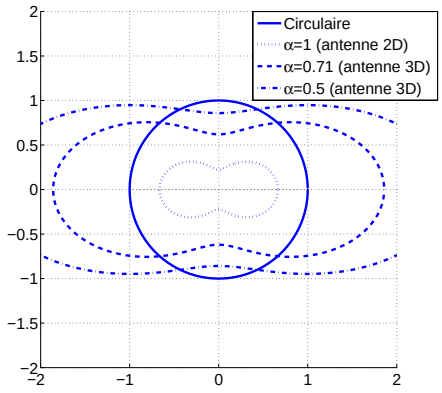
$$C_{\phi\phi}^{2D\perp} = \frac{2}{C_{RSB}S_1\sin^2\varphi}, \tag{2.115}$$

$$C_{\varphi\phi}^{2D\perp} = 0. \tag{2.116}$$

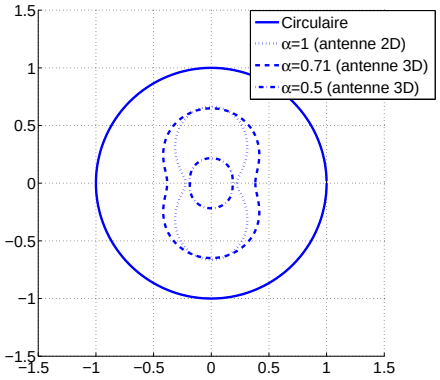
A partir de ces expressions, on peut remarquer que :

- lorsque la source se situe dans le plan, c'est-à-dire lorsque $\varphi = \frac{\pi}{2}$, $C_{\varphi\varphi}^{2D}$ tend vers l'infini tandis que $C_{\varphi\varphi}^{3D}$ reste fini. L'antenne 3D permet donc de lever l'ambiguïté.
- Dans le cas où $\Delta = \frac{\pi}{2}$, il y a découplage entre ϕ et φ ce qui confirme l'intuition. De plus, $C_{\phi\phi}^{3D\perp}$ et $C_{\phi\phi}^{2D\perp}$ ne dépendent plus de ϕ (propriété d'isotropie par rapport à ϕ). Si de plus $S_1 = 2S_2$, c'est-à-dire lorsque les trois branches de l'antenne 3D sont constituées par des antennes linéaires uniformes avec le même nombre de capteurs, l'estimation de φ ne depend plus de la position de la source (propriété d'isotropie par rapport à φ et ϕ) pour l'antenne 3D.
- Analyses et simulations

Des résultats de simulation concernant le comportement des BCR calculées précédemment en fonction du degré de liberté Δ. Toutes les branches, que le réseau soit 2D (deux branches) ou 3D (trois branches), sont des antennes linéaires uniformes avec un espacement inter-capteur



(a) BCR d'azimut normalisée en fonction de α à $\Delta = 60°$ et $\varphi = 45°$



(b) BCR d'élévation normalisée en fonction de α à $\Delta = 60°$ et $\varphi = 45°$

FIGURE 2.17 – Comparaison de performance des antennes en V avec une antenne circulaire

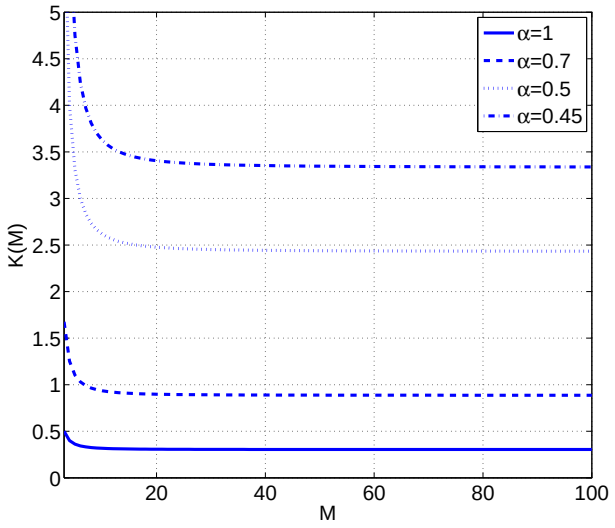


FIGURE 2.18 – Fraction $K(M)$ en fonction du nombre de capteurs M

d'une demi longueur d'onde. Pour toutes les simulations, le rapport signal sur bruit est de $10dB$ et le nombre d'observations est de $T = 50$.

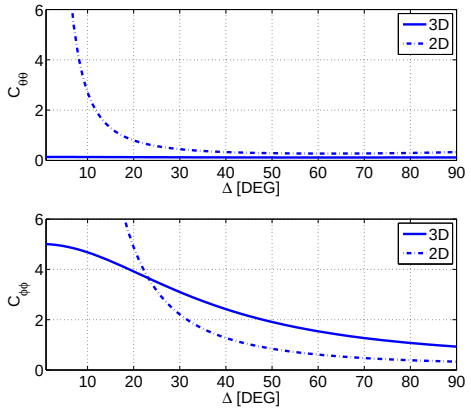
Il est intéressant de comparer les performances de l'antenne en V avec une antenne isotrope classique telle que l'antenne circulaire. Ces antennes ont le même nombre de capteurs. L'antenne circulaire ayant des capteurs séparés de manière équidistante $\frac{\lambda}{2}$, la valeur de son rayon est donc $r = \frac{\lambda}{4 \sin \frac{\pi}{M}}$. En posant $\alpha = \frac{N_1}{M} \leq 1$, il vient que la valeur de α associée à l'antenne planaire est égale à 1 tandis que celle associée à l'antenne 3D est strictement inférieure à 1. Les figures 2.17(a) et 2.17(b) montrent respectivement les BCR concernant l'azimut et l'élévation normalisées par la borne de l'antenne circulaire pour un angle d'ouverture $\Delta = 60°$ et pour une élévation $\varphi = 45°$. Les performances d'estimation concernant l'élévation des antennes en V sont toujours meilleures

par rapport à l'antenne circulaire, alors que celles concernant l'estimation d'azimut sont liées au nombre de capteurs que l'on place sur l'axe orthogonal, c'est-à dire au coefficient α. Pour des valeurs de α proches de 1, l'estimation de l'élévation des antennes en V est meilleure que celle de l'antenne circulaire.

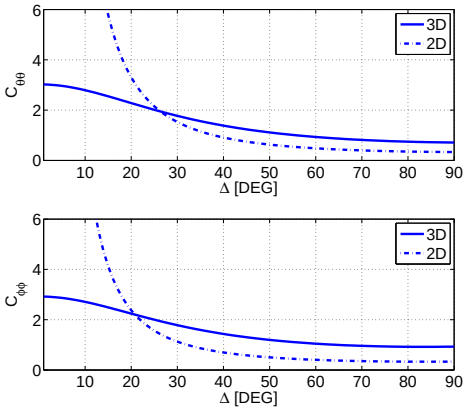
En particulier, si l'antenne en V et son extension 3D sont isotropes ($\Delta = \frac{\pi}{2}$), le rapport de la BCR sur l'azimut de ces antennes est donné par :

$$\begin{array}{l} K(M) = \frac{C_{\phi\phi}^{2D,3D\perp}}{C_{\phi\phi}^{Circulaire}} = \frac{3}{\alpha(\alpha^2 M^2 - 1)\sin^2\frac{\pi}{M}}. \\ \text{Si } \alpha = 1 \text{ et } M >> 1 \rightarrow K(M) = \frac{1}{3}. \\ \text{Si } \alpha < 1 \text{ et } \alpha M >> 1 \rightarrow K(M) = \frac{1}{3\alpha^2}. \end{array} \tag{2.117}$$

On peut dire que l'antenne en V a de meilleures performances en terme d'estimation d'azimut par rapport à l'antenne circulaire si et seulement si la fraction $K(M)$ est inférieure à 1. La figure 2.18 montre que l'antenne isotrope en V est meilleure que l'antenne circulaire à condition que la valeur de α satisfasse : $0.76 < \alpha < 1$.



(a) Comportement de $C_{\varphi\varphi}^{3D}$, $C_{\varphi\varphi}^{2D}$, $C_{\phi\phi}^{3D}$ et $C_{\phi\phi}^{2D}$ normalisées par la valeur de la BCR de l'antenne circulaire par rapport à Δ avec $\phi = 20°$ et $\varphi = 70°$



(b) Comportement de $C_{\varphi\varphi}^{3D}$, $C_{\varphi\varphi}^{2D}$, $C_{\phi\phi}^{3D}$ et $C_{\phi\phi}^{2D}$ normalisées par la valeur de la BCR de l'antenne circulaire par rapport à Δ avec $\phi = 50°$ et$\varphi = 30°$

FIGURE 2.19 – BCR normalisée en fonction de l'angle d'ouverture

Dans la suite, on compare les performances d'estimation entre le modèle 2D et 3D. Dans cette simulation, le réseau 2D est composé de $M = 7$ capteurs (un à l'origine plus trois sur chacune des deux branches). Le réseau 3D est également composé de $M = 7$ capteurs (un à l'origine plus deux sur chacune des trois branches). La figure 2.19(a) montre le comportement de $C_{\varphi\varphi}^{3D}$, $C_{\varphi\varphi}^{2D}$, $C_{\phi\phi}^{3D}$ et $C_{\phi\phi}^{2D}$ par rapport à l'angle d'ouverture Δ variant de 0 à $\frac{\pi}{2}$. Pour cette simulation, les valeurs de ϕ et φ sont respectivement de 20° et 70°. Nous qualifierons ce scénario de source rasante par rapport au plan de l'antenne. On observe que pour l'estimation de l'élévation, φ, le réseau 3D permet toujours d'obtenir de meilleures performances par rapport au réseau 2D. Ceci est toujours vrai pour une valeur d'élévation de $\varphi \geq 62.2°$ dans ce cas, car on montre facilement que :

$$\frac{C_{\varphi\varphi}^{3D}}{C_{\varphi\varphi}^{2D}} < 1 \Leftrightarrow \varphi > \arctan\sqrt{\max_{\Delta,\phi}\{\Gamma\}}. \tag{2.118}$$

avec $\Gamma = \frac{\sin^2\Delta\big((M^2-1)-\alpha(\alpha^2M^2-1)\big)}{(1-\cos\Delta\cos 2\phi)4(1-\alpha)((1-\alpha)M+1)(2(1-\alpha)M+1)}$, $\alpha = \frac{N_1}{M} = \frac{5}{7}, M = 7, \varphi \in [0°, 90°], \Delta \in (0°, 180°), \phi \in [0°, 360°]$.

Au contraire, il existe une valeur de Δ (environ 23° dans ce cas) en dessous de laquelle le réseau 3D permet de meilleures performances par rapport au réseau 2D pour l'estimation de l'azimut. Cette valeur critique peut être obtenue en résolvant numériquement l'équation $C^{3D}_{\phi\phi} = C^{2D}_{\phi\phi}$ en Δ. La figure 2.19(b) montre les mêmes courbes, mais pour des valeurs de ϕ et φ respectivement égales à 50° et 30°. Nous qualifierons ce scénario de source haute par rapport au plan de l'antenne. Dans ce cas, il convient, contrairement à l'intuition, de choisir le réseau 2D au dessus d'une certaine limite d'angle d'ouverture obtenue en résolvant numériquement $\max\left(C^{3D}_{\phi\phi} = C^{2D}_{\phi\phi}, C^{3D}_{\varphi\varphi} = C^{2D}_{\varphi\varphi}\right)$.

2.3.2 Bornes globales déterministes

2.3.2.1 Localisation de source en champ proche

Nous rappelons que la BCR est une borne non utilisable dans les zones non-asymptotiques, et donc non utilisable dans ces zones et que la zone dite de décrochement est un accroissement brusque de l'erreur quadratique moyenne (dont la cause est l'apparition des observations aberrantes). Ainsi, la zone de décrochement d'un estimateur délimite sa zone de fonctionnement optimal. Toujours dans le cadre de la thèse de Mohammed Nabil El Korso, nous avons donc complété l'analyse présentée au paragraphe 2.3.1.1.1 en calculant différentes bornes déterministes inférieures de l'erreur quadratique moyenne dans le cadre de la localisation passive d'une source en champ proche : la borne de McAulay-Seidman (BMS) [MS69], la borne de Hammersley-Chapman-Robbins (BHCR) [Ham50], la borne de McAulay-Hofstetter (BMH) [MH71] et, plus particulièrement, une borne récemment introduite dans la littérature dénommée la borne de Fourier Cramér-Rao (BFCR) [TT10]. A notre connaissance, aucun résultat concernant des bornes inférieures autres que la BCR (par rapport à la localisation en champ proche), n'était disponible dans la littérature. Ce travail est actuellement soumis à *IEEE* Transactions on Signal Processing.

Le modèle des observations est le même que celui présenté au paragraphe 2.3.1.1.1. Les hypothèses suivantes ont été utilisées :

- On admet que le bruit suit une loi normale complexe circulaire multidimensionnelle, de moyenne nulle et de matrice de covariance $\mathbf{\Sigma}_{\text{bruit}}$ connue et supposée de rang plein.
- Le vecteur de paramètres inconnus est défini par $\boldsymbol{\xi} = [\omega\ \phi]^T$ [LC93].

La fonction de vraisemblance relative au vecteur des observations $\boldsymbol{\chi} = [\mathbf{y}^T(1)\dots\mathbf{y}^T(T)]^T \sim \mathcal{CN}(\boldsymbol{\mu}(\boldsymbol{\xi}_0), \mathbf{\Sigma}(\boldsymbol{\xi}_0))$ pour un $\boldsymbol{\xi}_0$ donné, s'écrit simplement par :

$$p(\boldsymbol{\chi}|\boldsymbol{\xi}_0) = \frac{1}{\pi^{NT}\det\{\mathbf{\Sigma}(\boldsymbol{\xi}_0)\}} e^{-(\boldsymbol{\chi}-\boldsymbol{\mu}(\boldsymbol{\xi}_0))^H \mathbf{\Sigma}(\boldsymbol{\xi}_0)^{-1}(\boldsymbol{\chi}-\boldsymbol{\mu}(\boldsymbol{\xi}_0))}, \tag{2.119}$$

où $\boldsymbol{\xi}_0, \omega_0, \phi_0$ représentent les vraies valeurs des paramètres candidats $\boldsymbol{\xi}, \omega$ et ϕ, respectivement. Les valeurs de $\boldsymbol{\mu}(\boldsymbol{\xi}_0)$ et $\mathbf{\Sigma}(\boldsymbol{\xi}_0)$ seront spécifiées par la suite (selon le modéle déterminisite ou aléatoire).

Dans [Gla72, FL02, CGQL08, TT10] les auteurs ont proposé une unification de différentes bornes sur l'erreur quadratique moyenne. Plus précisément, s'appuyant sur un problème d'optimisation sous contraintes, Forster et Larzabal [FL02], ont présenté une unification de bornes sur l'erreur quadratique moyenne en imposant des contraintes sur le biais. Ils ont montré, par un choix judicieux de ces contraintes, qu'on pouvait ainsi donner une expression explicite de la borne de Cramér-Rao, la borne de Barankin ou la borne de Bhattacharyya. Notons que l'on peut trouver l'extension des travaux de Forster et Larzabal [FL02] dans le cas de plusieurs paramètres inconnus déterministes dans [CGQL08]. Dans [TT10], Todros et Tabrikian ont proposé

une nouvelle classe de bornes sur l'erreur quadratique moyenne en utilisant la transformation intégrale généralisée appliquée à la fonction de vraisemblance. Ainsi, ils ont montré que certaines bornes sur l'erreur quadratique moyenne (par exemple, la borne de Cramér-Rao, la borne de McAulay-Seidman et la borne de Battacharya) sont obtenues par un choix approprié du noyau de la transformation intégrale de la fonction de rapport de vraisemblance.

En utilisant l'une des deux approches, on constate que l'unification s'exprime à l'aide d'une matrice $\boldsymbol{K}$ inversible et d'une matrice $\boldsymbol{\Gamma}$ à valeurs complexes, comme suit :

$$\mathrm{MSE}(\hat{\boldsymbol{\xi}}) = \int_{\mathbb{C}^{NT}} \left(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0\right)\left(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0\right)^T p(\boldsymbol{\chi}|\boldsymbol{\xi}_0) d\boldsymbol{\chi} \succcurlyeq \boldsymbol{C} = \boldsymbol{\Gamma}\boldsymbol{K}^{-1}\boldsymbol{\Gamma}^H \tag{2.120}$$

où $\boldsymbol{K}$ peut être décomposée à l'aide de $\boldsymbol{\gamma}$ comme suit $\boldsymbol{K} = \int_{\mathbb{C}^{NT}} \boldsymbol{\gamma}\boldsymbol{\gamma}^H p(\boldsymbol{\chi}|\boldsymbol{\xi}_0) d\boldsymbol{\chi}$ et où $\hat{\boldsymbol{\xi}}$ est un estimateur du vrai paramètre déterministe $\boldsymbol{\xi}_0$ et où $\boldsymbol{A} \succcurlyeq \boldsymbol{B}$ signifie que la matrice $\boldsymbol{A} - \boldsymbol{B}$ est définie non négative. Par conséquent, pour différentes valeurs de $\boldsymbol{\Gamma}$ et de $\boldsymbol{\gamma}$, on aura différentes bornes inférieures de l'erreur quadratique moyenne. La BCR (notée dans cette section $\boldsymbol{C}_{\mathrm{BCR}}$ par souci d'uniformisation) est donnée par le couple :

$$\begin{cases} \boldsymbol{\Gamma}_{\mathrm{BCR}} &= \mathbf{1}_2, \\ \boldsymbol{\gamma}_{\mathrm{BCR}} &= \frac{\partial \ln p(\boldsymbol{x}|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}}|_{\boldsymbol{\xi}=\boldsymbol{\xi}_0}, \end{cases} \tag{2.121}$$

où $\mathbf{1}_l$ représente le vecteur de dimension $l \in \{1, \ldots, L\}$ rempli de 1. La borne de McAulay-Seidman (BMS) [MS69] peut être définie par le couple suivant :

$$\begin{cases} \boldsymbol{\Gamma}_{\mathrm{BMS}} &= \boldsymbol{\Phi}, \\ \boldsymbol{\gamma}_{\mathrm{BMS}} &= [\nu(\boldsymbol{x}|\boldsymbol{\xi}_1) \ \ldots \ \nu(\boldsymbol{x}|\boldsymbol{\xi}_L)]^T, \end{cases} \tag{2.122}$$

où $\nu(\boldsymbol{x}|\boldsymbol{\xi}_l) = \frac{p(\boldsymbol{x}|\boldsymbol{\xi}_l)}{p(\boldsymbol{x}|\boldsymbol{\xi}_0)}$, $\boldsymbol{\Phi} = [\boldsymbol{\xi}_1 - \boldsymbol{\xi}_0 \ \ldots \ \boldsymbol{\xi}_L - \boldsymbol{\xi}_0]^T$ avec $\{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_L\}$ qui représente un ensemble de points test appartenant à $\boldsymbol{\Theta}$. La borne de Hammersley-Chapman-Robbins (BHCR) [Ham50] est, quant à elle, donnée par :

$$\begin{cases} \boldsymbol{\Gamma}_{\mathrm{BHCR}} &= [\mathbf{0}_2 \ \boldsymbol{\Phi}], \\ \boldsymbol{\gamma}_{\mathrm{BHCR}} &= [1 \ \boldsymbol{\gamma}_{\mathrm{BMS}}]^T, \end{cases} \tag{2.123}$$

où $\mathbf{0}_l$ représente le vecteur de dimension $l \in \{1, \ldots, L\}$ rempli de 0 et enfin la borne de McAulay-Hofstetter (BMH) [MH71] peut être exprimée par :

$$\begin{cases} \boldsymbol{\Gamma}_{\mathrm{BMH}} &= [\boldsymbol{I}_2 \ \boldsymbol{\Phi}], \\ \boldsymbol{\gamma}_{\mathrm{BMH}} &= [\boldsymbol{\gamma}_{\mathrm{BCR}} \ \boldsymbol{\gamma}_{\mathrm{BMS}}]^T, \end{cases} \tag{2.124}$$

où $\boldsymbol{I}_2$ est la matrice identité d'ordre 2. La borne de Fourier Cramér-Rao (BFCR), récemment proposée [TT10], peut également être écrite sous la forme (2.120). Pour avoir un gain en temps de calcul, cette dernière utilise la transformée de Fourier discrète (TFD) des vecteurs $\boldsymbol{\Phi}$ et $\boldsymbol{\gamma}_{\mathrm{BMS}}$. La TFD peut être obtenue grâce à une multiplication matricielle notée $\boldsymbol{W}$. Ainsi, le couple $(\boldsymbol{\Gamma}, \boldsymbol{\gamma})$ adéquat pour la BFCR est donné par :

$$\begin{cases} \boldsymbol{\Gamma}_{\mathrm{BFCR}} &= [\boldsymbol{I}_2 \ \boldsymbol{\Phi}\boldsymbol{W}^H] \\ \boldsymbol{\gamma}_{\mathrm{BFCR}} &= [\boldsymbol{\gamma}_{\mathrm{BCR}} \ \boldsymbol{\gamma}_{\mathrm{BMS}}\boldsymbol{W}^T]^T. \end{cases} \tag{2.125}$$

où

$$[\boldsymbol{W}]_{p,l} = \exp(-i\boldsymbol{\Omega}_p^T \boldsymbol{\xi}_l) \tag{2.126}$$

représente la matrice de transformation relative à la TFD bi-dimensionelle et $\boldsymbol{\Omega}_p$ s'exprime à l'aide du $p^{ème}$ point test frequentiel $\boldsymbol{f}_p = [f_p \ f'_p]^T$ comme suit : $\boldsymbol{\Omega}_p = \left[\frac{2\pi f_p}{\delta([\boldsymbol{\xi}]_1)L_1} \ \frac{2\pi f'_p}{\delta([\boldsymbol{\xi}]_2)L_2}\right]^T$, avec $L = L_1L_2$, tel que $f_p \in \{1, \ldots, L_1\}$, $f'_p \in \{1, \ldots, L_2\}$, L_i est le nombre de points test associé à la variable $[\boldsymbol{\xi}]_i$ et $\delta([\boldsymbol{\xi}]_i)$ est la distance (constante) entre deux points test associée à la variable $[\boldsymbol{\xi}]_i$, $i = 1, 2$.

Après calcul, on peut montrer que les bornes précitées peuvent être écrites comme suit pour le modèle d'observation sus-mentionné

$$\boldsymbol{C}^{(L)}_{\text{BMS}} = \boldsymbol{\Phi}\boldsymbol{\Psi}^{-1}\boldsymbol{\Phi}^T, \tag{2.127}$$

$$\boldsymbol{C}^{(L)}_{\text{BHCR}} = \boldsymbol{\Phi}\left(\boldsymbol{\Psi} - \mathbf{1}\mathbf{1}^T\right)^{-1}\boldsymbol{\Phi}^T, \tag{2.128}$$

$$\boldsymbol{C}^{(L)}_{\text{BMH}} = \boldsymbol{C}_{\text{BCR}} + \boldsymbol{Q}\boldsymbol{R}^{-1}\boldsymbol{Q}^T, \tag{2.129}$$

$$\boldsymbol{C}^{(L,P)}_{\text{BFCR}} = \boldsymbol{C}_{\text{BCR}} + \boldsymbol{Q}\boldsymbol{W}^H\left(\boldsymbol{W}\boldsymbol{R}\boldsymbol{W}^H\right)^{-1}\boldsymbol{W}\boldsymbol{Q}^T, \tag{2.130}$$

où nous avons introduit la dépendance de ces bornes par rapport aux points test [2] (symboles L et P). Notons $\text{DKL}\left(p(\boldsymbol{x}|\boldsymbol{\xi}_l)||p(\boldsymbol{x}|\boldsymbol{\xi}_0)\right)$, la distance de Kullback-Leibler [Sch91] entre $p(\boldsymbol{x}|\boldsymbol{\xi}_l)$ et $p(\boldsymbol{x}|\boldsymbol{\xi}_0)$. On peut alors définir tous les éléments apparaissant dans les équations ci-dessus comme suit :

$$\boldsymbol{Q} = \boldsymbol{C}_{\text{BCR}}\boldsymbol{D} - \boldsymbol{\Phi}, \tag{2.131}$$

tel que

$$\boldsymbol{D} = \begin{bmatrix}\boldsymbol{d}(\boldsymbol{\xi}_1) & \ldots & \boldsymbol{d}(\boldsymbol{\xi}_L)\end{bmatrix},$$

et

$$\boldsymbol{d}(\boldsymbol{\xi}_l) = -\left(\frac{\partial \text{DKL}\left(p(\boldsymbol{x}|\boldsymbol{\xi}_l)||p(\boldsymbol{x}|\boldsymbol{\xi})\right)}{\partial \boldsymbol{\xi}}\right)^T|_{\boldsymbol{\xi}=\boldsymbol{\xi}_0} \tag{2.132}$$

où la dérivée vectorielle est donnée par $\left[\dot{\boldsymbol{\mu}}(\boldsymbol{\xi}_0)\right]_{i,j} = \frac{\partial[\ \boldsymbol{\mu}(\boldsymbol{\xi})]_i}{\partial[\boldsymbol{\xi}]_j}|_{\boldsymbol{\xi}=\boldsymbol{\xi}_0}$. De plus, la matrice $\boldsymbol{R}$ est donnée par

$$\boldsymbol{R} = \boldsymbol{\Psi} - \boldsymbol{D}^T\boldsymbol{C}_{\text{BCR}}\boldsymbol{D}, \tag{2.133}$$

et les éléments de la matrice $\boldsymbol{\Psi}$ sont définis par

$$[\boldsymbol{\Psi}]_{m,n} = E_{\boldsymbol{\chi}|\boldsymbol{\xi}_0}\left\{v(\boldsymbol{x}, \boldsymbol{\xi}_m)v(\boldsymbol{x}, \boldsymbol{\xi}_n)\right\}, \tag{2.134}$$

où $E_{\boldsymbol{\chi}|\boldsymbol{\xi}_l}\{.\}$ indique l'opérateur d'espérance mathématique relatif à $p(\boldsymbol{\chi}|\boldsymbol{\xi}_l)$.

Sachant que, pour un modèle Gaussien (circulaire) à moyenne paramétrée ou à covariance paramétrée, $\boldsymbol{C}_{\text{BCR}}$ est donné par l'inverse de la MIF, alors dans la suite nous ne donnerons que l'expression de $\boldsymbol{D}$ et $\boldsymbol{\Psi}$. Ainsi, en utilisant la valeur de $\boldsymbol{D}$ et de $\boldsymbol{\Psi}$ et (2.131) et (2.133), nous avons obtenu $\boldsymbol{C}^{(L)}_{\text{BMS}}$, $\boldsymbol{C}^{(L)}_{\text{BHCR}}$, $\boldsymbol{C}^{(L)}_{\text{BMH}}$ et $\boldsymbol{C}^{(L,P)}_{\text{BFCR}}$.

– Expressions analytiques pour le modèle déterministe

Pour le modèle déterministe, nous avons $\boldsymbol{\chi} \sim \mathcal{CN}(\boldsymbol{\mu}(\boldsymbol{\xi}_0), \boldsymbol{\Sigma}_{\text{bruit}})$ avec

$$\boldsymbol{\mu}(\boldsymbol{\xi}_0) = [s(1)\mathbf{a}^T(\omega_0, \phi_0) \ldots s(L)\mathbf{a}^T(\omega_0, \phi_0)]^T,$$

2. Notons par ailleurs que les valeurs des points tests qui maximisent les bornes (dites bornes optimales) sont celles où la fonction d'ambiguïté exhibe des maxima locaux [RM95, XBR04, RM97, TK99, Xu01, RAFL07]. Cela étant dit, il a été montré que, dans le cas où, les points tests couvrent les extrémités de $\boldsymbol{\Theta}$ et aussi la vraie valeur du paramètre $\boldsymbol{\xi}_0$, alors même si on obtient des bornes en dessous des bornes optimales, la différence est suffisamment faible pour que leur utilisation reste pertinente.

et

$$[\mathbf{FIM}]_{i,k} = 2\Re\left\{\frac{\partial \boldsymbol{\mu}(\boldsymbol{\xi}_0)^H}{\partial\,[\boldsymbol{\xi}_\mathbf{0}]_i}\boldsymbol{\Sigma}_{\text{bruit}}^{-1}\frac{\partial \boldsymbol{\mu}(\boldsymbol{\xi}_0)}{\partial\,[\boldsymbol{\xi}_\mathbf{0}]_k}\right\},\ i=1,2,\ k=1,2. \tag{2.135}$$

Notons que :

$$\begin{aligned}
&\text{KLD}\left(p(\boldsymbol{\chi}|\boldsymbol{\xi}_n)||p(\boldsymbol{\chi}|\boldsymbol{\xi})\right) = \int_{\mathbb{C}^{NT}} p(\boldsymbol{\chi}|\boldsymbol{\xi}_n)\text{Ln}\frac{p(\boldsymbol{\chi}|\boldsymbol{\xi}_n)}{p(\boldsymbol{\chi}|\boldsymbol{\xi})}d\boldsymbol{\chi}\\
&= \int_{\mathbb{C}^{NT}} \left[(\boldsymbol{\chi}-\boldsymbol{\mu}(\boldsymbol{\xi}))^H\boldsymbol{\Sigma}_{\text{bruit}}^{-1}(\boldsymbol{\chi}-\boldsymbol{\mu}(\boldsymbol{\xi})) - (\boldsymbol{\chi}-\boldsymbol{\mu}(\boldsymbol{\xi}_n))^H\boldsymbol{\Sigma}_{\text{bruit}}^{-1}(\boldsymbol{\chi}-\boldsymbol{\mu}(\boldsymbol{\xi}_n))\right] p(\boldsymbol{\chi}|\boldsymbol{\xi}_n)d\boldsymbol{\chi}\\
&= \int_{\mathbb{C}^{NT}} \left[\boldsymbol{\chi}^H\boldsymbol{\Sigma}_{\text{bruit}}^{-1}(\boldsymbol{\mu}(\boldsymbol{\xi}_n)-\boldsymbol{\mu}(\boldsymbol{\xi})) - \boldsymbol{\mu}(\boldsymbol{\xi})^H\boldsymbol{\Sigma}_{\text{bruit}}^{-1}(\boldsymbol{\chi}-\boldsymbol{\mu}(\boldsymbol{\xi}))\right] f(\boldsymbol{\chi}|\boldsymbol{\xi}_n)d\boldsymbol{\chi}\\
&= (\boldsymbol{\mu}(\boldsymbol{\xi}_n)-\boldsymbol{\mu}(\boldsymbol{\xi}))^H\boldsymbol{\Sigma}_{\text{bruit}}^{-1}(\boldsymbol{\mu}(\boldsymbol{\xi}_n)-\boldsymbol{\mu}(\boldsymbol{\xi}))\,.
\end{aligned} \tag{2.136}$$

De plus, les éléments de la matrice $\boldsymbol{\Psi}$ peuvent être donnés par :

$$\begin{aligned}
&[\boldsymbol{\Psi}]_{m,n} = \int_{\mathbb{C}^{NT}} \frac{1}{\pi\,|\boldsymbol{\Sigma}_{\text{bruit}}|}\exp\left((\boldsymbol{\chi}-\boldsymbol{\mu}(\boldsymbol{\xi}_0))^H\boldsymbol{\Sigma}_{\text{bruit}}^{-1}(\boldsymbol{\chi}-\boldsymbol{\mu}(\boldsymbol{\xi}_0))\right)\times\\
&\exp\left(-(\boldsymbol{\chi}-\boldsymbol{\mu}(\boldsymbol{\xi}_m))^H\boldsymbol{\Sigma}_{\text{bruit}}^{-1}(\boldsymbol{\chi}-\boldsymbol{\mu}(\boldsymbol{\xi}_m)) - (\boldsymbol{\chi}-\boldsymbol{\mu}(\boldsymbol{\xi}_n))^H\boldsymbol{\Sigma}_{\text{bruit}}^{-1}(\boldsymbol{\chi}-\boldsymbol{\mu}(\boldsymbol{\xi}_n))\right)d\boldsymbol{\chi}\\
&= \frac{\alpha(\boldsymbol{\xi}_m,\boldsymbol{\xi}_n)}{\pi\,|\boldsymbol{\Sigma}_{\text{bruit}}|}\underbrace{\int_{\mathbb{C}^{NT}}\exp-(\boldsymbol{\chi}-\boldsymbol{\mu}(\boldsymbol{\xi}_m)-\boldsymbol{\mu}(\boldsymbol{\xi}_n)+\boldsymbol{\mu}(\boldsymbol{\xi}_0))^H\boldsymbol{\Sigma}_{\text{bruit}}^{-1}(\boldsymbol{\chi}-\boldsymbol{\mu}(\boldsymbol{\xi}_m)-\boldsymbol{\mu}(\boldsymbol{\xi}_n)+\boldsymbol{\mu}(\boldsymbol{\xi}_0))\,d\boldsymbol{\chi}}_{\pi|\boldsymbol{\Sigma}_{\text{bruit}}|}\\
&= \alpha(\boldsymbol{\xi}_m,\boldsymbol{\xi}_n),
\end{aligned}$$

où

$$\begin{aligned}
\alpha(\boldsymbol{\xi}_m,\boldsymbol{\xi}_n) =& \exp(-2\boldsymbol{\mu}(\boldsymbol{\xi}_0)^H\boldsymbol{\Sigma}_{\text{bruit}}^{-1}\boldsymbol{\mu}(\boldsymbol{\xi}_0) - \boldsymbol{\mu}(\boldsymbol{\xi}_0)^H\boldsymbol{\Sigma}_{\text{bruit}}^{-1}(\boldsymbol{\mu}(\boldsymbol{\xi}_m)+\boldsymbol{\mu}(\boldsymbol{\xi}_n))\\
&+\boldsymbol{\mu}(\boldsymbol{\xi}_m)^H\boldsymbol{\Sigma}_{\text{bruit}}^{-1}(\boldsymbol{\mu}(\boldsymbol{\xi}_n)-\boldsymbol{\mu}(\boldsymbol{\xi}_0)) + \boldsymbol{\mu}(\boldsymbol{\xi}_m)^H\boldsymbol{\Sigma}_{\text{bruit}}^{-1}(\boldsymbol{\mu}(\boldsymbol{\xi}_n)-\boldsymbol{\mu}(\boldsymbol{\xi}_0))).
\end{aligned} \tag{2.137}$$

De ce fait $\boldsymbol{C}_{\text{BMH}}^{(L)}$, $\boldsymbol{C}_{\text{BHCR}}^{(L)}$, $\boldsymbol{C}_{\text{BMS}}^{(L)}$ et $\boldsymbol{C}_{\text{BCRF}}^{(L,P)}$ sont donnés en remplaçant (2.135), (2.136) et (2.137) dans (2.131) et (2.133).

– Expressions analytiques pour le modèle stochastique

Pour le modèle stochastique nous avons $\boldsymbol{\chi}\sim\mathcal{CN}(\mathbf{0},\boldsymbol{\Sigma}(\boldsymbol{\xi}_0))$ où $\boldsymbol{\Sigma}(\boldsymbol{\xi}_0) = \sigma_s^2\boldsymbol{I}_T\otimes\mathbf{a}(\omega_0,\phi_0)\mathbf{a}^H(\omega_0,\phi_0)+\boldsymbol{\Sigma}_{\text{bruit}}$, avec

$$[\mathbf{FIM}]_{i,k} = T\ \text{tr}\left\{\boldsymbol{\Sigma}(\boldsymbol{\xi}_0)^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\xi}_0)}{\partial\,[\boldsymbol{\xi}_\mathbf{0}]_i}\boldsymbol{\Sigma}(\boldsymbol{\xi}_0)^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\xi}_0)}{\partial\,[\boldsymbol{\xi}_\mathbf{0}]_k}\right\},\ i=1,2,\ k=1,2. \tag{2.138}$$

Notons que :

$$\begin{aligned}
&\text{KLD}\left(p(\boldsymbol{\chi}|\boldsymbol{\xi}_n)||p(\boldsymbol{\chi}|\boldsymbol{\xi})\right) = \int_{\mathbb{C}^{NT}} p(\boldsymbol{\chi}|\boldsymbol{\xi}_n)\text{Ln}\frac{p(\boldsymbol{\chi}|\boldsymbol{\xi}_n)}{p(\boldsymbol{\chi}|\boldsymbol{\xi})}d\boldsymbol{\chi}\\
&= \int_{\mathbb{C}^{NT}} \frac{1}{\pi^{NT}\,|\boldsymbol{\Sigma}(\boldsymbol{\xi}_n)|}\left(\boldsymbol{\chi}^H\left(\boldsymbol{\Sigma}(\boldsymbol{\xi}_n)^{-1}-\boldsymbol{\Sigma}(\boldsymbol{\xi})^{-1}\right)\boldsymbol{\chi}\right)\exp\left(-\boldsymbol{\chi}^H\boldsymbol{\Sigma}(\boldsymbol{\xi}_n)^{-1}\boldsymbol{\chi}\right)d\boldsymbol{\chi} + \text{Ln}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\xi})|}{|\boldsymbol{\Sigma}(\boldsymbol{\xi}_n)|}\\
&= E_{\boldsymbol{\chi}|\boldsymbol{\xi}_n}\left\{\boldsymbol{\chi}^H\boldsymbol{\Sigma}(\boldsymbol{\xi})\boldsymbol{\chi}\right\} + E_{\boldsymbol{\chi}|\boldsymbol{\xi}_n}\left\{\boldsymbol{\chi}^H\boldsymbol{\Sigma}(\boldsymbol{\xi}_n)\boldsymbol{\chi}\right\} + \text{Ln}\frac{|\boldsymbol{\Sigma}(\boldsymbol{\xi})|}{|\boldsymbol{\Sigma}(\boldsymbol{\xi}_n)|}.
\end{aligned} \tag{2.139}$$

Comme

$$E_{\boldsymbol{\chi}|\boldsymbol{\xi}_n}\left\{\boldsymbol{\chi}^H\boldsymbol{\Sigma}(\boldsymbol{\xi})\boldsymbol{\chi}\right\} = \sum_{i=1}^{NT}\sum_{j=1}^{NT}E_{\boldsymbol{\chi}|\boldsymbol{\xi}_n}\left\{[\boldsymbol{\chi}]_i^*\,[\boldsymbol{\chi}]_j\left[\boldsymbol{\Sigma}(\boldsymbol{\xi})^{-1}\right]_{i,j}\right\} = \text{tr}\left(\boldsymbol{\Sigma}(\boldsymbol{\xi}_n)\boldsymbol{\Sigma}(\boldsymbol{\xi})^{-1}\right) \tag{2.140}$$

et

$$E_{\boldsymbol{\chi}|\boldsymbol{\xi}_n}\left\{\boldsymbol{\chi}^H\boldsymbol{\Sigma}(\boldsymbol{\xi}_0)^{-1}\boldsymbol{\chi}\right\} = NT. \tag{2.141}$$

Alors, en remplaçant (2.140) et (2.141) dans (2.139) nous obtenons :

$$\mathrm{KLD}\left(p(\boldsymbol{\chi}|\boldsymbol{\xi}_n)||f(\boldsymbol{\chi}|\boldsymbol{\xi})\right) = \mathrm{Ln}\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}(\boldsymbol{\xi}_n)|} + \mathrm{tr}\left(\boldsymbol{\Sigma}(\boldsymbol{\xi}_n)\boldsymbol{\Sigma}(\boldsymbol{\xi})^{-1}\right) + NT. \tag{2.142}$$

De plus, on notera que :

$$[\boldsymbol{\Psi}]_{n,m} = \frac{|\boldsymbol{\Sigma}(\boldsymbol{\xi}_0)|}{\pi^{NT}\,|\boldsymbol{\Sigma}(\boldsymbol{\xi}_m)|\,|\boldsymbol{\Sigma}(\boldsymbol{\xi}_n)|}\underbrace{\int_{\mathbb{C}^{NT}}\exp\left(-\boldsymbol{\chi}^H\left(\boldsymbol{\Sigma}(\boldsymbol{\xi}_m)^{-1}+\boldsymbol{\Sigma}(\boldsymbol{\xi}_n)^{-1}-\boldsymbol{\Sigma}(\boldsymbol{\xi}_0)^{-1}\right)\boldsymbol{\chi}\right)d\boldsymbol{\chi}}_{\pi^{NT}\left|(\boldsymbol{\Sigma}(\boldsymbol{\xi}_m)^{-1}+\boldsymbol{\Sigma}(\boldsymbol{\xi}_n)^{-1}-\boldsymbol{\Sigma}(\boldsymbol{\xi}_0)^{-1})^{-1}\right|} \tag{2.143}$$

$$= \frac{|\boldsymbol{\Sigma}(\boldsymbol{\xi}_0)|}{|\boldsymbol{\Sigma}(\boldsymbol{\xi}_m)|\,|\boldsymbol{\Sigma}(\boldsymbol{\xi}_n)|\,|\boldsymbol{\Sigma}(\boldsymbol{\xi}_m)^{-1}+\boldsymbol{\Sigma}(\boldsymbol{\xi}_n)^{-1}-\boldsymbol{\Sigma}(\boldsymbol{\xi}_0)^{-1}|} \tag{2.144}$$

De ce fait, en utilisant $\frac{\partial \mathrm{Ln}|\boldsymbol{\Sigma}(\boldsymbol{\xi})|}{\partial \boldsymbol{\xi}} = \mathrm{tr}\left\{\boldsymbol{\Sigma}(\boldsymbol{\xi})^{-1}\frac{\partial\boldsymbol{\Sigma}(\boldsymbol{\xi})}{\partial\boldsymbol{\xi}}\right\}$ [PP06], $\boldsymbol{C}^{(L)}_{\mathrm{BMH}}$, $\boldsymbol{C}^{(L)}_{\mathrm{BHCR}}$, $\boldsymbol{C}^{(L)}_{\mathrm{BMS}}$ et $\boldsymbol{C}^{(L,P)}_{\mathrm{BCRF}}$ sont données en remplaçant (2.135), (2.136) et (2.137) dans (2.131) et (2.133).

– Analyse numérique

Pour ces simulations, nous avons considéré une antenne composée de $N = 10$ capteurs avec une distance inter-capteurs $d = \frac{\lambda}{2}$. La source, située dans la région de Fresnel, est repérée par les coordonnées suivantes $(\theta, r) = (30^\circ, 6\lambda)$. On supposera également que $\boldsymbol{\Sigma}_{\mathrm{bruit}} = \sigma^2\boldsymbol{I}$.

Il est à noter que les EQM empiriques de l'estimateur du MVD représentées dans les Fig. 2.20 et 2.21, ont été obtenues avec 1000 tirages de type Monte-Carlo. L'ensemble des points tests utilisés pour la BMS, la BHCR, la BMH et la BFCR est égal à $L = 2^{14}$ (plus précisément, l'ensemble des points test suivant le paramètre ω est fixé à $L_1 = 2^7$, de même que celui par rapport à ϕ qui est donné par $L_2 = 2^7$). La BFCR se calcule aussi en choisissant un ensemble de points de test fréquentiels. A cet effet, et pour garder une complexité de calcul sensiblement égale à la BMS, la BHCR et la BMH, on a choisi deux points tests fréquentiels parmi les 2^{14} maximisant la BFCR.

Pour le modèle déterministe, les Fig. 2.20 et Fig. 2.21 nous montrent les différentes bornes de l'EQM des deux paramètres d'intérêt ω et ϕ. On constate tout d'abord que l'EQM sur ϕ est inférieure à celle sur ω, ce qui était prévisible vue la plage de variation des deux paramètres. De plus, la BMS, la BHCR, la BMH et la BFCR décrivent bien le décrochement du MV. Cela étant dit, on remarque que la BMH est la plus pertinente (prédiction du décrochement à moins de 4 dB), vient ensuite la BHCR et la BMS (prédiction du décrochement à moins de 7 dB). Enfin, la BFCR nous fournit une prédiction du décrochement avoisinant les 10 dBs. Les mêmes conclusions peuvent être déduites pour le modèle aléatoire (voir les Fig. 2.22 et Fig. 2.23.)

De façon générale on constate que la BFCR demeure moins performante que toutes les bornes présentées dans cette contribution. Ceci est dû au fait que la BFCR "comprime" les contraintes en appliquant la TFD. Cette compression de contraintes est à l'origine de la dégradation de cette borne. Ce point n'a pas été mentionné dans [TT10] où la BFCR apparaît comme une borne plus pertinente par rapport à la BMS, BHCR et la BMH dans le cas particulier de l'analyse spectrale. Cela étant dit, il faut noter que la BFCR a été calculée à partir de 2^9 points tests, or les autres bornes (BMS, BHCR et la BMH) ont étés calculées en maximisant seulement 1 point test parmi 2^9. Ceci explique, pourquoi dans [TT10] la BFCR apparaît comme étant plus précise que les autres bornes contrairement à l'exemple traité ici.

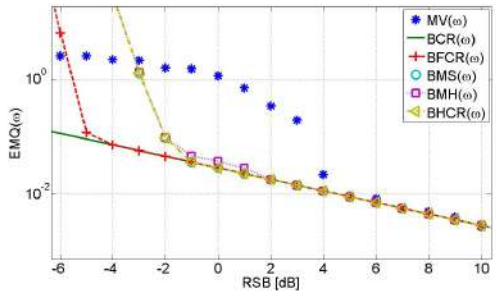


FIGURE 2.20 – Bornes inférieures de l'erreur quadratique moyenne pour le modèle déterministe en fonction de ω pour $(\theta, r) = (30^\circ, 6\lambda)$ et $T = 15$.

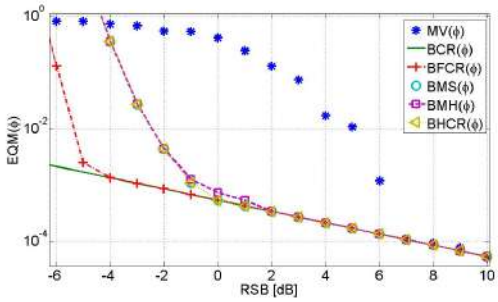


FIGURE 2.21 – Bornes inférieures de l'erreur quadratique moyenne pour le modèle déterministe en fonction de ϕ pour $(\theta, r) = (30^\circ, 6\lambda)$ et $T = 15$.

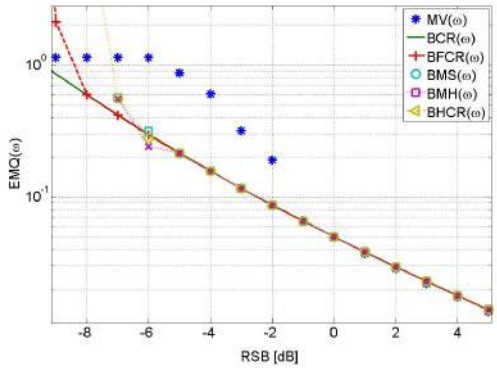


FIGURE 2.22 – Bornes inférieures de l'erreur quadratique moyenne pour le modèle aléatoire en fonction de ω pour $(\theta, r) = (30^\circ, 6\lambda)$ et $T = 100$.

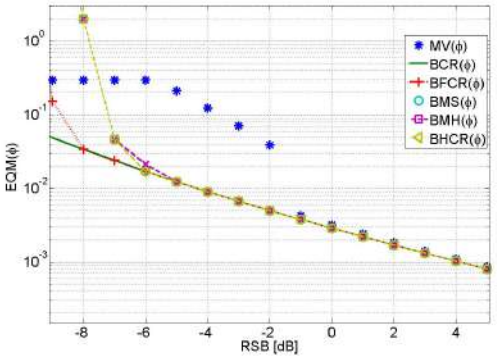


FIGURE 2.23 – Bornes inférieures de l'erreur quadratique moyenne pour le modèle aléatoire en fonction de ϕ pour $(\theta, r) = (30^\circ, 6\lambda)$ et $T = 100$.

2.3.2.2 Estimation de points de ruptures multiples

Avec Patricio La Rosa, Arye Nehorai, Washington University in St. Louis, USA, et Carlos Muravchick, Univérsité de La Plata, Argentine, nous avons calculé une borne inférieure de l'EQM pour un problème d'estimation de points de ruptures. Dans ce contexte, les paramètres sont discret et la classique borne de Cramér-Rao n'est plus applicable. En conséquence, nous nous sommes concentré sur la borne de Barankin qui ne requière pas les même conditions de régularités et est valide pour des paramètres discrets. En particulier, nous avons calculé la forme multivariée de la borne de Chapman–Robbins. Nous avons donné la structure de la matrice d'information de Barankin et montré que le cas particulier de deux points de ruptures est fondamental afin d'inverser la matrice complète. Nous avons proposé plusieurs forme analytiques de cette bornes pour l'estimation de points de ruptures dans des séquence de données gaussiennes et poissonniennes. Les détails sont donnés en **annexe E**.

2.3.3 Borne de Weiss-Weinstein

2.3.3.1 Localisation de sources et géometrie d'antenne

Dans le paragraphe 2.3.1.2.3, la BCR était utilisée pour étudier l'impact de la géométrie du réseau sur les performances d'estimation dans le contexte où les paramètres sont supposés déterministes. Afin de compléter cette étude, et toujours dans le cadre de la thèse de Dinh Thang Vu, nous nous sommes intéressés au développement d'un outil pour étudier cet impact dans le contexte Bayésien. La borne de Weiss-Weinstein qui est la borne la plus précise dans la famille des bornes de Weiss-Weinstein a été considérée. La BWW est une borne Bayésienne, donc, elle est utile pour les trois zones d'opération des estimateurs. Elle peut capturer le décrochement de l'EQM obtenue par l'estimateur maximum a posteriori. La BWW a été beaucoup moins étudiée que la BCR du fait de sa complexité. Les détails sont donnés en **annexe I**.

Encore une fois, nous avons étudié les deux modèles des signaux comme dans le chapitre précédent : le modèle stochastique et le modèle déterministe. Dans le contexte du traitement d'antenne, tandis que les expressions analytiques de la borne de Ziv-Zakai (plus précisément son extension proposée par Bell et. al. [BEV96]) ont été proposées il y a environ 15 ans pour le modèle stochastique et les résultats concernant la BWW étaient généralement réalisés par des simulations. Concernant le modèle stochastique, dans [NH88], la BWW a été évaluée par des simulations, et ensuite comparée avec l'EQM obtenue par l'algorithme MUSIC et l'algorithme de formation des voies en utilisant un réseau de capteurs de 8×8 éléments. Dans [NV94], les auteurs ont introduit un comparaison numérique entre la BCR Bayésienne, la borne Ziv-Zakai, et la BWW pour le problème de l'estimation de la DDA. Dans [Ath01], des simulations numériques de la BWW pour optimiser la position des capteurs dans un réseau linéaire non-uniforme ont été proposées. Toujours concernant le modèle stochastique, dans [XBR04], les auteurs ont dérivé des expressions quasi-analytiques d'une version simplifiée de la BWW pour le problème de l'estimation de la DDA dans le contexte du traitement champ adapté. En effet, l'intégration sur la fonction de distribution *a priori* n'a pas été effectuée. Le modèle déterministe avec la séquence des signaux supposée connue, a été étudié seulement dans [Ren07], où les expressions analytiques de la BWW étaient données pour le cas de l'analyse spectrale.

Bien que l'objectif principal était de donner des expressions analytiques de la BWW pour le problème de l'estimation de la DDA d'une seule source en utilisant un réseau de capteurs planaire quelconque, et pour les modèles déterministe et stochastique, nous avons également fourni des expressions quasi-analytiques de la borne qui pourraient être utiles pour d'autres problèmes. Premièrement, nous avons étudié le modèle général des observations Gaussiennes avec moyenne ou matrice de covariance paramétrée. En effet, pour ce modèle, l'un des succès de la BCR est dû à l'expression analytique de la matrice d'information de Fisher qui est appelée la formule de Slepian-Bang [Kay93]. Une telle formule n'avait pas été proposée pour la BWW. Deuxièmement, vu que l'un des objectif de cette partie est de fournir des formules qui pourraient être appliquées dans d'autres domaines, nous avons proposé quelques résultats pour le contexte de sources multiple sans avoir besoin de préciser la structure de la matrice directionnelle et de la matrice de covariance du bruit. Plus particulièrement, pour obtenir les expressions analytiques de la BWW, il faut calculer des intégrales sur les observations et sur les paramètres. Dans cette partie, l'intégration sur l'espace des observations a été proposée. Pour la deuxième intégration sur l'espace des paramètres, il a été nécessaire de détailler la structure de la matrice de covariance du bruit et de la matrice directionnelle afin d'obtenir des expressions analytiques. Ensuite, ces résultats ont été appliqués dans le cas particulier d'une seule source avec deux types de la géométrie du réseau : le réseau linéaire non-uniforme (estimer seulement l'angle d'élévation), et le réseau planaire (estimer l'angle azimut et l'angle d'élévation). Grâce à la structure exponentielle

du vecteur directionnel, nous avons obtenu des expressions de la BWW plus compactes. En effet, nous avons trouvé que pour le cas d'une seule source, l'intégration sur les paramètres était obtenue directement. Il faut noter qu'une des particularité de cette contribution par rapport aux publications précédentes concernant la BWW est que nous n'avons pas utilisé l'hypothèse $s = 1/2$. Finalement, quelques simulations ont été présentées. Nous avons utilisé la BWW pour étudier la valeur optimal de l'angle d'ouverture de l'antenne en V mentionné précédemment. Nous avons montré que pour le modèle déterministe, cette valeur est égale à 90°.

Nous nous sommes limité au cas de l'estimation passive d'une DDA pour une source, située en champ lointain et dont le signal est supposé à bande étroite à partir d'un réseau linéaire (non-uniforme) constitué de N capteurs (les différentes généralisations présentées ci-dessus sont détaillées dans l'annexe I). La position des capteurs dans le réseau est caractérisée par rapport à un réferentiel par le vecteur $\mathbf{d} = [d_1 \dots d_N]$. La réponse du i^{eme} capteur à l'instant t est un vecteur donné par : $y_i(t) = [\mathbf{a}(\varphi)]_i s(t) + \mathbf{b}_i(t), \quad t = 1, \dots, T,$. $s(t)$ est le signal de la source, T est le nombre des observations, et $[\mathbf{a}(\varphi)]_i = \exp\left(j\frac{2\pi}{\lambda} d_i \sin\varphi\right)$ est le i^{eme} élément du vecteur directionnel $\mathbf{a}(\varphi)$ où φ représente l'angle d'élévation. $\mathbf{b}_i(t)$ est un bruit additif supposé complexe, circulaire, gaussien de moyenne nulle et de covariance $\sigma^2\mathbf{I}$. Concernant le signal source, on considérera les deux modèles suivants :

- $\mathcal{M}_1$: Le modèle déterministe ou conditionnel où le signal est supposé connu [OVSN93].
- $\mathcal{M}_2$: Le modèle stochastique ou non-conditionnel où le signal est supposé aléatoire, complexe, circulaire, gaussien de moyenne nulle et de matrice de covariance $\sigma_s^2\mathbf{I}$ connue. Pour ce modèle, le signal est également supposé indépendant du bruit [OVSN93].

On considérera plus particulièrement l'estimation de l'angle électrique $\omega = \sin\varphi$. Cette étude de performance se déroulant dans le contexte Bayésien, on supposera que le paramètre ω est aléatoire avec une loi uniforme $\omega \sim \mathcal{U}[-1, 1]$ *a priori* :

$$p(\omega) = \begin{cases} \frac{1}{2} \text{ si } -1 \leq \omega \leq 1, \\ 0 \text{ si non.} \end{cases} \tag{2.145}$$

Donc, le modèle des observations à l'instant t s'écrit

$$\mathbf{y}(t) = \mathbf{a}(\omega)s(t) + \mathbf{b}(t), \tag{2.146}$$

où $\mathbf{y}(t) = [y_1(t) \dots y_N(t)]^T$. A partir des hypothèses précédentes, la fonction de vraisemblance de toutes les observations, *i.e.*, du vecteur $\mathbf{y} = [\mathbf{y}^T(1) \dots \mathbf{y}^T(T)]^T$, pour le modèle $\mathcal{M}_1$ est donnée par

$$p(\mathbf{y}|\omega) = \frac{1}{(\pi\sigma^2)^{2NT}} e^{\left(-\frac{1}{\sigma^2}\sum_{t=1}^{T}\|\mathbf{y}(t)-\mathbf{a}(\omega)s(t)\|^2\right)}, \tag{2.147}$$

et la fonction de vraisemblance pour le modèle $\mathcal{M}_2$ est donnée par

$$p(\mathbf{y}|\omega) = \frac{1}{\pi^{2NT}\left|\mathbf{R}(\omega)\right|^T} e^{\left(-\sum_{t=1}^{T}\mathbf{y}(t)^H\mathbf{R}(\omega)^{-1}\mathbf{y}(t)\right)}, \tag{2.148}$$

où $\mathbf{R}(\omega) = \sigma_s^2\mathbf{a}(\omega)\mathbf{a}(\omega)^H + \sigma^2\mathbf{I}_{2N}$ représente la matrice de covariance pour le modèle $\mathcal{M}_2$. La BWW sera dérivée pour le modèle $\mathcal{M}_1$ et $\mathcal{M}_2$.

– Borne de Weiss-Weinstein pour le réseau linéaire

La BWW est obtenue, en général, en cherchant le maximum d'une fonction sur un ensemble de points test et sur un ensemble de paramètres $s \in [0, 1]$. Concernant le paramètre s, on utilise

souvent l'hypothèse[3] $s = 1/2$ [VRBM10, XBR04, Ren07]. Ω et Θ représentent respectivement l'espace des observations et l'espace des paramètres, la BWW pour $s = 1/2$ s'écrit [WW88] :

$$\int_{\Theta}\int_{\Omega}(\hat{\omega}-\omega)^2 p(\mathbf{y},\omega)d\mathbf{y}d\omega \geq WWB = \sup_{h}\frac{h^2\eta(h,0)\eta(0,h)}{2(\eta(h,h)-\eta(h,-h))}, \tag{2.149}$$

où $\hat{\omega}$ est un estimateur de ω, où $p(\mathbf{y},.)$ représente la loi jointe entre le vecteur des observations et le paramètre (ou un point de test), et où h représente la différence entre le paramètre d'intérêt et un point de test appartenant à l'espace des paramètres (c'est-à-dire qu'il faut respecter $\omega + h \in \Theta$). On a défini

$$\eta(\alpha,\beta) = \int_{\Theta}\int_{\Omega}\sqrt{p(\mathbf{y},\omega+\alpha)p(\mathbf{y},\omega+\beta)}d\mathbf{y}d\omega = \int_{\Theta}\sqrt{p(\omega+\alpha)p(\omega+\beta)}\zeta(\omega,\alpha,\beta)d\omega \tag{2.150}$$

et $\zeta(\omega,\alpha,\beta) = \int_{\Omega}\sqrt{p(\mathbf{y}|\,\omega+\alpha)p(\mathbf{y}|\,\omega+\beta)}d\mathbf{y}$ où $p(.)$ représente la distribution *a priori* du paramètre.

– Modèle déterministe $\mathcal{M}_1$

À partir de l'équation (2.147), l'expression de $\zeta(\alpha,\beta)$ est donnée par :

$$\zeta(\alpha,\beta) = \int_{\Omega}\frac{1}{(\pi\sigma^2)^{2NT}}\times e^{\left(-\frac{1}{2\sigma^2}\sum_{t=1}^{T}\left(\|\mathbf{y}(t)-\mathbf{a}(\omega+\alpha)s(t)\|^2+\|\mathbf{y}(t)-\mathbf{a}(\omega+\beta)s(t)\|^2\right)\right)}d\mathbf{y}. \tag{2.151}$$

Par le changement de variable

$$\mathbf{x}(t) = \mathbf{y}(t) - \frac{1}{2}\left(\mathbf{a}(\omega+\alpha)s(t) + \mathbf{a}(\omega+\beta)s(t)\right),$$

on obtient

$$\begin{aligned}&-\tfrac{1}{2\sigma^2}\sum_{t=1}^{T}\left(\|\mathbf{y}(t)-\mathbf{a}(\omega+\alpha)s(t)\|^2+\|\mathbf{y}(t)-\mathbf{a}(\omega+\beta)s(t)\|^2\right)\\ &=-\tfrac{1}{\sigma^2}\sum_{t=1}^{T}\left(\|\mathbf{x}(t)\|^2+\tfrac{1}{4}\|\mathbf{a}(\omega+\alpha)-\mathbf{a}(\omega+\beta)\|^2\right)\end{aligned} \tag{2.152}$$

Puisque

$$\int_{\Omega}\frac{1}{(\pi\sigma^2)^{2NT}}\exp\left(\sum_{t=1}^{T}-\frac{1}{\sigma^2}\|\mathbf{x}(t)\|^2\right)d\mathbf{x} = 1, \tag{2.153}$$

on a

$$\zeta(\alpha,\beta) = \exp\left(-\frac{\|\mathbf{s}\|^2}{4\sigma^2}\|\mathbf{a}(\omega+\alpha)-\mathbf{a}(\omega+\beta)\|^2\right). \tag{2.154}$$

Grâce à la structure du vecteur $\mathbf{a}(\omega)$, et sachant que les expression analytiques de $\|\mathbf{a}(\omega+\alpha)-\mathbf{a}(\omega+\beta)\|^2$ sont données par

$$\|\mathbf{a}(\omega+\alpha)\|^2 = \|\mathbf{a}(\omega+\beta)\|^2 = N, \tag{2.155}$$

$$\mathbf{a}(\omega+\alpha)^H\mathbf{a}(\omega+\beta) = \sum_{i=1}^{N}e^{\left(j\frac{2\pi}{\lambda}d_k(\beta-\alpha)\right)}, \tag{2.156}$$

3. Voir **annexe I** pour les résultats sans cette hypothèse.

et par

$$\mathbf{a}(\omega+\beta)^H\mathbf{a}(\omega+\alpha)=\sum_{i=1}^{N}e^{\left(j\frac{2\pi}{\lambda}d_k(\alpha-\beta)\right)}. \tag{2.157}$$

On trouve que les fonctions $\zeta(\alpha,\beta)$ ne dépendent plus du paramètre ω. Par conséquent, $\eta(\alpha,\beta)$ est donné par

$$\eta(\alpha,\beta)=\zeta(\alpha,\beta)\int_{\Theta}\sqrt{p(\omega+\alpha)p(\omega+\beta)}d\omega. \tag{2.158}$$

Sous l'hypothèse d'un distribution *a priori* uniforme, on obtient

$$\int_{\Theta}\sqrt{p(\omega+\alpha)p(\omega+\beta)}=1-\frac{|\alpha|+|\beta|}{2}. \tag{2.159}$$

À partir de (2.149), (2.155), (2.156), (2.157), (2.158) et (2.159), l'expression analytique de la BWW est donnée par (2.160).

$$BWW=\sup_{h}\frac{h^2\left(1-\frac{|h|}{2}\right)^2\exp\left(-\frac{\|\mathbf{s}\|^2}{\sigma^2}\left(N-\sum\limits_{k=1}^{N}\cos\left(\frac{2\pi}{\lambda}d_kh\right)\right)\right)}{2\left(1-\frac{|h|}{2}\right)-2\left(1-|h|\right)\exp\left(-\frac{\|\mathbf{s}\|^2}{2\sigma^2}\left(N-\sum\limits_{k=1}^{N}\cos\left(\frac{4\pi}{\lambda}d_kh\right)\right)\right)}. \tag{2.160}$$

– Modèle stochastique $\mathcal{M}_2$

À partir de (2.148), l'expression analytique de $\zeta(\alpha,\beta)$ est donnée par :

$$\zeta(\alpha,\beta)=\int_{\Omega}\frac{1}{\pi^{2NT}\left|\mathbf{R}(\omega+\alpha)\right|^{T/2}\left|\mathbf{R}(\omega+\beta)\right|^{T/2}}e^{\left(-\sum\limits_{t=1}^{T}\mathbf{y}(t)^H\left(\frac{\mathbf{R}(\omega+\alpha)^{-1}+\mathbf{R}(\omega+\beta)^{-1}}{2}\right)\mathbf{y}(t)\right)}d\mathbf{y} \tag{2.161}$$

En posant $\mathbf{\Gamma}^{-1}=\frac{\mathbf{R}(\omega+\alpha)^{-1}+\mathbf{R}(\omega+\beta)^{-1}}{2}$, on obtient, $|\mathbf{\Gamma}|=\frac{2^{2N}}{|\mathbf{R}(\omega+\alpha)^{-1}+\mathbf{R}(\omega+\beta)^{-1}|}$, ce qui donne

$$\zeta(\alpha,\beta)=\frac{\left|\mathbf{\Gamma}\right|^T}{\left|\mathbf{R}(\omega+\alpha)\right|^{T/2}\left|\mathbf{R}(\omega+\beta)\right|^{T/2}}\times\int_{\Omega}\frac{1}{\pi^{2NT}\left|\mathbf{\Gamma}\right|^T}\exp\left(-\sum_{t=1}^{T}\mathbf{y}(t)^H\mathbf{\Gamma}^{-1}\mathbf{y}(t)\right)d\mathbf{y}. \tag{2.162}$$

Puisque $\int_{\Omega}\frac{1}{\pi^{2NT}|\mathbf{\Gamma}|^T}\exp\left(-\sum\limits_{t=1}^{T}\mathbf{y}(t)^H\mathbf{\Gamma}^{-1}\mathbf{y}(t)\right)d\mathbf{y}=1$, on a

$$\zeta(\alpha,\beta)=\frac{\left|\mathbf{\Gamma}\right|^T}{\left|\mathbf{R}(\omega+\alpha)\right|^{T/2}\left|\mathbf{R}(\omega+\beta)\right|^{T/2}}. \tag{2.163}$$

Grâce à la structure de la matrice $\mathbf{R}(\omega+\delta)=\sigma_s^2\mathbf{a}(\omega+\delta)\mathbf{a}(\omega+\delta)^H+\sigma^2\mathbf{I}_{2N}$, on a

$$\left|\mathbf{R}(\omega+\delta)\right|=\sigma^{4N}\left(1+\frac{\sigma_s^2}{\sigma^2}\left\|\mathbf{a}(\omega+\delta)\right\|^2\right). \tag{2.164}$$

En outre, par l'identité de Woodbury, on obtient

$$\mathbf{R}(\omega+\delta)^{-1}=\frac{1}{\sigma^2}\left(\mathbf{I}_{2N}-\frac{\sigma_s^2\mathbf{a}(\omega+\delta)\mathbf{a}(\omega+\delta)^H}{\sigma_s^2\left\|\mathbf{a}(\omega+\delta)\right\|^2+\sigma^2}\right), \tag{2.165}$$

donc,

$$\mathbf{R}(\omega+\alpha)^{-1}+\mathbf{R}(\omega+\beta)^{-1}=\frac{1}{\sigma^2}\left(2\mathbf{I}_{2N}-\frac{\sigma_s^2\mathbf{a}(\omega+\alpha)\mathbf{a}(\omega+\alpha)^H}{\sigma_s^2\left\|\mathbf{a}(\omega+\alpha)\right\|^2+\sigma^2}-\frac{\sigma_s^2\mathbf{a}(\omega+\beta)\mathbf{a}(\omega+\beta)^H}{\sigma_s^2\left\|\mathbf{a}(\omega+\beta)\right\|^2+\sigma^2}\right). \tag{2.166}$$

Le déterminant de la matrice $\mathbf{R}(\omega+\alpha)^{-1}+\mathbf{R}(\omega+\beta)^{-1}$ est obtenu par une analyse des valeurs propres. En particulier, il y a $2N-2$ valeurs propres qui sont égales à $2/\sigma^2$, et les vecteurs propres correspondant aux deux dernières valeurs propres forment une combinaison linéaire $\mathbf{a}(\omega+\alpha)+q\mathbf{a}(\omega+\beta)$. De plus, ces deux valeurs propres ν sont des solutions de l'équation suivante :

$$\left(\mathbf{R}(\omega+\alpha)^{-1}+\mathbf{R}(\omega+\beta)^{-1}\right)\left(\mathbf{a}(\omega+\alpha)+q\mathbf{a}(\omega+\beta)\right)=\nu\left(\mathbf{a}(\omega+\alpha)+q\mathbf{a}(\omega+\beta)\right), \tag{2.167}$$

ce qui se réduit à

$$\begin{aligned}&\mathbf{a}(\omega+\alpha)\left(\tfrac{1}{\sigma^2}\left(2-A\left\|\mathbf{a}(\omega+\alpha)\right\|^2-qAC\right)-\nu\right)\\&+\mathbf{a}(\omega+\beta)\left(\tfrac{1}{\sigma^2}\left(2q-Bq\left\|\mathbf{a}(\omega+\beta)\right\|^2-BC^H\right)-q\nu\right)=0,\end{aligned} \tag{2.168}$$

où $A=\frac{\sigma_s^2}{\sigma_s^2\|\mathbf{a}(\omega+\alpha)\|^2+\sigma^2}$, $B=\frac{\sigma_s^2}{\sigma_s^2\|\mathbf{a}(\omega+\beta)\|^2+\sigma^2}$ et $C=\mathbf{a}(\omega+\alpha)^H\mathbf{a}(\omega+\beta)$. On obtient l'équation

$$\begin{aligned}&\nu^2\sigma^4+\nu\sigma^2\left(2-A\left\|\mathbf{a}(\omega+\alpha)\right\|^2-2+B\left\|\mathbf{a}(\omega+\beta)\right\|^2\right)\\&-4+2A\left\|\mathbf{a}(\omega+\alpha)\right\|^2+2B\left\|\mathbf{a}(\omega+\beta)\right\|^2\\&-AB\left\|\mathbf{a}(\omega+\alpha)\right\|^2\left\|\mathbf{a}(\omega+\beta)\right\|^2+ABCC^H=0.\end{aligned} \tag{2.169}$$

En résolvant (2.169) pour ν, et vu que

$$\left\|\mathbf{a}(\omega+\alpha)\right\|^2=\left\|\mathbf{a}(\omega+\beta)\right\|^2=\left\|\mathbf{a}(\omega)\right\|^2,$$

on obtient

$$\left|\mathbf{R}(\omega+\alpha)^{-1}+\mathbf{R}(\omega+\beta)^{-1}\right|=\prod_{i=1}^{2N}\nu_i=\frac{2^{2N}}{\sigma^{4N}}\left(\frac{\sigma^2}{\left\|\mathbf{a}(\omega)\right\|^2\sigma_s^2+\sigma^2}+\frac{1}{4}\frac{\sigma_s^4\left(\left\|\mathbf{a}(\omega)\right\|^4-\|C\|^2\right)}{\left(\left\|\mathbf{a}(\omega)\right\|^2\sigma_s^2+\sigma^2\right)^2}\right). \tag{2.170}$$

Finalement, en remplaçant (2.164), (2.170) dans (2.163), on a

$$\zeta(\alpha,\beta)=\left(1+\frac{\sigma_s^2(\left\|\mathbf{a}(\omega)\right\|^4-\left|\mathbf{a}(\omega+\alpha)^H\mathbf{a}(\omega+\beta)\right|^2)}{4\sigma^2(\left\|\mathbf{a}(\omega)\right\|^2\sigma_s^2+\sigma^2)}\right)^{-T}. \tag{2.171}$$

Dans (2.155), (2.156), et (2.157), on trouve que $\zeta(\alpha,\beta)$ ne dépend pas du paramètre ω, comme dans le cas déterministe. Par conséquent, l'expression analytique de la BWW est donnée par (2.172).

$$BWW=\sup_h\frac{h^2\left(1-\frac{|h|}{2}\right)^2\left(1+\frac{\sigma_s^2\left(N^2-\left\|\sum_{k=1}^{N}\exp\left(j\frac{2\pi}{\lambda}d_kh\right)\right\|^2\right)}{4\sigma^2(N\sigma_s^2+\sigma^2)}\right)^{-2T}}{2\left(1-\frac{|h|}{2}\right)-2\left(1-|h|\right)\left(1+\frac{\sigma_s^2\left(N^2-\left\|\sum_{k=1}^{N}\exp\left(j\frac{4\pi}{\lambda}d_kh\right)\right\|^2\right)}{4\sigma^2(N\sigma_s^2+\sigma^2)}\right)^{-T}}. \tag{2.172}$$

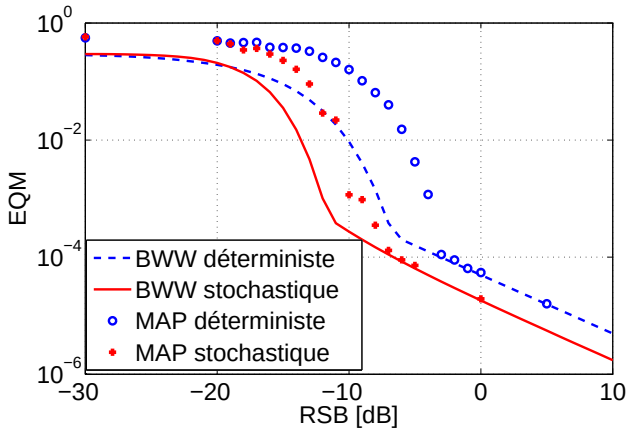


FIGURE 2.24 – MAP par rapport à la BWW.

– Résultats de simulations

On considère un réseau linéaire uniforme composé de $N = 10$ capteurs avec une distance entre capteurs de $\lambda/2$. Le nombre des observations est égal à $T = 20$. Enfin, l'EQM empirique de l'estimateur du maximum *a posteriori* est réalisé à partir de 1000 tirages de Monte Carlo. La Fig. 2.24 montre que la BWW donne une bonne approximation du décrochement de l'estimateur du maximum *a posteriori* pour les deux modèles de signaux considérés ici.

2.3.3.2 Estimation de la phase des signaux de pulsar

Une solution potentielle à la navigation spatiale autonome réside dans l'utilisation des signaux de pulsar à rayons X. Le challenge réside en l'estimation précise de la phase du signal du pulsar arrivant sur le detecteur. Les études récentes à ce sujet ont montrées que le phénomène de décrochement apparaît lui aussi pour l'EQM des estimateurs. C'est pourquoi, dans le cadre de la thèse de Nguyen Duy Tran, nous avons proposés plusieurs formes analytiques de bornes inférieures de l'EQM (déterministes et bayésiennes) afin de prédire ce phénomène. Les détails sont donnés en **appendix H**.

Chapitre 3

Conclusions et perspectives

Ma thématique générale de recherche concerne l'estimation et la détection de paramètres (généralement physiques) à partir d'un jeu de données bruitées obtenues à l'aide d'un réseau de capteurs (c'est-à-dire en exploitant à la fois la diversité temporelle et spatiale). En particulier, je me focalise sur les performances ultimes que de tels systèmes peuvent atteindre en utilisant principalement des outils statistiques appelés bornes inférieures de l'erreur quadratique moyenne.

Dans ce contexte, mes recherches s'articulent autour des axes suivants :

– L'étude des performances asymptotiques (en nombre d'observations et/ou en Rapport Signal sur Bruit(RSB)) de l'estimateur du maximum de vraisemblance pour le traitement d'antenne. Le traitement d'antenne concerne l'estimation des directions d'arrivées (localisation) de signaux multiples à l'aide d'un réseau de capteurs. (i) Dans le cadre du maximum de vraisemblance stochastique, c'est-à-dire lorsque les signaux sources sont modelisés par un processus aléatoire gaussien, nous avons démontré que l'estimateur pour les directions d'arrivées est non-gaussien et non-efficace (il n'atteint pas la borne de Cramér-Rao stochastique) lorsque le nombre d'observations T est fini et le RSB infini [J2][CI1][CN1]. Ce résultat est valide quel que soit le nombre de sources et quel que soit la géométrie de l'antenne. Dans le cas d'un scenario à une seule source, nous avons démontré que la loi de l'estimateur est une loi de student à $2T$ degrés de liberté et que la variance de l'estimé était égale à $T/(T-1)$ fois la borne de Cramér-Rao stochastique. Dans le cas d'un scenario à deux sources nous avons établi une forme analytique de la matrice de covariance des estimés [CI2]. (ii) Lorsque les signaux sources sont supposés déterministes, nous avons établi la gaussianité et l'efficacité asymptotique en RSB de l'estimateur du maximum de vraisemblance déterministe quel que soit le nombre de sources et quel que soit la géométrie de l'antenne [J1]. Ce travail a été réalisé en collaboration avec Eric Chaumette, ONERA Palaiseau. (iii) Dans le cadre de la thèse de Mohammed Nabil El Korso, commencée en septembre 2008, nous avons calculé et analysé les bornes de Cramér-Rao stochastiques et déterministes dans le contexte de sources situées en champs proche, c'est-à-dire, lorsque les fronts d'ondes ne peuvent plus être considérés comme plans [J9][CI9][CN5].

– Le développement théorique et l'analyse de nouvelles bornes inférieures pouvant prendre en compte des scénarios non standards (c'est-à-dire lorsque la classique borne de Cramér-Rao n'est plus valide). Ces scénarios comprennent des situations où le nombre d'observations est faible et où le rapport signal sur bruit est faible. (i) Nous avons unifié les bornes Bayésiennes de la famille Weiss-Weinstein à l'aide d'un problème d'optimisation sous contraintes [J7][CI3][CN2]. Particulièrement, nous avons établi un pont entre les bornes susmentionnées

et la “meilleure” borne bayésienne : l'erreur quadratique moyenne (EQM) de l'estimateur de la moyenne conditionnelle. Ceci nous a permis d'établir une relation d'ordre entre les différentes bornes souvent inaccessible dans les développements traditionnels exploitant une inégalité de covariance. (ii) Cette unification nous a permis de proposer une nouvelle borne Bayésienne plus précise qu'une large classe d'autres bornes pour une charge de calcul équivalente [CI4]. Ces différents travaux ont été réalisés en collaboration avec Christ Richmond du MIT Lincoln Laboratory (USA) et Arye Nehorai de Washington University in St. Louis (USA). (iii) Nous avons proposé une modification des bornes classiques dans le cas où le processus d'estimation est conditionné à un test de détection ce qui est généralement le cas dans les systèmes radar où la localisation d'une cible est opérée seulement après une phase préalable de détection [CI5]. (iv) Nous nous sommes intéressés à l'obtention de nouvelles bornes inférieures par le biais d'un mélange de contraintes linéaires et non-linéaires. Ces premiers résultats nous ont permis de montrer que toutes les bornes inférieures de l'EQM peuvent être rendues plus précises à l'aide d'une simple transformation de la fonction de vraisemblance [CI8]. Nous avons également proposé une borne conjecturale issue du contexte Bayésien pour analyser les performances non-asymptotiques des estimateurs non-Bayésien [CI13]. Ces deux derniers points ont été réalisés en collaboration avec Eric Chaumette, ONERA Palaiseau. (v) Ces résultats théoriques ont été appliqués pour l'étude non-asymptotique des estimateurs de fréquences porteuses en communications numériques entraînées pour laquelle nous avons proposé une forme simplifiée de la borne d'Abel et de la borne de Weiss-Weinstein [J3][J4].

– L'utilisation de ces bornes inférieures non plus comme benchmark pour connaître les performances possibles pour un estimateur, mais pour le design optimal des systèmes multi-capteurs (en collaboration avec Sylvie Marcos et Rémy Boyer). (i) Dans le cadre de la thèse de Dinh Thang VU, commencée en septembre 2008 (projet DIGITEO), nous avons étudié l'impact de la géométrie d'un réseau de capteurs sur la localisation d'une ou plusieurs sources [J13][CI11][CN4]. Ces travaux ayant commencé dans le domaine asymptotique (en nombre d'observations et/ou en RSB), nous avons poursuivi cette étude dans le domaine non-asymptotique par le biais de l'étude de la borne de Weiss-Weinstein [CI12][CI18](et [J19] soumis). (ii) Dans le cadre de la thèse de Nguyen Duy TRAN, commencée en septembre 2009 (projet DIGITEO), nous travaillons sur le concept de radar MIMO. En particulier, nous avons montré que les performances de l'estimateur du maximum de vraisemblance tendent vers une limite finie en présence d'erreurs de modèle sur la géométrie de l'antenne radar [CI16][CN6] contrairement au cas classique (sans erreurs de modèle) où la précision des estimateurs devient infinie. Dans un cadre non-asymptotique, nous avons proposé une borne inférieure de l'erreur quadratique moyenne globale que nous avons utilisée pour optimiser la géométrie de l'antenne ainsi que celui des formes d'ondes ([J16] soumis).

– L'application des bornes inférieures de l'erreur quadratique moyenne à des problèmes d'estimation non-standard, c'est-à-dire hors cadre asymptotique Gaussien i.i.d., lorsque la classique borne de Cramér-Rao n'existe pas ou encore lorsqu'elle permet d'obtenir d'autres renseignements sur le système étudié (résolution et convergence). (i) Nous avons travaillé sur le thème des bornes minimales récursives pour les communications numériques dans le cadre de modèles non i.i.d. [J6][CI6][CN3]. (ii) L'hypothèse classique d'un bruit Gaussien pour les mesures est erronée dans le contexte radar puisque celui-ci doit faire face à l'environnement situé tout autour de l'objet à détecter. Les radars sont alors perturbés par les échos indésirables des réflecteurs constitutifs du sol, du sur-sol ou encore de la mer. Dans ce contexte, avec Fréderic Pascal du laboratoire SONDRA (Supelec), nous avons déterminé

les propriétés statistiques de l'estimateur du maximum de vraisemblance ainsi que la borne de Cramér-Rao pour l'estimation de matrice de covariance lorsque le bruit (appelé clutter dans le contexte radar) suit une distribution de Bessel-K [J8]. (iii) Avec Arye Nehorai, Washington University in St. Louis (USA) et Carlos Muravchik, University of La Plata (Argentine), un travail conjoint a porté sur le thème de la détection de contractions chez la femme enceinte. Le but est, à l'aide d'un réseau de capteurs, de mesurer le signal mécanomyographique et de détecter les différents points de ruptures (contractions) dans ce signal. Le challenge vient ici du fait que la borne de Cramér-Rao ne peut pas être calculée dans ce genre de situation car la fonction de vraisemblance intervenant dans celle-ci n'est pas dérivable. Nous avons donc proposé une étude mettant en jeu la borne de Barankin dont l'obtention ne nécessite pas une telle régularité. Cette étude a été réalisée dans le contexte de données Gaussiennes multivariées avec changement de moyenne et/ou covariance ainsi que dans le cadre de processus Poissonniens [J10][CI7]. (iv) Avec Cédric Herzet, INRIA, nous avons prouvé que la convergence de certains algorithmes d'approximation du maximum de vraisemblance pouvait être calculée à l'aide de différents types de bornes de Cramér-Rao [J5]. (v) Dans le cadre de la thèse de Mohammed Nabil El Korso, nous nous sommes attaqués aux liens entre la borne de Cramér-Rao et les limites de résolutions statistiques des estimateurs. Après avoir appliqué ces concepts aux cas de sources polarisées [J11][CI14], dans un context MIMO [J14][CI15], nous avons étendu de manière théorique la notion de resolution statistique au cas multidimensionnel (plusieurs paramètres d'intérêts par signaux et plusieurs signaux) [CI10][CI19]. Nous avons appliqué cette extension au domaine de la localisation de sources en champ proche [J15] ainsi que pour le modèle multi-harmoniques [J12]. (vi) Enfin, dans le cadre de la thèse de Nguyen Duy Tran, nous avons montré le potentiel des bornes autres que la borne de Cramér-Rao pour des domaines autres que celui du traitement d'antenne. Nous avons commencé à étudier le problème de l'estimation de la phase d'objets astrophysiques appelé Pulsar à rayon X. Un Pulsar est une étoile à neutrons tournant très rapidement sur elle-même. La période de rotation d'un Pulsar présente une stabilité plus forte que les horloges atomiques et pourrait donc s'avérer utile afin d'éviter d'embarquer une horloge atomique dans un satellite. En conséquence, l'estimation de cette période de rotation constitue une piste intéressante dans le domaine des technologies spatiales. Les résultats récents de la littérature (Emadzadeh et Speyer, *IEEE* Transactions on Signal Processing 2010) montrent que le modèle des observations suit un processus Poissonnien parametré par la période de rotation. Si quelques exemples d'estimateurs (sous-optimaux) sont déjà disponibles, l'étude des performances de tels estimateurs n'a pour le moment pas été entreprise. De plus, des simulations numériques ont clairement montré que ces estimateurs étaient sujet au phénomène de décrochement. En conséquence, nous avons calculé les bornes de Barankin et de Weiss-Weinstein ([J17] soumis) pour ce problème et nous espérons que l'utilisations de ces bornes inférieures donneront des résultats utiles aux astrophysiciens pour quantifier les performances ultimes des tels systèmes.

Depuis mon post-doctorat (2006-2007) et mon recrutement en temps que Maître de Conférence en octobre 2007, mon activité de recherche à principalement consisté à montrer à la communauté scientifique l'utilité des bornes inférieures de l'erreur quadratique moyenne pour le traitement du signal. En particulier, j'ai souhaité accentuer mes recherches vers le fait que ces bornes ne sont pas de simple outils théoriques mais qu'elles peuvent être utiles en pratique dans les systèmes multi-capteurs du futur. Il s'agit de travaux de nature plus appliqués que ceux effectués durant ma thèse que je souhaite terminer (perspectives à court terme) pour me réorienter progressivement vers les outils théoriques (perspectives à long terme). Dans ce contexte, mes

perspectives de recherches sont donc les suivantes :

- *Court terme* : (i) Curieusement, le premier résultat que j'ai obtenu durant ma thèse (ainsi que mon stage de DEA) était un non résultat. En effet, comme dit préalablement, dans le cadre du traitement d'antenne et du maximum de vraisemblance stochastique, c'est-à-dire lorsque les signaux sources sont modelisés par un processus aléatoire gaussien, nous avons démontré que l'estimateur pour les directions d'arrivées est non-gaussien et non-efficace (il n'atteint pas la borne de Cramér-Rao stochastique) lorsque le nombre d'observations T est fini et le RSB infini [J2][CI1][CN1]. Il s'agit d'un non résultat car nous n'avons pas été en mesure de trouver la densité de probabilité de cet estimateur et donc d'en déduire sa matrice de covariance. J'ai récemment découvert que cette loi semble reliée à une loi complexe circulaire Beta multivariée. (ii) Dans le contexte non-Bayésien et Bayésien, de nouvelles bornes ont récemment été proposées dans la littérature (Todros et Tabrikian, *IEEE* Transactions on Information Theory 2010). Nous étudions actuellement ces résultats afin de les replacer dans le cadre unifié d'un problème d'optimisation sous contraintes que nous avions proposé dans [J7][CI3] et [CN2]. Grâce à ce cadre, nous avons montré que ces bornes sont obtenues en appliquant une transformation de Fourier aux contraintes des bornes classiques. Nous cherchons donc à utiliser d'autres transformations unitaires à appliquer aux contraintes afin de diminuer la charge de calcul. De plus, nous cherchons à appliquer ces bornes au problème d'estimation de directions d'arrivées des sources situées en champs proches (un résultat à déjà été publié dans [CN7]). De tels résultats nous permettraient d'étendre le travaux décrit dans [J9] pour l'étude des plages non-asymptotiques des estimateurs. (iii) L'un des point clé du projet Digiteo TIMuCa auquel je participe concerne la séléction de formes d'ondes dans le contexte des radars MIMO. C'est un point que nous avons encore peu abordé dans le cadre de la thèse de Nguyen Duy Tran et que nous allons donc poursuivre. Le but sera de proposer des critères de sélection des formes d'ondes ou de certains de leur paramètres afin d'améliorer les performances d'estimation en termes d'erreur quadratique moyenne de tels systèmes. Nous nous focaliserons dans un premier temps sur les bornes minimales d'estimation récursives ce qui nous permettra d'obtenir des critères optimaux en terme d'EQM, c'est-à-dire qui conduiront à une meilleure estimation des paramètres physiques de ou des objets présents dans l'environnement (distance, vitesse, position, etc.) pour la localisation, la poursuite et l'identification. Puis nous étudierons des critères optimaux en termes de probabilité de détection et de fausse alarme.

- *Long terme* : (i) L'inégalité de Cramér-Rao et toutes les bornes inférieures de l'erreur quadratique moyenne basées sur une inégalité de covariance en générale ne sont en fait qu'une (petite) partie des travaux de Fisher, Rao, Fréchet, Darmois, Cramér, etc dont la portée était plus générale. Par exemple, dans l'article de Rao de 1945 (Bull.Calcutta Math. Soc.), la moitié de l'article a pour objet la borne de Cramér-Rao et la matrice de Fisher, mais son objectif final concernait la deuxième partie de l'article sur la géométrie de l'information. La matrice d'information de Fisher apparaît dans ce contexte comme la métrique d'une géométrie Riemannienne et nous donne une indication sur la déformation de la variété dans l'espace des paramètres. Il s'agit d'un domaine qui provient à la base des statistiques et des mathématiques en général, mais depuis environ 10 ans, les applications en traitement du signal ont connu une forte croissance. On citera, par exemple, l'analyse de flux audio, la détection radar par moyennage de la matrice de covariance des observations dans l'espace des matrices Hermitiennes, ou encore la vision par ordinateur. Dans ce cadre, la démonstration d'une inégalité de type Cramér-Rao sur une variété arbitraire (qui n'est pas munie d'un ensemble de coordonnées intrinsèques) a été proposée en 2005 (S. Smith, *IEEE* Tran-

sactions on Signal Processing) et utilisée avec succès en traitement d'antenne. L'extension aux autres bornes inférieures de l'erreur quadratique moyenne reste un problème ouvert. Dans le cadre de la géométrie de l'information, un outil fondamental est la métrique de Fisher basée sur l'information de Fisher. La borne de Cramér-Rao est simplement l'inverse de l'information de Fisher. Il paraît donc naturel que les matrices d'informations utilisées pour l'établissement des autres bornes inférieures de l'erreur quadratique moyenne puissent être vues comme une métrique au même titre que la métrique de Fisher. Si nous arrivons à mettre en place ces métriques, un exemple pratique d'application sera encore une fois le traitement d'antenne où une utilisation de la géométrie différentielle pour l'optimisation de la position des réseaux de capteurs a déjà été suggérée dans Manikas, Differential Geometry in Array Processing, Imperial College Press, 2004. Il reste cependant beaucoup de zones d'ombre à explorer sur ce sujet. (ii) Avec Patricio La Rosa et Arye Nehorai (Washington university in St. Louis) et Carlos Muravchik (University of La Plata) nous travaillons actuellement sur l'application des bornes Bayésiennes au problème d'estimation de points de ruptures. En effet, dans [J10] et [CI7], nous avons déjà appliqué la borne de Barankin (qui est une borne déterministe) pour ce type de problème, mais cette borne apparaît comme optimiste pour l'application visée (détection de contraction chez la femme enceinte). Nous avons identifié que le problème venait du fait que cette borne s'applique pour des estimateurs non biaisés tandis que l'estimateur du maximum de vraisemblance pour ce problème présente un biais. Afin de lever cette contrainte, il nous a donc semblé judicieux de nous placer dans un contexte Bayésien où les bornes inférieures de l'erreur quadratique moyenne ne nécessitent pas une telle hypothèse. Mais il est intéressant de noter que, pour le moment, nous n'avons travaillé que sur des séquences indépendantes et identiquement distribuées ainsi que lorsque les paramètres des densités de probabilités étaient connus. Il reste donc beaucoup de travaux à réaliser afin d'obtenir des limites inférieures de l'erreur quadratique moyenne des estimateurs à distance finie (plusieurs résultats sont déjà disponible dans la littérature concernant les performances asymptotiques) par exemple lorsque les processus entre chaque points de ruptures suivent un modèle autorégressif ou GARCH. De plus, les algorithmes actuels dans ce domaine ne ce cantonnent pas à la simple estimation des positions des points de ruptures mais aussi à leurs détection (leurs nombre). A ce sujet, je pense que les travaux entrepris avec Eric Chaumette sur les bornes inférieures de l'erreur quadratique moyenne conditionnée à un test de détection [CI5] nous seront alors d'une grande utilité lorsque nous nous attaquerons à ce type de problème.

Bibliographie

[Abe90] J. S. Abel. A bound on mean square estimate error. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 3, pages 1345–1348, Albuquerque, NM, USA, April 1990.

[Abe93] J. S. Abel. A bound on mean square estimate error. *IEEE Transactions on Information Theory*, 39(5) :1675–1680, September 1993.

[Abe06] H. Abeida. *Imagerie d'antenne pour signaux non circulaires : bornes de performance et algorithmes*. PhD thesis, université Paris 6, France, November 2006.

[AD08] H. Abeida and J.-P. Delmas. Statistical performance of MUSIC-like algorithms in resolving noncircular sources. *IEEE Transactions on Signal Processing*, 56(6) :4317–4329, September 2008.

[AGGS96] Y. Abramovich, D.A. Gray, A.Y. Gorokhov, and N.K. Spencer. Comparaison of DOA estimation performance for various types of sparse antenna array geometries. In *EUSIPCO*, pages 1968–1972, Trieste, Italy, September 1996.

[ASG99] Y. I. Abramovich, N. K. Spencer, and A. Y. Gorokhov. Positive-definite toeplitz completion in DOA estimation for nonuniform linear antenna arrays. II : Partially augmentable arrays. *IEEE Transactions on Signal Processing*, 47(6) :1502–1521, June 1999.

[Ath01] F. Athley. Optimization of element positions for direction finding with sparse arrays. In *Proc. of IEEE Workshop on Statistical Signal Processing (SSP)*, volume 1, pages 516–519, 2001.

[Ath05] F. Athley. Threshold region performance of maximum likelihood direction of arrival estimators. *IEEE Transactions on Signal Processing*, 53(4) :1359–1373, April 2005.

[AW08] A. Amar and A.J. Weiss. Fundamental limitations on the resolution of deterministic signals. *IEEE Transactions on Signal Processing*, 56(11) :5309–5318, November 2008.

[Bar49] E. W. Barankin. Locally best unbiased estimates. *The Annals of Mathematical Statistics*, 20(4) :477–501, December 1949.

[BC07] A. Barbieri and G. Colavolpe. On the Cramér-Rao bound for carrier frequency estimation in the presence of phase noise. 6(2) :575–582, February 2007.

[Beh90] R.T. Behrens. *Subspace signal processing in structured noise*. PhD thesis, University of Colorado, Boulder, US-CO, 1990.

[Bel95] K. L. Bell. *Performance bounds in parameter estimation with application to bearing estimation*. PhD thesis, George Mason University, Fairfax, VA, USA, 1995.

[BEV96] K. L. Bell, Y. Ephraim, and H. L. Van Trees. Explicit Ziv-Zakaï lower bound for bearing estimation. *IEEE Transactions on Signal Processing*, 44(11) :2810–2824, November 1996.

[Bha46] A. Bhattacharyya. On some analogues of the amount of information and their use in statistical estimation. *Sankhya Indian J. of Stat.*, 8 :1–14, 201–218, 315–328, 1946.

[BM03] U. Baysal and R. L. Moses. On the geometry of isotropic arrays. *IEEE Transactions on Signal Processing*, 51 :1469–1477, 2003.

[BMWZ87] B. Z. Bobrovsky, E. Mayer-Wolf, and M. Zakaï. Some classes of global Cramér-Rao bounds. *The Annals of Statistics*, 15(4) :1421–1438, December 1987.

[Böh84] J. F. Böhme. Estimation of source parameters by maximum likelihood and non linear regression. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 9, pages 271–274, San Diego, CA, USA, March 1984.

[Böh86] J. F. Böhme. Estimation of spectral parameters of correlated signals in wavefields. *ELSEVIER Signal Processing*, 11(4) :329–337, December 1986.

[Boy08] R. Boyer. Deterministic asymptotic Cramér-Rao bound for the multidimensional harmonic model. *Signal Processing*, 88 :2869–2877, 2008.

[BS94] R. T. Behrens and L. L. Scharf. Signal processing applications of oblique projection operators. *IEEE Transactions on Signal Processing*, 42(6) :1413–1424, June 1994.

[BSET97] K. Bell, Y. Steinberg, Y. Ephraim, and H. L. Van Trees. Extended Ziv Zakai lower bound for vector parameter estimation. *IEEE Transactions on Signal Processing*, 43 :624–638, March 1997.

[BT74] S. Bellini and G. Tartara. Bounds on error in signal parameter estimation. *IEEE Transactions on Communications*, 22(3) :340–342, March 1974.

[BT06] I. Bekkerman and J. Tabrikian. Target detection and localization using MIMO radars and sonars. *IEEE Transactions on Signal Processing*, 54 :3873–3883, October 2006.

[BZ76] B. Z. Bobrovsky and M. Zakaï. A lower bound on the estimation error for certain diffusion processes. *IEEE Transactions on Information Theory*, 22(1) :45–52, January 1976.

[Cap69] J. Capon. High resolution frequency wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8) :1408–1418, August 1969.

[CGQL08] E. Chaumette, J. Galy, A. Quinlan, and P. Larzabal. A new Barankin bound approximation for the prediction of the threshold region performance of maximum likelihood estimators. *IEEE Transactions on Signal Processing*, 56(11) :5319–5333, November 2008.

[CGV+07] E. Chaumette, J. Galy, F. Vincent, A. Renaux, and P. Larzabal. Mse bounds conditioned by the energy detector. In *Proc. of European Signal Processing Conference (EUSIPCO)*, Poznan, Poland, 2007.

[Cho04] Y. H. Choi. Unified approach to Cramer-Rao bounds in direction estimation with known signal structures. *Signal Processing*, 84(10) :1875 – 1882, 2004.

[CM97] M. Cedervall and R. L. Moses. Efficient maximum likelihood DOA estimation for signals with known waveforms in presence of multipath. *IEEE Transactions on Signal Processing*, 45(3) :808–811, March 1997.

[Cox73] H. Cox. Resolving power and sensitivity to mismatch of optimum array processors. *J. Acoust. Soc. Am*, 54(3) :771–785, 1973.

[CR51] D. G. Chapman and H. Robbins. Minimum variance estimation without regularity assumptions. *The Annals of Mathematical Statistics*, 22(4) :581–586, December 1951.

[Cra46] H. Cramér. *Mathematical Methods of Statistics*, volume 9 of *Princeton Mathematics*. Princeton University Press, New-York, September 1946.

[CZZ75] D. Chazan, M. Zakaï, and J. Ziv. Improved lower bounds on signal parameter estimation. *IEEE Transactions on Information Theory*, 21(1) :90–93, January 1975.

[DA06] J.-P. Delmas and H. Abeida. Statistical resolution limits of DOA for discrete sources. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 4, pages 889–892, Toulouse, France, 2006.

[Dar45] G. Darmois. Sur les limites de la dispersion de certaines estimations. *Revue de l'Institut International de Statistique*, 13(1/4) :9–15, 1945.

[Dil98] E. Dilaveroglu. Nonmatrix Cramér-Rao bound expressions for high-resolution frequency estimators. *IEEE Transactions on Signal Processing*, 46(2) :463–474, February 1998.

[DMR00] A. Demir, A. Mehrotra, and J. Roychowdhury. Phase noise in oscillators : a unifying theory and numerical methods for characterization. 47 :655–674, May 2000.

[Dug37] D. Dugué. Application des propriétés de la limite au sens du calcul des probabilités à l'étude des diverses questions d'estimation. *Journal de l'Ecole Polytechnique*, 3(4) :305–372, 1937.

[EBBM09] M. N. El Korso, G. Bouleux, R. Boyer, and S. Marcos. Sequential estimation of the range and the bearing using the zero-forcing MUSIC approach. In *Proc. EUSIPCO*, pages 1404–1408, Glasgow, Scotland, August 2009.

[EBRM09] M. N. El Korso, R. Boyer, A. Renaux, and S. Marcos. Nonmatrix closed-form expressions of the Cramér-Rao bounds for near-field localization parameters. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Taipei, Taiwan, 2009.

[EBRM10] M. N. El Korso, R. Boyer, A. Renaux, and S. Marcos. Conditional and unconditional Cramér-Rao bounds for near-field source localization. *IEEE Transactions on Signal Processing*, 58(5) :2901–2907, May 2010.

[EBRM11a] M. N. El Korso, R. Boyer, A. Renaux, and S. Marcos. A GLRT-based framework for the multidimensional statistical resolution limit. In *Proc. of IEEE Workshop on Satistical Signal Processing*, Nice, France, June 2011.

[EBRM11b] M. N. El Korso, R. Boyer, A. Renaux, and S. Marcos. Statistical resolution limit for source localization in a MIMO context. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Prague, Czeh, May 2011.

[FBLR03] P. Forster, E. Boyer, P. Larzabal, and A. Renaux. Non-efficacité et non-Gaussianité asymptotiques d'un estimateur du maximum de vraisemblance à fort rapport signal sur bruit. In *Proc. du colloque GRETSI sur le traitement du signal et des images*, volume 1, pages 125–128, Paris, FR, September 2003.

[FC09] A. Ferréol and P. Chevalier. High resolution direction finding : from performance toward antenna array optimization -the mono-source case. In *Proc. of European Signal Processing Conference (EUSIPCO)*, pages 1973–1977, Glasgow, Scotland, August 2009.

[FG52] D. A. S. Fraser and I. Guttman. Bhattacharyya bounds without regularity assumptions. *The Annals of Mathematical Statistics*, 23(4) :629–632, December 1952.

[Fis22] R. A. Fisher. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 222 :309–368, 1922.

[FL02] Ph. Forster and P. Larzabal. On lower bounds for deterministic parameter estimation. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Orlando, Fl, 2002.

[FLV08] A. Ferreol, P. Larzabal, and M. Viberg. On the resolution probability of MUSIC in presence of modeling errors. *IEEE Transactions on Signal Processing*, 56(5) :1945–1953, May 2008.

[Fre43] M. Frechet. Sur l'extension de certaines évaluations statistiques au cas de petit échantillons. *Revue de l'Institut International de Statistique*, 11(3/4) :182–205, 1943.

[FT08] T. Filik and T. E. Tuncer. Uniform and nonuniform V-shaped isotropic planar arrays. In *Proc. of IEEE Workshop on Sensor Array and Multi-channel Processing (SAM)*, pages 99–103, Darmstadt, Germany, July 2008.

[GAM09] H. Gazzah and K. Abed-Meraim. Optimum ambiguity free directional and omni directional planar antenna arrays for DOA estimation. *IEEE Transactions on Signal Processing*, 57(10) :3942–3953, October 2009.

[GAMH05] E. Grosicki, K. Abed-Meraim, and Y. Hua. A weighted linear prediction method for near-field source localization. *IEEE Transactions on Signal Processing*, 53 :3651–3660, 2005.

[GC81] L. C. Godara and A. Cantoni. Uniqueness and linear independence of sterring vectors in array space. 70(2) :467– 475, August 1981.

[Gla72] F. E. Glave. A new look at the Barankin lower bound. *IEEE Transactions on Information Theory*, 18(3) :349–356, May 1972.

[GM06] H. Gazzah and S. Marcos. Cramér-Rao bounds for antenna array design. *IEEE Transactions on Signal Processing*, 54(1) :336–345, January 2006.

[God97] L.C. Godara. Applications of antenna arrays to mobile communications : II. Beamforming and direction of arrival considerations. 85(8) :1195–1245, August 1997.

[GS05] A.B. Gershman and N.D. Sidiropoulos. *Space-time processing for MIMO communications*. Wiley, New York, 2005.

[GSPL02] A.B. Gershman, P. Stoica, M. Pesavento, and E.G. Larsson. Stochastic Cramér-Rao bound for direction estimation in unknown noise fields. *IEE Proceedings-Radar, Sonar and Navigation*, 149 :2–8, January 2002.

[GW91] M. Gavish and A. J. Weiss. Array geometry for ambiguity resolution in direction finding. 44(6) :143– 146, February 1991.

[Ham50] J. M. Hammersley. On estimating restricted parameters. *Journal of the Royal Statistical Society. Series B (Methodological)*, 12(2) :192–240, 1950.

[Hay85] S. Haykin. *Array signal processing*. Englewood Cliffs, NJ, Prentice-Hall, 1985.

[HB91] Y. D. Huang and M. Barkat. Near-field multiple source localization by passive sensor array. 39 :968–975, 1991.

[HN98] M. Haardt and J.A. Nossek. Simultaneous schur decomposition of several nonsymmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems. *IEEE Transactions on Signal Processing*, 46(1) :161–169, January 1998.

[HS91] Y. Hua and T. K. Sarkar. A note on the Cramér-Rao bound for 2-D direction finding based on 2-D array. *IEEE Transactions on Signal Processing*, 39(5) :1215–1218, May 1991.

[HSW91] Y. Hua, T. K. Sarkar, and D. D. Weiner. An L-shaped array for estimating 2D directions of wave arrival. 39 :143– 146, February 1991.

[JLL09] M. Jin, G. Liao, and J. Li. Joint DOD and DOA estimation for bistatic MIMO radar. *Elsevier Signal Processing*, 2 :244–251, February 2009.

[JStB01] T. Jiang, N.D. Sidiropoulos, and J.M.F. ten Berge. Almost-sure identifiability of multidimensional harmonic retrieval. *IEEE Transactions on Signal Processing*, 49(9) :1849–1859, September 2001.

[Kay93] S. M. Kay. *Fundamentals of Statistical Signal Processing : Estimation Theory*, volume 1. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, March 1993.

[Kay98] S. M. Kay. *Fundamentals of Statistical Signal Processing : Detection Theory*, volume 2. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, January 1998.

[KB86] M. Kaveh and A. Barabell. The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise. In *Proc. ASSP Workshop on Spectrum Estimation and Modeling*, volume 34, pages 331–341, 1986.

[KBRM11a] M. N. El Korso, R. Boyer, A. Renaux, and S. Marcos. Statistical resolution limit for the multidimensional harmonic retrieval model : Hypothesis test and cramér-rao bound approaches. *EURASIP Journal on Advances in Signal Processing, special issue on Advances in Angle-of-Arrival and Multidimensional Signal Processing for Localization and Communications*, (5), May 2011.

[KBRM11b] M. N. El Korso, R. Boyer, A. Renaux, and S. Marcos. Statistical resolution limit of the uniform linear cocentered orthogonal loop and dipole array. *IEEE Transactions on Signal Processing*, 59(1) :425–431, January 2011.

[KG09] J. Kusuma and V.K. Goyal. On the accuracy and resolution of powersum-based sampling methods. *IEEE Transactions on Signal Processing*, 57(1) :182–193, January 2009.

[Kie52] J. Kiefer. On minimum variance estimators. *The Annals of Mathematical Statistics*, 23(4) :627–629, December 1952.

[KT83] R. Kumaresan and D. W. Tufts. Estimating the angles of arrival of multiple plane waves. *IEEE Transactions on Aerospace and Electronic Systems*, AES-19(1) :134–139, January 1983.

[KV96] H. Krim and M. Viberg. Two decades of array signal processing research : the parametric approach. *IEEE Signal Processing Magazine*, 13(4) :67–94, July 1996.

[LC91] J. Li and R. Compton. Angle and polarization estimation using esprit with a polarizationsensitive array. 39(9) :1376–1383, September 1991.

[LC03a] J. Lebrun and P. Comon. An algebraic approach to blind identification of communication channels. In *Seventh International Symposium on Signal Processing and Its Applications.*, 2003.

[LC03b] E. L. Lehmann and G. Casella. *Theory of Point Estimation*. Springer Texts in Statistics. Springer, New-York, NY, USA, 2 edition, September 2003.

[LC93] J. Li and R. T. Compton. Maximum likelihood angle estimation for signals with known waveforms. *IEEE Transactions on Signal Processing*, 41(9) :2850–2862, September 93.

[Lee92] H. B. Lee. The Cramér-Rao bound on frequency estimates of signals closely spaced in frequency. *IEEE Transactions on Signal Processing*, 40(6) :1507–1517, 1992.

[Lee94] H. B. Lee. The Cramér-Rao bound on frequency estimates of signals closely spaced in frequency (unconditional case). *IEEE Transactions on Signal Processing*, 42(6) :1569–1572, 1994.

[LHSV95] J. Li, B. Halder, P. Stoica, and M. Viberg. Computationally efficient angle estimation for signals with known waveforms. *IEEE Transactions on Signal Processing*, 43(9) :2154–2163, September 1995.

[LJ92] J. T. H. Lo and S. L. Marple. Jr. Observability conditions for multiple signal direction finding and array sensor localization. *IEEE Transactions on Signal Processing*, 40(11) :2641– 2650, November 1992.

[LL93] H. B. Lee and F. Li. Quantification of the difference between detection and resolution thresholds for multiple closely spaced emitters. *IEEE Transactions on Signal Processing*, 41(6) :2274–2277, 1993.

[LN07] Z. Liu and A. Nehorai. Statistical angular resolution limit for point sources. *IEEE Transactions on Signal Processing*, 55(11) :5521–5527, November 2007.

[LS09] K. W. K. Lui and H. C. So. A study of two-dmensional sensor placement using time-difference-of-arrival measurements. *Digital Signal Processing*, 19 :650– 659, 2009.

[LSZ96] J. Li, P. Stoica, and D. Zheng. Efficient direction and polarization estimation with a cold array. 44(4) :539–547, April 1996.

[LV99] A. Leshem and A-J. Van der Veen. Direction-of-arrival estimation for constant modulus signals. *IEEE Transactions on Signal Processing*, 47(11) :3125–3129, November 1999.

[LVT89] F. Li, R. J. Vaccaro, and D. W. Tuft. Min-Norm linear prediction for arbitrary sensor arrays. volume 4, pages 2613–2616, Glasgow, UK, May 1989.

[Man04] A. Manikas. *Differential Geometry in Array Processing.* Imperial College Press, 2004.

[Mar98] S. Marcos, editor. *Les Méthodes à Haute Résolution : Traitement d'Antenne et Analyse Spectrale.* Hermès, Paris, FR, February 1998.

[McN94] J. A. McNeill. *Jitter in Ring Oscillators.* PhD thesis, Boston University, 1994.

[MD01] Y. Meurisse and J.P. Delmas. Bounds for sparse planar and volume arrays. *IEEE Transactions on Signal Processing*, 47 :464–468, January 2001.

[Mes06] H. Messer. The hybrid Cramér-Rao lower bound – from practice to theory. In *Proc. of IEEE Workshop on Sensor Array and Multi-channel Processing (SAM)*, pages 304–307, Waltham, MA, USA, July 2006.

[MH71] R. J. McAulay and E. M. Hofstetter. Barankin bounds on parameter estimation. *IEEE Transactions on Information Theory*, 17(6) :669–676, November 1971.

[Mof68] A. Moffet. Minimum-redundancy linear arrays. 16(2) :172–175, June 1968.

[MS69] R. J. McAulay and L. P. Seidman. A useful form of the Barankin lower bound and its application to PPM threshold analysis. *IEEE Transactions on Information Theory*, 15(2) :273–279, March 1969.

[MS91] A. Mirkin and L. H. Sibul. Cramér-Rao bounds on angle estimation with a two-dimensional array. *IEEE Transactions on Signal Processing*, 39 :515–517, February 1991.

[MSPM04] K.N. Mokios, N.D. Sidiropoulos, M. Pesavento, and C.F. Mecklenbrauker. On 3-D harmonic retrieval for wireless channel sounding. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 2, pages 89–92, Philadelphia, U.S.A., 2004.

[NH88] T. J. Nohara and S. Haykin. Application of the Weiss-Weinstein bound to a two dimensional antenna array. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(9) :1533–1534, September 1988.

[Nie94] R. O. Nielsen. Azimuth and elevation angle estimation with a three dimensional array. 19 :84–86, 1994.

[NS09] D. Nion and N.D. Sidiropoulos. A PARAFAC-based technique for detection and localization of multiple targets in a MIMO radar system. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Taipei, Taiwan, 2009.

[NS10] D. Nion and D. Sidiropoulos. Tensor algebra and multi-dimensional harmonic retrieval in signal processing for MIMO radar. *IEEE Transactions on Signal Processing*, 58 :5693–5705, November 2010.

[NV94] H. Nguyen and H. L. Van Trees. Comparison of performance bounds for DOA estimation. In *Proc. of IEEE Workshop on Statistical Signal and Array Processing (SSAP)*, volume 1, pages 313–316, June 1994.

[OM05] U. Oktel and R. L. Moses. A Bayesian approach to array geometry design. *IEEE Transactions on Signal Processing*, 53(5) :1919– 1923, May 2005.

[OVSN93] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai. Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing. In S. S. Haykin, J. Litva, and T. J. Shepherd, editors, *Radar Array Processing*, chapter 4, pages 99–151. Springer-Verlag, Berlin, 1993.

[PG01] M. Pesavento and A.B. Gershman. Maximum-likelihood direction-of-arrival estimation in the presence of unknown nonuniform noise. *IEEE Transactions on Signal Processing*, 49 :1310–1324, July 2001.

[PMB04] M. Pesavento, C.F. Mecklenbrauker, and J.F. Bohme. Multidimensional rank reduction estimator for parametric MIMO channel models. *EURASIP Journal on Applied Signal Processing*, 9 :1354–1363, 2004.

[PP06] K.B. Petersen and M.S. Pedersen. The matrix cookbook. *Citeseer*, 2006.

[PRK86] A. Paulraj, R. Roy, and T. Kailath. A subspace rotation approach to signal parameter estimation. *Proceedings of the IEEE*, 74(7) :1044–1046, July 1986.

[RAFL07] A. Renaux, L. N. Atallah, Ph. Forster, and P. Larzabal. A useful form of the Abel bound and its application to estimator threshold prediction. *IEEE Transactions on Signal Processing*, 55(5) :2365–2369, May 2007.

[Rao45] C. R. Rao. Information and accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37 :81–91, 1945.

[RB74] D. C. Rife and R. R. Boorstyn. Single tone parameter estimation from discrete time observations. *IEEE Transactions on Information Theory*, 20(5) :591–598, September 1974.

[Ren07] A. Renaux. Weiss-Weinstein bound for data aided carrier estimation. *IEEE Signal Processing Letters*, 14(4) :283–286, April 2007.

[RFB04] A. Renaux, P. Forster, and E. Boyer. Non asymptotic efficiency of a maximum likelihood estimator at finite number of samples. In *Proc. of European Signal Processing Conference (EUSIPCO)*, pages 2247–2250, Vienna, AT, September 2004.

[RFBL] A. Renaux, Ph. Forster, E. Boyer, and P. Larzabal. Unconditional maximum likelihood performance at finite number of samples and high signal to noise ratio. *to appear in IEEE Trans. on Signal Processing.*

[RFBL04] A. Renaux, P. Forster, E. Boyer, and P. Larzabal. Non efficiency and non Gaussianity of a maximum likelihood estimator at high signal to noise ratio and finite number of samples. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 2, pages 121–124, Montreal, QC, CA, May 2004.

[RFBL07] A. Renaux, P. Forster, E. Boyer, and P. Larzabal. Unconditional maximum likelihood performance at finite number of samples and high signal-to-noise ratio. *IEEE Transactions on Signal Processing*, 55(5, Part 2) :2358–2364, May 2007.

[RFCL06] A. Renaux, P. Forster, E. Chaumette, and P. Larzabal. On the high-SNR conditional maximum-likelihood estimator full statistical characterization. *IEEE Transactions on Signal Processing*, 54(12) :4840–4843, December 2006.

[RFL05a] A. Renaux, P. Forster, and P. Larzabal. A new derivation of the Bayesian bounds for parameter estimation. In *Proc. of IEEE Workshop on Statistical Signal Processing (SSP)*, pages 567–572, Bordeaux, FR, July 2005.

[RFL05b] A. Renaux, P. Forster, and P. Larzabal. Une nouvelle approche des bornes bayésiennes. In *Proc. du colloque GRETSI sur le traitement du signal et des images*, volume 1, pages 141–144, Louvain-la-Neuve, BE, September 2005.

[RFL+08] A. Renaux, P. Forster, P. Larzabal, C. D. Richmond, and A. Nehorai. A fresh look at the Bayesian bounds of the Weiss-Weinstein family. *IEEE Transactions on Signal Processing*, 56(11) :5334–5352, November 2008.

[RFLR06] A. Renaux, Ph. Forster, P. Larzabal, and C.D. Richmond. The Bayesian Abel bound on the mean square error. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Toulouse, FR, May 2006.

[RHG07] F. Roemer, M. Haardt, and G. Del Galdo. Higher order SVD based subspace estimation to improve multi-dimensional parameter estimation algorithms. In *Proc. of Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, 2007.

[RM95] I. Reuven and H. Messer. The use of the Barankin bound for determining the threshold SNR in estimating the bearing of a source in the presence of another. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 3, pages 1645–1648, Detroit, MI, USA, May 1995.

[RM97] I. Reuven and H. Messer. A Barankin-type lower bound on the estimation error of a hybrid parameter vector. *IEEE Transactions on Information Theory*, 43(3) :1084–1093, May 1997.

[RS87a] Y. Rockah and P. Schultheiss. Array shape calibration using sources in unknown locations–Part I : Far-field sources. 35 :286–299, March 1987.

[RS87b] Y. Rockah and P. Schultheiss. Array shape calibration using sources in unknown locations–Part II : Near-field sources and estimator implementation. 35 :724–735, March 1987.

[Sch81] R. O. Schmidt. *A signal subspace approach to multiple emitter location and spectral estimation.* PhD thesis, Stanford University, Stanford, CA, USA, November 1981.

[Sch91] L. L. Scharf. *Statistical Signal Processing : Detection, Estimation, and Time Series Analysis*. Addison-Wesley, New-York, NY, USA, 1991.

[SD95] K. Sharman and T. Durrani. Resolving power of signal subspace methods for fnite data lengths. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 1501–1504, Florida, USA, 1995.

[SLG01] P. Stoica, E.G. Larsson, and A.B. Gershman. The stochastic CRB for array processing : a textbook derivation. 8 :148–150, May 2001.

[SM04] M. Shahram and P. Milanfar. Imaging below the diffraction limit : A statistical analysis. 13(5) :677–689, May 2004.

[SM05a] M. Shahram and P. Milanfar. On the resolvability of sinusoids with nearby frequencies in the presence of noise. *IEEE Transactions on Signal Processing*, 53(7) :2579–2585, July 2005.

[SM05b] P. Stoica and R.L. Moses. *Spectral Analysis of Signals*. Prentice Hall, NJ, 2005.

[Smi98] S. T. Smith. Accuracy and resolution bounds for adaptive sensor array processing. In *Proceedings in the ninth IEEE SP Workshop on Statistical Signal and Array Processing*, pages 37–40, 1998.

[Smi05] S. T. Smith. Statistical resolution limits and the complexified Cramér-Rao bound. *IEEE Transactions on Signal Processing*, 53 :1597–1609, May 2005.

[SN89] P. Stoica and A. Nehorai. MUSIC, maximum likelihood and the Cramér-Rao bound. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(5) :720–741, May 1989.

[SN90a] P. Stoica and A. Nehorai. MUSIC, maximum likelihood and the Cramér-Rao bound : further results and comparisons. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(12) :2140–2150, December 1990.

[SN90b] P. Stoica and A. Nehorai. Performances study of conditional and unconditional direction of arrival estimation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(10) :1783–1795, October 1990.

[SSL68] D. Sengupta, T. Smith, and R. Larson. Radiation characteristics of a spherical array of circularly polarized elements. 16(1) :2– 7, January 1968.

[SSS95] P. Stoica, V. Simonyte, and T. Soderstrom. On the resolution performance of spectral analysis. *Elsevier Signal Processing*, 44 :153–161, January 1995.

[Ste76] B.D. Steinberg. Principles of aperture and array system design : Including random and adaptive arrays. *New York, Wiley-Interscience*, vol. 1 :p. 374, 1976.

[STWT06] C. Schneider, U. Trautwein, W. Wirnitzer, and R.S. Thoma. Performance verification of MIMO concepts using multi-dimensional channel sounding. In *Proc. EUSIPCO*, Florence, Italy, September 2006.

[TGT96] K. C. Tan, S. S. Goh, and E. C. Tan. A study of the rank-ambiguilty issues in direction-of-arrival estimation. *IEEE Transactions on Signal Processing*, 44(4) :880–887, April 1996.

[THL+01] R.S. Thoma, D. Hampicke, M. Landmann, G. Sommerkorn, and A. Richter. MIMO measurement for double-directional channel modelling. In *COLLOQUIUM DIGEST-IEE*, pages 1–7, 2001.

[TK99] J. Tabrikian and J. L. Krolik. Barankin bounds for source localization in an uncertain ocean environment. *IEEE Transactions on Signal Processing*, 47(11) :2917–2927, November 1999.

[TT10] K. Todros and J. Tabrikian. General classes of performance lower bounds for parameter estimation Part I : Non-bayesian bounds for unbiased estimators. 56 :5045–5063, October 2010.

[Van68] H. L. VanTrees. *Detection, Estimation and Modulation Theory*, volume 1. Wiley, New York, 1968.

[Van95] F. Vanpoucke. *Algorithms and Architectures for Adaptive Array Signal Processing.* Ph. D. dissertation, Universiteit Leuven, Leuven, Belgium, 1995.

[Van01] H. L. VanTrees. *Detection, Estimation and Modulation Theory : Radar-Sonar Signal Processing and Gaussian Signals in Noise*, volume 3. Wiley, New York, 2001.

[Van02] H. L. VanTrees. *Detection, Estimation and Modulation theory : Optimum Array Processing*, volume 4. Wiley, New York, 2002.

[VEB+11] D. T. Vu, M. N. El Korso, R. Boyer, A. Renaux, and S. Marcos. Angular resolution limit for vector sensor arrays : Detection and information theory approaches. In *Proc. of IEEE Workshop on Satistical Signal Processing*, Nice, France, June 2011.

[VH86] E. Vertatschitsch and S. Haykin. Nonredundant arrays. *IEEE Transactions on Signal Processing*, 74(1) :217–217, June 1986.

[VO91] M. Viberg and B. Ottersten. Sensor array processing based on subspace fitting. *IEEE Transactions on Signal Processing*, 39(5) :1110–1121, May 1991.

[VOK91] M. Viberg, B. Ottersten, and T. Kailath. Detection and estimation in sensor arrays processing using Weighted Subspace Fitting. *IEEE Transactions on Signal Processing*, 39(11) :2436–2449, November 1991.

[VRBM10] D. T. Vu, A. Renaux, R. Boyer, and S. Marcos. Closed-form expression of the Weiss-Weinstein bound for 3D source localization : the conditional case. In *Proc. of IEEE Workshop on Sensor Array and Multi-channel Processing (SAM)*, Kibutz Ma'ale Hahamisha, Israel, October 2010.

[Wei88] E. Weinstein. Relations between Belini-Tartara, Chazan-Zakaï-Ziv, and Wax-Ziv lower bounds. *IEEE Transactions on Information Theory*, 34(2) :342–343, March 1988.

[WF93] A. J. Weiss and B. Friedlander. Range and bearing estimation using polynomial rooting. 18 :130–137, July 1993.

[WW85] A. J. Weiss and E. Weinstein. A lower bound on the mean square error in random parameter estimation. *IEEE Transactions on Information Theory*, 31(5) :680–682, September 1985.

[WW88] E. Weinstein and A. J. Weiss. A general class of lower bounds in parameter estimation. *IEEE Transactions on Information Theory*, 34(2) :338–342, March 1988.

[WZ97] K.T. Wong and M.D. Zoltowski. Uni-vector-sensor ESPRIT for multisource azimuth, elevation, andpolarization estimation. 45(10) :1467–1474, October 1997.

[XBR04] W. Xu, A. B. Baggeroer, and C. D. Richmond. Bayesian bounds for matched-field parameter estimation. *IEEE Transactions on Signal Processing*, 52(12) :3293–3305, December 2004.

[Xu01] W. Xu. *Performances bounds on matched-field methods for source localization and estimation of ocean environmental parameters.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, June 2001.

[YF98] N. Yuen and B. Friedlander. Performance analysis of higher order ESPRIT for localization of near-field sources. *IEEE Transactions on Signal Processing*, 46 :709–719, 1998.

[YS05] B. Yang and J. Scheuing. Cramér-Rao bound and optimum sensor array for source localization from the differences of arrival. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 4, pages 961–964, Philadenphia, USA, March 2005.

[ZC07] W. Zhi and M.Y.W. Chia. Near-field source localization via symmetric subarrays. 14(6) :409–412, 2007.

[ZZ69] J. Ziv and M. Zakai. Some lower bounds on signal parameter estimation. *IEEE Transactions on Information Theory*, 15(3) :386–391, May 1969.

Annexe A

A Cramér-Rao bound characterization of the EM-algorithm mean speed of convergence

IEEE Transactions on Signal Processing, Volume : 56, Issue : 6, Jun. 2008, pp. 2218-2228

2218 IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 56, NO. 6, JUNE 2008

A Cramér-Rao Bound Characterization of the EM-Algorithm Mean Speed of Convergence

Cédric Herzet, Valéry Ramon, Alexandre Renaux, and Luc Vandendorpe, *Fellow, IEEE*

***Abstract*—This paper deals with the *mean* speed of convergence of the expectation-maximization (EM) algorithm. We show that the asymptotic behavior (in terms of the number of observations) of the EM algorithm can be characterized as a function of the Cramér-Rao bounds (CRBs) associated to the so-called *incomplete* and *complete* data sets defined within the EM-algorithm framework. We particularize our result to the case of a complete data set defined as the concatenation of the observation vector and a vector of nuisance parameters, independent of the parameter of interest. In this particular case, we show that the CRB associated to the complete data set is nothing but the well-known modified CRB. Finally, we show by simulation that the proposed expression enables to properly characterize the EM-algorithm mean speed of convergence from the CRB behavior when the size of the observation set is large enough.**

***Index Terms*—Convergence of numerical methods, iterative methods, maximum-likelihood estimation.**

I. Introduction

SINCE its first statement by Dempster, Laird, and Rubin [1], the expectation-maximization (EM) algorithm has become a popular numerical method to compute maximum-likelihood (ML) estimates, see, e.g., [2] and references therein. Among the reasons of its success, its low complexity of implementation and its robustness are usually pointed out [2]. The main drawback of the EM algorithm is however its speed of convergence which, in some situations, may be extremely slow. In their seminal paper, Dempster, Laird, and Rubin [1] showed that the EM algorithm exhibits a linear speed of convergence, with a rate of convergence obtained from the information matrices associated to the *missing* and *complete* data sets. More recently, some authors [3], [4] have given further insights into the EM-algorithm convergence. In particular, in [3], [4], the authors emphasize that the EM algorithm may locally achieve quasi-Newton behavior in some specific situations.

In this paper, we address[1] the problem of the characterization of the EM-algorithm *mean* speed of convergence. We derive an expression relating the *asymptotic* (in terms of number of observations) mean speed of convergence of the EM algorithm to the Cramér-Rao bounds (CRBs) associated to the *incomplete* and *complete* data sets. In particular, we emphasize that, as long as the number of observation is large, the proposed expression enables to have a good intuition of the EM-algorithm speed of convergence simply by looking at the CRB behavior. The paper is organized as follows. In Section II, we recall the basics of the EM algorithm and we give the general expressions of the CRB and modified CRB (MCRB) [6]. Section III is the core of the paper: we derive an expression relating the EM-algorithm mean speed of convergence to the CRBs associated to the *incomplete* and *complete* data estimation problems. In a first part, we show that the proposed CRB-based expression is asymptotically valid under some conditions. Then, we particularize the proposed expression to the case where the complete data set is the concatenation of the received observation and a vector of *nuisance* parameters independent of the parameter of interest. In this particular case, we show that the CRB associated to the complete data set is equal to the MCRB. Finally, we briefly discuss the complexity associated to the evaluation of the proposed expression. In Section IV, we illustrate the relevance of the proposed approach in two different examples. We consider the estimation of the mean in a Gaussian mixture problem and the estimation of the carrier phase offset in a digital communication system. In particular, we emphasize that the EM algorithm behavior may be well predicted from the knowledge of the CRB and MCRB when the number of observations is large.

II. ML Estimation, EM Algorithm, and CRBs

In this section, we briefly review some notions which will be useful in the remainder of this paper. In Section II-A, we present the EM algorithm and discuss some of its features. In Section II-B, we recall the main equations of the standard and the modified CRBs.

A. ML Estimation and the EM Algorithm

Let us consider an observation vector $\mathbf{r}$ depending on an unknown deterministic scalar parameter b. The ML estimate of b given $\mathbf{r}$ is the solution of the following maximization problem:

$$\hat{b}_{\mathrm{ML}} = \arg\max_{\tilde{b}} \; p(\mathbf{r}\,|\,\tilde{b}) \tag{1}$$

where $\tilde{b}$ is a trial value of b. The ML estimator enjoys very good asymptotic statistical properties but its evaluation is unfortunately quite complex in a number of practical problems [7].

In order to circumvent this issue, iterative ML search methods have been proposed in the literature. In particular, the EM algorithm, proposed by Dempster, Laird, and Rubin in [1], is a

Manuscript received April 3, 2007; revised December 6, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Andreas Jakobsson.

C. Herzet is with the INRIA/IRISA, Campus Universitaire de Beaulieu, 35042 Rennes, France (e-mail: cherzet@inria.fr).

V. Ramon and L. Vandendorpe are with the Communications and Remote Sensing Laboratory, Université Catholique de Louvain, B1348 Louvain-la-Neuve, Belgium (e-mail: ramon@tele.ucl.ac.be; vandendorpe@tele.ucl.ac.be).

A. Renaux is with the Laboratoire des Signaux et Systemes, University Paris-Sud 11, Plateau de Moulon, 91192 Gif-sur-Yvette, France (e-mail: renaux @ieee.org).

Digital Object Identifier 10.1109/TSP.2008.917024

[1]This contribution is the extension of the following paper [5] by the same authors.

powerful iterative method which has been shown to provide a suitable solution to a number of problems encountered in the technical literature [2]. Formally, the EM algorithm is based on the following two steps:

$$\text{E}-\text{step:}\quad \mathcal{Q}(\tilde{b},\hat{b}^{(n)}) = \int_{\mathcal{Z}} p(\mathbf{z}|\mathbf{r},\hat{b}^{(n)})\,\log p(\mathbf{z},\mathbf{r}|\tilde{b})\,d\mathbf{z} \quad (2)$$

$$\text{M}-\text{step:}\quad \hat{b}^{(n+1)} = \arg\max_{\tilde{b}}\ \mathcal{Q}(\tilde{b},\hat{b}^{(n)}) \quad (3)$$

where $\hat{b}^{(n)}$ is the estimate computed by the EM algorithm at iteration n and $\mathbf{r} = f(\mathbf{z})$, where $f(\mathbf{z})$ is a *many-to-one* mapping. $\mathcal{Z}$ is the set of values that $\mathbf{z}$ can take on. Vectors $\mathbf{r}$ and $\mathbf{z}$ are often referred to as the *incomplete* and the *complete* data sets, respectively.

Since the EM algorithm is an iterative method, the question of its speed of convergence naturally arises. Dempster, Laird, and Rubin showed in their seminal paper [1] that the convergence of the EM algorithm is locally linear, i.e., we have in a neighborhood of $\hat{b}_{\text{ML}}$ that

$$e^{(n+1)}(\mathbf{r}) = C(\mathbf{r})\,e^{(n)}(\mathbf{r}) \quad (4)$$

where $e^{(n)}(\mathbf{r}) = |\hat{b}^{(n)}(\mathbf{r}) - \hat{b}_{\text{ML}}(\mathbf{r})|$ and $C(\mathbf{r})$ is the *rate of convergence* of the EM algorithm. The authors showed moreover that the rate of convergence of the EM algorithm may be related to the amount of *missing information*[2] in the considered problem, i.e.,

$$C(\mathbf{r}) = I_c^{-1}(\mathbf{r})\,I_m(\mathbf{r}) \quad (5)$$

where $I_c(\mathbf{r})$ and $I_m(\mathbf{r})$ are, respectively, the information matrices associated to the complete and the missing data, i.e.,

$$I_c(\mathbf{r}) \triangleq -\left(\int_{\mathcal{Z}} p(\mathbf{z}\,|\,\mathbf{r},\tilde{b})\,\frac{\partial^2}{\partial\tilde{b}^2}\,\log p(\mathbf{z},\mathbf{r}|\tilde{b})\,d\mathbf{z}\right)_{|\tilde{b}=\hat{b}_{\text{ML}}(\mathbf{r})} \quad (6)$$

$$I_m(\mathbf{r}) \triangleq -\left(\int_{\mathcal{Z}} p(\mathbf{z}\,|\,\mathbf{r},\tilde{b})\,\frac{\partial^2}{\partial\tilde{b}^2}\,\log p(\mathbf{z}\,|\,\mathbf{r},\tilde{b})\,d\mathbf{z}\right)_{|\tilde{b}=\hat{b}_{\text{ML}}(\mathbf{r})}. \quad (7)$$

B. The Standard and the Modified CRBs

The standard [7] and the modified [6] CRBs are lower bounds on the mean square estimation error of any unbiased estimator. In the case of the estimation of a scalar parameter b from a received observation vector $\mathbf{r}$, the (standard) CRB may be expressed as the inverse of the Fisher information matrix, i.e.,

$$\text{CRB}_{\mathbf{r}}(b) = -\left(E_{\mathbf{r}\,|\,b}\left[\frac{\partial^2}{\partial\tilde{b}^2}\log p(\mathbf{r}\,|\,\tilde{b})_{|\tilde{b}=b}\right]\right)^{-1} \quad (8)$$

$$= \left(E_{\mathbf{r}\,|\,b}\left[\left(\frac{\partial}{\partial\tilde{b}}\log p(\mathbf{r}\,|\,\tilde{b})_{|\,\tilde{b}=b}\right)^2\right]\right)^{-1} \quad (9)$$

where $\tilde{b}$ is a derivation variable.

[2]The missing information may actually be seen as the difference between the amount of information contained in the complete data set and the incomplete data set.

In some situations, the standard CRB may be quite tedious to evaluate. In particular, when the received observations $\mathbf{r}$ also depend on a nuisance parameter vector $\mathbf{a}$, the evaluation of $p(\mathbf{r}\,|\,b)$ may require a huge summation, see, e.g., [8]. In such situations, we may use the modified CRB (MCRB) [6], which is easier to compute but looser than the standard CRB, i.e., $\text{CRB}_{\mathbf{r}}(b) \geq \text{MCRB}(b)$. Formally, the MCRB is defined as

$$\text{MCRB}(b) = -\left(E_{\mathbf{r},\mathbf{a}|b}\left[\frac{\partial^2}{\partial\tilde{b}^2}\log p(\mathbf{r}|\mathbf{a},\tilde{b})_{|\tilde{b}=b}\right]\right)^{-1}. \quad (10)$$

To conclude this subsection, let us mention that the MCRB has been shown in [9] to be the high-SNR asymptote of the standard CRB, i.e., the CRB and the MCRB coincide when the SNR tends to infinity.

III. A CRB-Based Expression of the EM Mean Speed of Convergence

As discussed in the introduction, the local convergence of the EM algorithm given a *particular* observation vector $\mathbf{r}$ has already been well studied in the literature. In some situations, however, one may been interested in the *average* speed of convergence of the EM algorithm; the average being taken over the distribution of the observations $\mathbf{r}$. In this section, we will focus on this problem. In particular, we will consider the evolution of the following quantity throughout the iterations:

$$E_{\mathbf{r}\,|\,b}\left[e^{(n+1)}(\mathbf{r})\right] = \int_{\mathcal{R}} e^{(n+1)}(\mathbf{r})\,p(\mathbf{r}\,|\,b)d\mathbf{r} \quad (11)$$

where $\mathcal{R}$ is the set of values that $\mathbf{r}$ can take on.

In Section III-A, we will show that this quantity may be expressed as

$$E_{\mathbf{r}\,|\,b}[e^{(n+1)}(\mathbf{r})] = M_C\,E_{\mathbf{r}\,|\,b}[e^{(n)}(\mathbf{r})] \quad (12)$$

where

$$M_C = 1 - \frac{\text{CRB}_{\mathbf{z}}(b)}{\text{CRB}_{\mathbf{r}}(b)} \quad (13)$$

when the size of the observation vector $\mathbf{r}$ tends to infinity and under some regularity conditions. In other words, we will emphasize that, in the asymptotic regime, the evolution of the mean absolute error $E_{\mathbf{r}\,|\,b}[e^{(n)}(\mathbf{r})]$ is linear and that the factor of proportionality M_C is only a function of the ratio of two CRBs. In the sequel, we will refer to M_C as the *mean convergence rate* (MCR).

In Section III-B, we will emphasize that the MCR can be related to the well-known MCRB when the complete data set is made up of the concatenation of the observation vector $\mathbf{r}$ and a vector of nuisance parameter $\mathbf{a}$ independent of b. In such a case, we show that

$$M_C = 1 - \frac{\text{MCRB}(b)}{\text{CRB}_{\mathbf{r}}(b)}. \quad (14)$$

Finally, in Section III-C we discuss the practical evaluation of the MCR and the associated complexity.

A. Asymptotic Speed of Convergence of the EM Algorithm

In this section, we show that (12)–(13) enable to characterize the EM-algorithm behavior in the asymptotic regime. In order to do so, we will proceed in two steps. First, we will show that, $\forall\,\epsilon>0$, $E_{\mathbf{r}\,|\,b}[e^{(n+1)}(\mathbf{r})]$ may be lower and upper bounded as follows:

$$\begin{aligned}\frac{E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]}{E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]}\,K^-\left(\int_{\mathcal{R}} e^{(n)}(\mathbf{r})\,p(\mathbf{r}\,|\,b)\,d\mathbf{r}-\frac{\sigma}{\epsilon}\sqrt{\mathrm{MSE}_{\mathrm{EM}}^{(n)}}\right)\\ \le E_{\mathbf{r}\,|\,b}[e^{(n+1)}(\mathbf{r})]\\ \le \frac{E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]}{E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]}\,K^+\int_{\mathcal{R}} e^{(n)}(\mathbf{r})\,p(\mathbf{r}\,|\,b)\,d\mathbf{r}+\frac{\sigma}{\epsilon}\sqrt{\mathrm{MSE}_{\mathrm{EM}}^{(n)}}.\end{aligned} \tag{15}$$

where σ is the maximum of the standard deviations of $I_m(\mathbf{r})$ and $I_c(\mathbf{r})$, and

$$\mathrm{MSE}_{\mathrm{EM}}^{(n)} \triangleq \int_{\mathcal{R}} \left|\hat{b}_{\mathrm{ML}}(\mathbf{r})-\hat{b}^{(n)}(\mathbf{r})\right|^2 p(\mathbf{r}\,|\,b)\,d\mathbf{r}, \tag{16}$$

$$K^- \triangleq \frac{1-\frac{\epsilon}{E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]}}{1+\frac{\epsilon}{E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]}}$$

$$K^+ \triangleq \frac{1+\frac{\epsilon}{E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]}}{1-\frac{\epsilon}{E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]}}. \tag{17}$$

Then, in a second part, we will show that these bounds tend to $M_C\,E_{\mathbf{r}\,|\,b}[e^{(n)}]$ when the size of the observation vector tends to infinity.

Let us first show (15). We will assume that $\hat{b}^{(n)}$ is in a neighborhood of $\hat{b}_{\mathrm{ML}}$ with probability 1. Therefore, using (4) we have

$$E_{\mathbf{r}\,|\,b}\left[e^{(n+1)}\right]=\int_{\mathcal{R}} C(\mathbf{r})\,e^{(n)}(\mathbf{r})\,p(\mathbf{r}\,|\,b)\,d\mathbf{r}. \tag{18}$$

Let us define, $\forall\,\epsilon>0$, the following subspaces of the observation space $\mathcal{R}$

$$\begin{aligned}\mathcal{R}_\epsilon=\{\mathbf{r}\,|\,&|I_m(\mathbf{r})-E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]|\le\epsilon;|I_c(\mathbf{r})\\ &-E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]|\le\epsilon\}\end{aligned} \tag{19}$$

$$\bar{\mathcal{R}}_\epsilon=\mathcal{R}\backslash\mathcal{R}_\epsilon. \tag{20}$$

Since $\mathcal{R}_\epsilon\cup\bar{\mathcal{R}}_\epsilon=\mathcal{R}$ and $\mathcal{R}_\epsilon\cap\bar{\mathcal{R}}_\epsilon=\emptyset$, we have

$$\begin{aligned}\int_{\mathcal{R}} C(\mathbf{r})\,e^{(n)}(\mathbf{r})\,p(\mathbf{r}\,|\,b)\,d\mathbf{r}\\ =\int_{\mathcal{R}_\epsilon} C(\mathbf{r})\,e^{(n)}(\mathbf{r})\,p(\mathbf{r}\,|\,b)\,d\mathbf{r}\\ +\int_{\bar{\mathcal{R}}_\epsilon} C(\mathbf{r})\,e^{(n)}(\mathbf{r})\,p(\mathbf{r}\,|\,b)\,d\mathbf{r}.\end{aligned} \tag{21}$$

In order to find a lower and an upper bound on the left-hand side of (21), we will derive a lower and an upper bound on each term of the right-hand side. Let us first consider the first term. Using the definition of $\mathcal{R}_\epsilon$, we can lower-bound and upper-bound $C(\mathbf{r})$ as follows:

$$\begin{aligned}C(\mathbf{r})&=\frac{I_m(\mathbf{r})}{I_c(\mathbf{r})}\\ &\le\frac{E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]+\epsilon}{E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]-\epsilon}=\frac{E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]}{E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]}\,\frac{1+\frac{\epsilon}{E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]}}{1-\frac{\epsilon}{E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]}} \quad (22)\\ &\ge\frac{E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]-\epsilon}{E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]+\epsilon}=\frac{E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]}{E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]}\,\frac{1-\frac{\epsilon}{E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]}}{1+\frac{\epsilon}{E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]}}. \quad (23)\end{aligned}$$

Hence, using the definition of K^- and K^+ in (17) we have

$$\begin{aligned}\frac{E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]}{E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]}\,K^-\int_{\mathcal{R}_\epsilon} e^{(n)}(\mathbf{r})\,p(\mathbf{r}\,|\,b)\,d\mathbf{r}\\ \le\int_{\mathcal{R}_\epsilon} C(\mathbf{r})\,e^{(n)}(\mathbf{r})\,p(\mathbf{r}\,|\,b)\,d\mathbf{r}\\ \le\frac{E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]}{E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]}\,K^+\int_{\mathcal{R}_\epsilon} e^{(n)}(\mathbf{r})\,p(\mathbf{r}\,|\,b)\,d\mathbf{r}.\end{aligned} \tag{24}$$

Let us now derive a bound on the second term in (21). Since $0\le C(\mathbf{r})\le 1$, $e^{(n)}(\mathbf{r})\ge 0$, and $p(\mathbf{r}\,|\,b)\ge 0$, we have

$$\begin{aligned}0&\le\int_{\bar{\mathcal{R}}_\epsilon} C(\mathbf{r})\,e^{(n)}(\mathbf{r})\,p(\mathbf{r}\,|\,b)\,d\mathbf{r}\\ &\le\int_{\bar{\mathcal{R}}_\epsilon} e^{(n)}(\mathbf{r})\,p(\mathbf{r}\,|\,b)\,d\mathbf{r}.\end{aligned} \tag{25}$$

Moreover, by Cauchy-Schwarz we have that

$$\begin{aligned}\int_{\bar{\mathcal{R}}_\epsilon} e^{(n)}(\mathbf{r})\,p(\mathbf{r}\,|\,b)\,d\mathbf{r}\\ \le\sqrt{\int_{\bar{\mathcal{R}}_\epsilon}\left(e^{(n)}(\mathbf{r})\right)^2 p(\mathbf{r}\,|\,b)\,d\mathbf{r}\int_{\bar{\mathcal{R}}_\epsilon} p(\mathbf{r}\,|\,b)\,d\mathbf{r}}, \quad (26)\\ \le\sqrt{\int_{\mathcal{R}}\left(e^{(n)}(\mathbf{r})\right)^2 p(\mathbf{r}\,|\,b)\,d\mathbf{r}\int_{\bar{\mathcal{R}}_\epsilon} p(\mathbf{r}\,|\,b)\,d\mathbf{r}}. \quad (27)\end{aligned}$$

Now we have that

$$\begin{aligned}\int_{\bar{\mathcal{R}}_\epsilon} p(\mathbf{r}\,|\,b)\,d\mathbf{r}&=\Pr\{\mathbf{r}\in\bar{\mathcal{R}}_\epsilon\}, \quad (28)\\ &\le\max(\Pr\{|I_m(\mathbf{r})-E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]|>\epsilon\}\\ &\quad\Pr\{|I_c(\mathbf{r})-E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]|>\epsilon\}) \quad (29)\end{aligned}$$

since $\mathbf{r}\in\bar{\mathcal{R}}_\epsilon$ is defined by the intersection of two events. Moreover, by Chebychev we have that the probability that $I_m(\mathbf{r})$ and $I_c(\mathbf{r})$ are "ϵ-away" from their mean is upper bounded by

$$\Pr\{|I_m(\mathbf{r})-E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]|>\epsilon\}\le\frac{\sigma_m^2}{\epsilon^2} \tag{30}$$

$$\Pr\{|I_c(\mathbf{r})-E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]|>\epsilon\}\le\frac{\sigma_c^2}{\epsilon^2} \tag{31}$$

where σ_m^2 (respectively, σ_c^2) is the variance of $I_m(\mathbf{r})$ (respectively, $I_c(\mathbf{r})$). Defining

$$\sigma=\max(\sigma_m,\sigma_c) \tag{32}$$

we have, therefore

$$0 \leq \int_{\bar{\mathcal{R}}_\epsilon} C(\mathbf{r})\, e^{(n)}(\mathbf{r})\, p(\mathbf{r}\,|\,b)\, d\mathbf{r} \leq \frac{\sigma}{\epsilon}\sqrt{\mathrm{MSE}_{\mathrm{EM}}^{(n)}}. \quad (33)$$

Combining (24) and (33), we can, therefore, bound $E_{\mathbf{r}\,|\,b}[e^{(n+1)}(\mathbf{r})]$ as follows:

$$\begin{aligned}&\frac{E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]}{E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]}\, K^- \int_{\mathcal{R}_\epsilon} e^{(n)}(\mathbf{r})\, p(\mathbf{r}\,|\,b)\, d\mathbf{r}\\ &\quad\leq E_{\mathbf{r}\,|\,b}\left[e^{(n+1)}(\mathbf{r})\right]\\ &\quad\leq \frac{E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]}{E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]}\, K^+ \int_{\mathcal{R}_\epsilon} e^{(n)}(\mathbf{r})\, p(\mathbf{r}\,|\,b)\, d\mathbf{r} + \frac{\sigma}{\epsilon}\sqrt{\mathrm{MSE}_{\mathrm{EM}}^{(n)}}.\end{aligned} \quad (34)$$

Furthermore, we have

$$\int_{\mathcal{R}_\epsilon} e^{(n)}(\mathbf{r})\, p(\mathbf{r}\,|\,b)\, d\mathbf{r} \leq \int_{\mathcal{R}} e^{(n)}(\mathbf{r})\, p(\mathbf{r}\,|\,b)\, d\mathbf{r} \quad (35)$$

$$\begin{aligned}&\int_{\mathcal{R}_\epsilon} e^{(n)}(\mathbf{r})\, p(\mathbf{r}\,|\,b)\, d\mathbf{r}\\ &\quad= \int_{\mathcal{R}} e^{(n)}(\mathbf{r})\, p(\mathbf{r}\,|\,b)\, d\mathbf{r} - \int_{\bar{\mathcal{R}}_\epsilon} e^{(n)}(\mathbf{r})\, p(\mathbf{r}\,|\,b)\, d\mathbf{r} \quad (36)\\ &\quad\geq \int_{\mathcal{R}} e^{(n)}(\mathbf{r})\, p(\mathbf{r}\,|\,b)\, d\mathbf{r} - \frac{\sigma}{\epsilon}\sqrt{\mathrm{MSE}_{\mathrm{EM}}^{(n)}}. \quad (37)\end{aligned}$$

Inequality (35) follows from the nonnegativity of $e^{(n)}(\mathbf{r})$ and $p(\mathbf{r}\,|\,b)$. Equations (36) and (37), respectively, follows from the fact that $\mathcal{R}_\epsilon \cup \bar{\mathcal{R}}_\epsilon = \mathcal{R}$ and from (33). Using (35) and (37), we finally end up with (15).

Let us now consider the bounds in (15) when the number of observations, say N, tends to infinity and under the following conditions:

$$\lim_{N\to\infty} E_{\mathbf{r}\,|\,b}[(\hat{b}_{\mathrm{ML}} - b)^2] = 0 \quad (38)$$

$$\lim_{N\to\infty} \frac{\mathrm{CRB}_{\mathbf{r}}}{\mathrm{CRB}_{\mathbf{z}}} \neq 1 \quad (39)$$

$$\sigma < \infty. \quad (40)$$

First, taking (38) into account, we also have asymptotically that

$$E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})] = \mathrm{CRB}_{\mathbf{z}}^{-1} \quad (41)$$

$$E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})] = \mathrm{CRB}_{\mathbf{z}}^{-1} - \mathrm{CRB}_{\mathbf{r}}^{-1}. \quad (42)$$

Relations (42) and (41) can be shown as follows. Using the definition of the complete-data information matrix (6) and taking the expectation with respect to $p(\mathbf{r}\,|\,b)$, we have

$$\begin{aligned}E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})] = -\int_{\mathcal{R}} p(\mathbf{r}\,|\,b) \Bigg[\int_{\mathcal{Z}} p(\mathbf{z}|\mathbf{r},\tilde{b})\\ \times \frac{\partial^2}{\partial\tilde{b}^2}\log p(\mathbf{z},\mathbf{r}|\tilde{b})\, d\mathbf{z}\Bigg]_{|\tilde{b}=\hat{b}_{\mathrm{ML}}(\mathbf{r})} d\mathbf{r}.\end{aligned} \quad (43)$$

From (38), we have that $\hat{b}_{\mathrm{ML}} \simeq b$ asymptotically and therefore

$$\begin{aligned}&E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})]\\ &\quad= \left[-\int_{\mathcal{R}}\int_{\mathcal{Z}} p(\mathbf{z},\mathbf{r}\,|\,\tilde{b})\, \frac{\partial^2}{\partial\tilde{b}^2}\log p(\mathbf{z},\mathbf{r}\,|\,\tilde{b})\, d\mathbf{z}\, d\mathbf{r}\right]_{|\tilde{b}=b}\\ &\quad= \left[-\int_{\mathcal{Z}} p(\mathbf{z}\,|\,\tilde{b})\, \frac{\partial^2}{\partial\tilde{b}^2}\log p(\mathbf{z},\mathbf{r}\,|\,\tilde{b})\, d\mathbf{z}\right]_{|\tilde{b}=b}. \quad (44)\end{aligned}$$

Now, we have that

$$\begin{aligned}&\frac{\partial^2}{\partial\tilde{b}^2}\log p(\mathbf{z},\mathbf{r}\,|\,\tilde{b})\\ &\quad= \frac{\partial^2}{\partial\tilde{b}^2}\log p(\mathbf{r}\,|\,\mathbf{z},\tilde{b}) + \frac{\partial^2}{\partial\tilde{b}^2}\log p(\mathbf{z}\,|\,\tilde{b}) \quad (45)\\ &\quad= \frac{\partial^2}{\partial\tilde{b}^2}\log p(\mathbf{z}\,|\,\tilde{b}) \quad (46)\end{aligned}$$

since[3] $p(\mathbf{r}|\mathbf{z},\tilde{b}) = \mathbb{I}\{\mathbf{r} = f(\mathbf{z})\}$ and is, therefore, not a function of $\tilde{b}$. Plugging (46) in (44), we get

$$E_{\mathbf{r}\,|\,b}[I_c(\mathbf{r})] = \left[-\int_{\mathcal{Z}} p(\mathbf{z}\,|\,\tilde{b})\, \frac{\partial^2}{\partial\tilde{b}^2}\log p(\mathbf{z}|\tilde{b})\, d\mathbf{z}\right]_{|\tilde{b}=b}. \quad (47)$$

Comparing (47) with (8), we see that the right-hand side (RHS) of (47) is nothing but the inverse of the CRB associated to b and based on the complete data set $\mathbf{z}$. This shows (41). Let us now consider (42). First, notice that $I_m(\mathbf{r})$ and $I_c(\mathbf{r})$ may be related [2] as

$$-\left(\frac{\partial}{\partial\tilde{b}^2}\log p(\mathbf{r}\,|\,\tilde{b})\right)_{|\,\tilde{b}=\hat{b}_{\mathrm{ML}}(\mathbf{r})} = I_c(\mathbf{r}) - I_m(\mathbf{r}). \quad (48)$$

Based on (48), we may write

$$E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})] = E_{\mathbf{r}\,|\,b}\left[I_c(\mathbf{r}) + \left(\frac{\partial}{\partial\tilde{b}^2}\log p(\mathbf{r}\,|\,\tilde{b})\right)_{|\,\tilde{b}=\hat{b}_{\mathrm{ML}}(\mathbf{r})}\right]. \quad (49)$$

Using the fact that $\hat{b}_{\mathrm{ML}} \simeq b$ from (38), we finally have

$$\begin{aligned}&E_{\mathbf{r}\,|\,b}[I_m(\mathbf{r})]\\ &\quad= \mathrm{CRB}_{\mathbf{z}}^{-1}(b) + E_{\mathbf{r}\,|\,b}\left[\left(\frac{\partial}{\partial\tilde{b}^2}\log p(\mathbf{r}\,|\,\tilde{b})\right)_{|\,\tilde{b}=b}\right] \quad (50)\\ &\quad= \mathrm{CRB}_{\mathbf{z}}^{-1}(b) - \mathrm{CRB}_{\mathbf{r}}^{-1}(b) \quad (51)\end{aligned}$$

where (50) follows from (41), and (51) follows from the definition of the CRB (8). This shows (42).

Plugging (38), (39), and (40) into (15), we get

$$\begin{aligned}&M_C\, K_\infty^- \left(E_{\mathbf{r}\,|\,b}[e^{(n)}(\mathbf{r})] - \frac{\sigma}{\epsilon}\sqrt{\mathrm{MSE}_{\mathrm{EM}}^{(n)}}\right)\\ &\quad\leq E_{\mathbf{r}\,|\,b}\left[e^{(n+1)}(\mathbf{r})\right]\\ &\quad\leq M_C\, K_\infty^+ E_{\mathbf{r}\,|\,b}\left[e^{(n)}(\mathbf{r})\right] + \frac{\sigma}{\epsilon}\sqrt{\mathrm{MSE}_{\mathrm{EM}}^{(n)}}.\end{aligned} \quad (52)$$

[3] $\mathbb{I}\{S\}$ is the indicator function which is equal to 1 if the statement S is true and 0 otherwise.

where

$$K_{\infty}^{-} \triangleq \frac{1-\frac{\epsilon}{\mathrm{CRB}_{\mathbf{r}}^{-1}}}{1+\frac{\epsilon}{\mathrm{CRB}_{\mathbf{z}}^{-1}-\mathrm{CRB}_{\mathbf{r}}^{-1}}}$$
$$K_{\infty}^{+} \triangleq \frac{1+\frac{\epsilon}{\mathrm{CRB}_{\mathbf{r}}^{-1}}}{1-\frac{\epsilon}{\mathrm{CRB}_{\mathbf{z}}^{-1}-\mathrm{CRB}_{\mathbf{r}}^{-1}}}. \tag{53}$$

Note that (38) also implies that $\lim_{N\to\infty}\mathrm{CRB}_{\mathbf{r}}=0$ since the CRB is a lower bound on the mean square error. Therefore, as long as (39) holds we have for any $\epsilon<\infty$ that

$$K_{\infty}^{-}=1, \quad K_{\infty}^{+}=1. \tag{54}$$

Finally, notice that our initial assumption that $\hat{b}^{(n)}$ is (with probability 1) in a neighborhood of $\hat{b}_{\mathrm{ML}}$ ensures that $\mathrm{MSE}_{\mathrm{EM}}^{(n)}<\infty$. Therefore, since (52) is valid $\forall\,\epsilon<\infty$, we have

$$\begin{aligned} &M_C\,E_{\mathbf{r}\,|\,b}[e^{(n)}(\mathbf{r})]-\epsilon' \\ &\quad\le E_{\mathbf{r}\,|\,b}[e^{(n+1)}(\mathbf{r})]\le M_C\,E_{\mathbf{r}\,|\,b}[e^{(n)}(\mathbf{r})]+\epsilon' \qquad \forall\,\epsilon'>0. \end{aligned} \tag{55}$$

This proves (12).

Let us pause a moment to discuss this result. As far as our building assumptions are valid, we just showed that (12) and (13) establish a relationship between the rate of improvement of the mean absolute error $E_{\mathbf{r}\,|\,b}[e^{(n)}(\mathbf{r})]$ and the CRBs associated to the complete and incomplete data sets. In particular, we see from (13) that the (mean) rate at which the EM algorithm converges to the ML estimate decreases a function of the ratio $\mathrm{CRB}_{\mathbf{z}}/\mathrm{CRB}_{\mathbf{r}}$. This ratio is actually a measure of the improvement of the estimation quality which can be achieved by observing the complete-data set instead of the incomplete-data set. Note that (see the Appendix),

$$0\le\mathrm{CRB}_{\mathbf{z}}\le\mathrm{CRB}_{\mathbf{r}} \tag{56}$$

and, therefore, by (13)

$$0\le M_C\le 1. \tag{57}$$

This implies that the mean absolute estimation error will be nonincreasing at each iteration, which is in good accordance with the convergence properties of the EM algorithm [1], [10]. Note also from (13) that the larger the ratio $\mathrm{CRB}_{\mathbf{z}}/\mathrm{CRB}_{\mathbf{r}}$, the slower the mean convergence of the EM algorithm. Interestingly, this result is in good accordance with the existing convergence results [1], [2] according to which the speed of convergence of the EM algorithm depends on the amount of *missing* information between the incomplete and the complete data sets.

B. Mean Convergence Rate: Nuisance Parameter Case

In this section, we particularize the MCR expression (13) to the case where the complete data set is defined as $\mathbf{z}\triangleq[\mathbf{r}^T,\mathbf{a}^T]^T$, where $\mathbf{a}$ is a vector of *nuisance parameters* affecting $\mathbf{r}$ and independent of b. This particular case often occurs in digital communication systems where we have to estimate some channel parameter b independent of the transmitted symbol sequence $\mathbf{a}$. We show that, in this particular case, the MCR is written as

$$M_C=1-\frac{\mathrm{MCRB}(b)}{\mathrm{CRB}_{\mathbf{r}}(b)}. \tag{58}$$

To prove (58), we show that $\mathrm{CRB}_{\mathbf{z}}^{-1}(b)=\mathrm{MCRB}^{-1}(b)$ when $\mathbf{z}\triangleq[\mathbf{r}^T,\mathbf{a}^T]^T$. Using the definition of both the CRB and the complete data set $\mathbf{z}$, we, respectively, have

$$\begin{aligned} &\mathrm{CRB}_{\mathbf{z}}^{-1}(b) \\ &\quad=\left[-\int_{\mathcal{Z}}p(\mathbf{z}|\tilde{b})\,\frac{\partial^2}{\partial\tilde{b}^2}\log p(\mathbf{z}|\tilde{b})\,d\mathbf{z}\right]_{|\tilde{b}=b} \qquad (59)\\ &\quad=\left[-\int_{\mathcal{A}}\int_{\mathcal{R}}p(\mathbf{r},\mathbf{a}|\tilde{b})\,\frac{\partial^2}{\partial\tilde{b}^2}\log p(\mathbf{r},\mathbf{a}|\tilde{b})\,d\mathbf{r}\,d\mathbf{a}\right]_{|\tilde{b}=b}. \qquad (60)\end{aligned}$$

Using the Bayes rule and the independence between $\mathbf{a}$ and b, we finally have

$$\begin{aligned} &\mathrm{CRB}_{\mathbf{z}}^{-1}(b) \\ &\quad=\left[-\int_{\mathcal{A}}\int_{\mathcal{R}}p(\mathbf{r},\mathbf{a}|\tilde{b})\right. \\ &\qquad\left.\times\,\frac{\partial^2}{\partial\tilde{b}^2}\Big(\log p(\mathbf{r}\,|\,\mathbf{a},\tilde{b})+\log p(\mathbf{a})\Big)\,d\mathbf{r}\,d\mathbf{a}\right]_{|\tilde{b}=b} \qquad (61)\\ &\quad=\left[-\int_{\mathcal{A}}\int_{\mathcal{R}}p(\mathbf{r},\mathbf{a}|\tilde{b})\,\frac{\partial^2}{\partial\tilde{b}^2}\log p(\mathbf{r}\,|\,\mathbf{a},\tilde{b})\,d\mathbf{r}\,d\mathbf{a}\right]_{|\tilde{b}=b}. \qquad (62)\end{aligned}$$

Comparing (62) with (10), we see that the RHS of (62) corresponds to the definition of $\mathrm{MCRB}^{-1}(b)$. This shows (58).

C. Practical Evaluation of the MCR

In this section, we briefly discuss the complexity associated to the evaluation of the MCR. In particular, it is interesting to compare the complexity associated to the computation of the MCR to the complexity pertaining to the evaluation the standard convergence rate $C(\mathbf{r})$ defined in (5).

On the one hand, the evaluation of $C(\mathbf{r})$ implies: i) the computation of $\hat{b}_{\mathrm{ML}}$; ii) the evaluation of the information matrices $I_m(\mathbf{r})$ and $I_c(\mathbf{r})$. Unfortunately, these two operations imply most of the time a large computational burden in scenarios of practical interest. First, the computation of $\hat{b}_{\mathrm{ML}}$ implicitly requires a complexity equivalent to running the EM algorithm until convergence. Indeed, if there exists an algorithm able to compute $\hat{b}_{\mathrm{ML}}$ with a complexity lower than the EM algorithm, considering the EM algorithm has only little interest. Moreover, the evaluation of $I_c(\mathbf{r})$ and $I_m(\mathbf{r})$ is also usually a complex task. Indeed, the EM algorithm is often used in situations where Newton-type algorithms (which requires the computation of the Hessian of the incomplete-data likelihood function) is too complex [2]. Now, if the evaluation of $\frac{\partial}{\partial\tilde{b}^2}\log p(\mathbf{r}\,|\,\tilde{b})$ is complex, so is the evaluation of $I_c(\mathbf{r})$ and $I_m(\mathbf{r})$ since these quantities are related as in (48).

On the other hand, the computation of the MCR (13) usually exhibits a reasonable complexity. Indeed, the evaluation of (13) requires the evaluation of $\mathrm{CRB}_{\mathbf{z}}$ and $\mathrm{CRB}_{\mathbf{r}}$. Now, $\mathrm{CRB}_{\mathbf{z}}$ has usually an explicit expression very easy to evaluate (see [6] and Section IV for some examples) and its complexity is, therefore,

negligible. Moreover, $\mathrm{CRB}_{\mathbf{r}}$ can for example be efficiently evaluated as follows [8].

1) Generate K independent realizations of $\mathbf{r}_i$ according to $p(\mathbf{r}\,|\,b)$.
2) For each vector $\mathbf{r}_i$, $1 \leq i \leq K$, evaluate $\frac{\partial}{\partial \tilde{b}} \log p(\mathbf{r}_i|\tilde{b})$ as follows:

$$\begin{aligned}
&\frac{\partial}{\partial \tilde{b}} \log p(\mathbf{r}_i|\tilde{b}) \\
&\quad = \frac{\partial}{\partial \tilde{b}} \log \int_{\mathcal{Z}} p(\mathbf{z}, \mathbf{r}_i|\tilde{b}) d\mathbf{z} && (63)\\
&\quad = \frac{\int_{\mathcal{Z}} \frac{\partial}{\partial \tilde{b}} p(\mathbf{z}, \mathbf{r}_i|\tilde{b}) d\mathbf{z}}{p(\mathbf{r}_i|\tilde{b})} && (64)\\
&\quad = \int_{\mathcal{Z}} \frac{p(\mathbf{z}, \mathbf{r}_i|\tilde{b})}{p(\mathbf{r}_i|\tilde{b})} \frac{\partial}{\partial \tilde{b}} \log p(\mathbf{z}, \mathbf{r}_i|\tilde{b}) d\mathbf{z} && (65)\\
&\quad = \int_{\mathcal{Z}} p(\mathbf{z}|\mathbf{r}_i, \tilde{b}) \frac{\partial}{\partial \tilde{b}} \log p(\mathbf{z}, \mathbf{r}_i|\tilde{b}) d\mathbf{z} && (66)
\end{aligned}$$

where we have used the fact that

$$\frac{\partial}{\partial \tilde{b}} \log f(\tilde{b}) = \frac{\frac{\partial}{\partial \tilde{b}} f(\tilde{b})}{f(\tilde{b})}$$

in (64) and (65).

3) Compute $\mathrm{CRB}_{\mathbf{r}}(b)$ as follows:

$$\mathrm{CRB}_{\mathbf{r}}(b) \simeq \frac{1}{K} \sum_{i=1}^{K} \left(\frac{\partial}{\partial \tilde{b}} \log p(\mathbf{r}_i|\tilde{b})_{|\tilde{b}=b} \right)^2. \quad (67)$$

Comparing (66) with (2)–(3), we can notice that the evaluation of (66) has more or less the same complexity as *one* EM-algorithm iteration (indeed, maximizing a function is in general roughly as complex as computing its gradient). The number of realizations K depends on the accuracy we want to achieve on the speed of convergence of the EM algorithm. In our simulations (see Section IV), we have noticed that if $\mathrm{CRB}_{\mathbf{r}}$ is not very close to $\mathrm{CRB}_{\mathbf{z}}$, we can already achieve a good accuracy for very small values of K (around 10–20). On the contrary, when $\mathrm{CRB}_{\mathbf{r}}$ is close to $\mathrm{CRB}_{\mathbf{z}}$, one needs to increase K to achieve the required accuracy.

Let us conclude this section by mentioning that in a number of situations, the proposed expression (13) enables to have a good insight into the EM-algorithm without making any computation. Indeed, first it is worth noticing that the behavior of the CRBs in many estimation problems has already been studied and is available in the literature. We can therefore take benefit from these results to predict the EM-algorithm convergence via (13). Moreover, as it will become clear from our examples in Section IV, the behavior of the CRB is often predictable by some intuitive reasoning. In such cases, it is therefore easy from (13) to have a *qualitative* idea of the evolution of the EM-algorithm speed of convergence when some parameters of the problem at hand are modified.

IV. Examples

In this section, we illustrate by simulation that the proposed expression enables to properly predict the convergence of the EM algorithm when the size of the observation set is large enough. In particular, we will consider the estimation of the mean in a Gaussian mixture scenario and the problem of estimating the carrier phase in a digital communication receiver.

A. Gaussian Mixture

As a first example, we consider the ML estimation of a parameter b in the following Gaussian mixture problem:

$$\begin{aligned}
p(\mathbf{r}\,|\,b) = \prod_i & \, p(r_i|a_i = 1, b)\, p(a_i = 1) \\
& + p(r_i|a_i = 2, b)\, p(a_i = 2) \quad (68)
\end{aligned}$$

where

$$p(r_i|a_i = 1, b) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{ -\frac{(r_i + b)^2}{2\sigma_1^2} \right\} \quad (69)$$

$$p(r_i|a_i = 2, b) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left\{ -\frac{(r_i - b)^2}{2\sigma_2^2} \right\}. \quad (70)$$

In words, the observation r_i is generated by first selecting either Gaussian distribution (69) or (70) with probability $p(a_i = 1)$ and $p(a_i = 2)$, respectively, and then drawing the observation according to $p(r_i|a_i, b)$. Equation (68)–(70), therefore, defines a particular Gaussian mixture problem where the two Gaussian distributions have, up to a sign, the same mean b but can have different variances. In the sequel we will assume that the distribution $p(a_i)$ is the same $\forall i$ and equal to $p(a)$.

Let us consider the ML estimation of b by means of the EM algorithm. Defining the complete data set $\mathbf{z} \triangleq [\mathbf{r}^T, \mathbf{a}^T]^T$, where $\mathbf{a}$ is the vector made up of all the a_i, and using the standard EM (2)–(3) we easily get the following update equation:

$$\hat{b}^{(n+1)} = \frac{\sum_i r_i\, \alpha_i^{(n)}}{\sum_i \beta_i^{(n)}} \quad (71)$$

where

$$\alpha_i^{(n)} = \frac{p\left(r_i|a_i = 2, \hat{b}^{(n)}\right)}{\sigma_2^2} - \frac{p\left(r_i|a_i = 1, \hat{b}^{(n)}\right)}{\sigma_1^2} \quad (72)$$

$$\beta_i^{(n)} = \frac{p\left(r_i|a_i = 2, \hat{b}^{(n)}\right)}{\sigma_2^2} + \frac{p\left(r_i|a_i = 1, \hat{b}^{(n)}\right)}{\sigma_1^2}. \quad (73)$$

According to our previous derivations, the EM-algorithm behavior should be asymptotically well characterized by means of the CRB and MCRB associated to the estimation problem [see (58)]. The CRB and MCRB can be computed as follows. On the one hand, using (10) one readily obtains that

$$\mathrm{MCRB} = \frac{1}{N} \left(\frac{p(a = 1)}{\sigma_1^2} + \frac{p(a = 2)}{\sigma_2^2} \right)^{-1}. \quad (74)$$

In particular, if $\sigma_1^2 = \sigma_2^2 = \sigma^2$, we have that

$$\mathrm{MCRB} = \frac{\sigma^2}{N} \quad (75)$$

i.e., the MCRB is equal to the CRB associated to the estimation of the mean of *one* single Gaussian of variance σ^2 [7]. On the other hand, the CRB can be computed following the procedure described in Section III-C. The CRB and the MCRB are

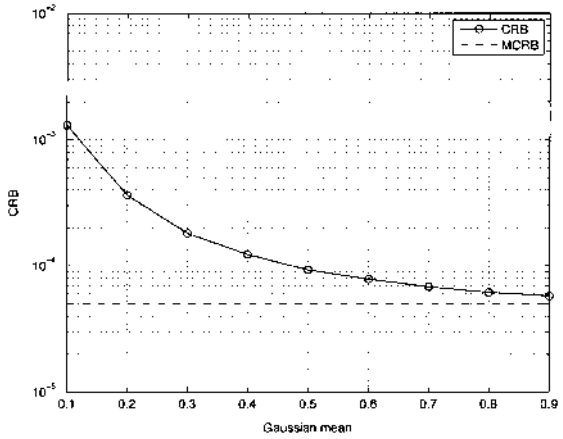


Fig. 1. Standard and modified CRBs versus the absolute mean of the Gaussians of the mixture.

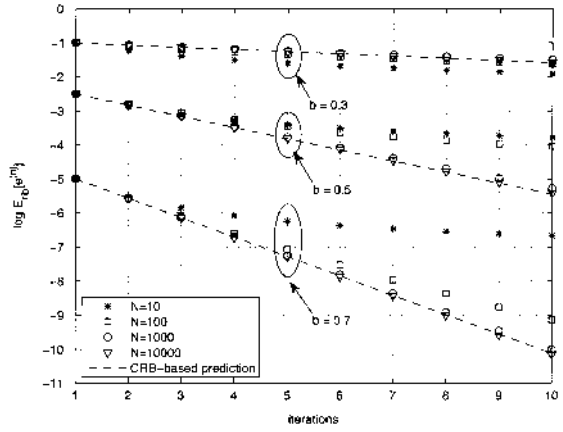


Fig. 2. Evolution of the mean absolute distance between the EM and the ML estimates for different number of observations.

represented in Fig. 1 for $\sigma_1^2 = \sigma_2^2 = 0.5$, $0.1 \leq b \leq 0.9$, and $p(a = 1) = p(a = 2) = 1/2$.

We have represented[4] in Fig. 2 the evolution of $\log E_{\mathbf{r}\,|\,b}[e^{(n)}(\mathbf{r})]$ versus the number of iterations. The dashed lines represent the performance predicted via (12) and (58) whereas the triangles, circles, squares, and stars represent the actual EM-algorithm performance[5] for different number of observations. We can notice that the simulation results are in good accordance with the proposed prediction when the number of observations increases. We can also notice that, as expected from the behavior of the CRB and MCRB in Fig. 1, the mean speed of convergence of the EM algorithm increases when the value of b increases.

[4]for the sake of clarity, the curves corresponding to different values of b have been vertically shifted.

[5]These curves have been computed by "Monte Carlo simulations," i.e., by averaging $e^{(n)}(\mathbf{r})$ for a large number of independent realizations of $\mathbf{r}$ drawn according to $p(\mathbf{r}\,|\,b)$.

Note that the behavior of the CRB can often be easily predicted by a little thought. For example, in the considered scenario, we can easily predict that the CRB tends to the MCRB when b is sufficiently large. Indeed, on the one hand, if b increases, the overlapping between the two Gaussians decreases. On the other hand, the MCRB [see (75)] actually corresponds to the CRB associated to the estimation of the mean of one single Gaussian. Now, it is intuitively clear that we can achieve the same degree of accuracy when estimating the mean of one Gaussian or the mean of two nonoverlapping Gaussians having opposite means. From this simple reasoning and using the proposed CRB-based expression, it is therefore easy to have a *qualitative* idea of the EM-algorithm behavior in the limit of large N. For example, in the considered case, we can predict that the mean speed of convergence will increase when b increases.

In the next section, we will give more examples in which we can intuitively predict how the EM-algorithm speed of convergence will evolve when some parameters of the problem at hand are modified.

B. Carrier Phase Estimation

In this section, we consider the practical example of the carrier-phase synchronization in a digital receiver. The observation model may be expressed as follows:

$$\mathbf{r} = \sqrt{\frac{E_s}{N_0}}\,\mathbf{a}\,e^{jb} + \mathbf{v} \tag{76}$$

where $\mathbf{a}$ is a vector of data symbols belonging to constellation alphabet $\mathcal{A}$, E_s is the mean energy per symbol, N_0 is the noise spectral density, b is the carrier-phase offset, and $\mathbf{v}$ is a vector of zero-mean white Gaussian noise samples with complex variance equal to 1.

The EM algorithm is applied to the problem of computing the ML estimate of the carrier phase offset. Defining the complete data set as $\mathbf{z} \triangleq [\mathbf{r}^T, \mathbf{a}^T]^T$, we get the following EM update equation [11]:

$$\hat{b}^{(n+1)} = \arg\left\{\sum_{i=1}^{N} \eta_i^{(n)}\, r_i\right\} \tag{77}$$

where

$$\eta_i^{(n)} = \sum_{a \in \mathcal{A}} a\; p\left(a_k = a\,|\,\mathbf{r}, \hat{b}^{(n)}\right) \tag{78}$$

and $\arg\{x\}$ denotes the argument of the complex number x. Note that the definition of the complete data set is similar to the one made in Section III-B. Hence, in the remainder of this section, we will compare the EM-algorithm speed of convergence to the one predicted using (58).

The MCR expression in (58) is only a function of the CRB/MCRB-ratio. On the other hand, note that the behavior of the CRBs and MCRB associated to carrier-phase estimation has been extensively studied in the literature, see, e.g., [12], [8], [6], and [9]. In [12], the authors derive the CRB expression for uncoded transmissions, i.e., assuming that all transmitted symbols are equiprobable. In [8], a semianalytical method is proposed

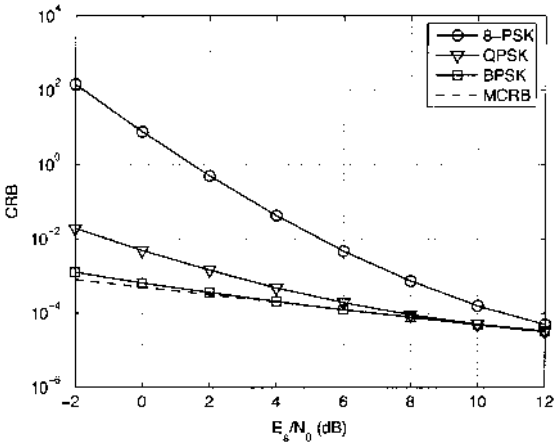


Fig. 3. CRBs versus E_s/N_0-ratio for uncoded PSK transmission with different constellation sizes.

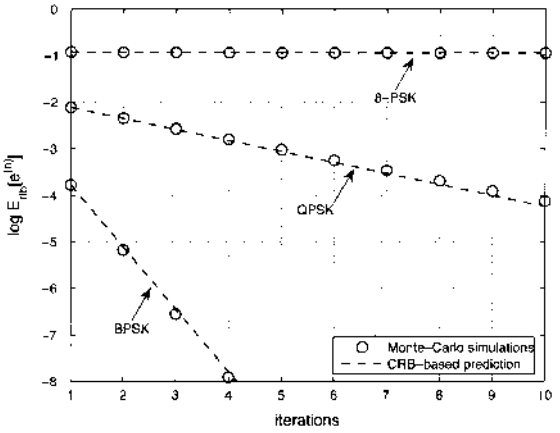


Fig. 4. Absolute mean distance between the ML estimate and the EM-algorithm estimate at a given iteration. Monte Carlo simulations are compared with the performance predicted by (58). The plot illustrates the sensitivity of the EM-algorithm speed of convergence to the choice of the constellation size.

for the evaluation of the CRB in coded scenarios. In [6], the authors derived the expression of the MCRB

$$\mathrm{MCRB} = \frac{1}{2N}\left(\frac{E_s}{N_0}\right)^{-1}. \tag{79}$$

Finally, in [9], the author shows that the CRB is equal to the MCRB at high signal-to-noise ratios (SNRs).

In the remainder of this section, we will illustrate that (58) may help predicting the EM-algorithm behavior by taking benefit from the knowledge of the CRB and MCRB behavior. We will consider the sensitivity of the EM-algorithm convergence to, respectively, the symbol-constellation size, the SNR and the symbol-sequence *a priori* information $p(\mathbf{a})$. In each scenario, the EM-algorithm performance computed via Monte Carlo simulations will be compared to the one predicted by means of (58).

a) Effect of the Constellation Size: We first investigate the EM-algorithm behavior when the size of the symbol constellation alphabet varies. We consider the following setup. The transmitted frames consist of 1000 uncoded PSK symbols. The size of the constellation alphabet is set to either 2 (BPSK), 4 (QPSK) or 8 (8-PSK). We use a Gray mapping.

The CRBs and the MCRB[6] associated to this setup are represented versus the E_s/N_0-ratio in Fig. 3. We see that for a given E_s/N_0-ratio, the larger the constellation size the worse the achievable estimation quality. This behavior is predictable from some intuitive reasoning. Indeed, it is easy to show that the MCRB defined in (79) is equal to the CRB associated to the carrier phase estimation problem when the symbols $\mathbf{a}$ are perfectly known at the receiver (i.e., there is no uncertainty about $\mathbf{a}$). On the other hand, for a given E_s/N_0-ratio, it is clear that the more elements we have in the constellation alphabet $\mathcal{A}$, the more uncertainty we have about the transmitted symbols. As a consequence, this implies that increasing the constellation size can only decrease the estimation quality. Based on the behavior of the CRBs and MCRB and using (58), we can therefore expect the EM algorithm to exhibit (asymptotically) a faster mean speed of converge when the size of the constellation decreases.

[6]Note that the MCRB does not depend on the constellation size.

Fig. 4 illustrates the validity of the proposed approach. We compare the EM-algorithm performance as predicted by (12) and (58) with actual performance computed via Monte Carlo simulations. More particularly, we have represented $\log E_{\mathbf{r}\,|\,b}[e^{(n)}(\mathbf{r})]$ versus the number of EM iterations. The E_s/N_0-ratio has been set to 4 dB. The EM algorithm has been initialized by means of a phase estimate computed by a Viterbi and Viterbi [13] synchronizer. The dashed curves correspond to the prediction computed from (12) and (58) whereas the circles correspond to the actual performance computed via Monte Carlo simulations. We see from Fig. 4 that the simulated points are very close to the performance predicted via (58). Therefore, as far as our simulation setup is concerned, the proposed CRB-based expression (58) of the MCR enables to accurately predict the EM algorithm convergence. In particular, we can easily predict from the CRB behavior in Fig. 3 that an increase of the constellation size will decrease the speed of convergence of the EM-based phase synchronizer.

b) Effect of the SNR: We now illustrate the sensitivity of the EM algorithm speed of convergence to the system operating SNR. We keep the same setup as in the previous point. The CRB and MCRB plotted in Fig. 3 are therefore still valid for computing the MCR via (58). We see that, as emphasized in [9], the CRBs tend to the MCRB at high SNR irrespective of the constellation size. From our previous derivations we can, therefore, conclude that, as long as N is large enough, the EM algorithm will exhibit a faster speed of convergence when the SNR increases.

In Fig. 5, we compare the performance predicted via (12) and (58) with the one computed by Monte Carlo simulations. The CRB-based predictions are plotted with dashed curves and the simulated points with circles. The constellation alphabet is a Gray-mapped 8-PSK and we have considered E_s/N_0 equal to 4, 8, and 12 dB, respectively. We see from this figure that the behavior predicted by (12) and (58) is in good accordance with the results computed by Monte Carlo simulations. We also see that, as expected, the EM algorithm converges all the faster as the SNR is large.

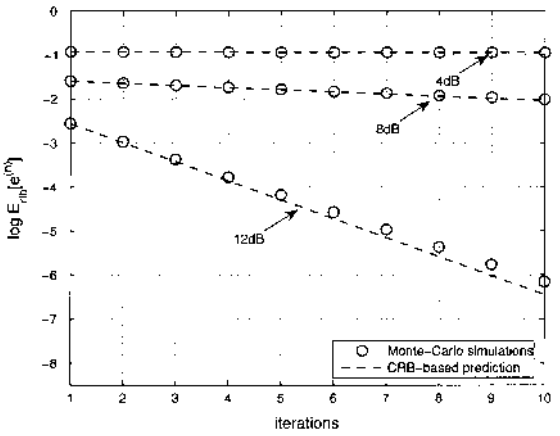


Fig. 5. Mean distance between the ML estimate and the EM-algorithm estimate at a given iteration. Monte Carlo simulations are compared with the performance predicted by (58). The plot illustrates the sensitivity of the EM-algorithm speed of convergence to the operating SNR.

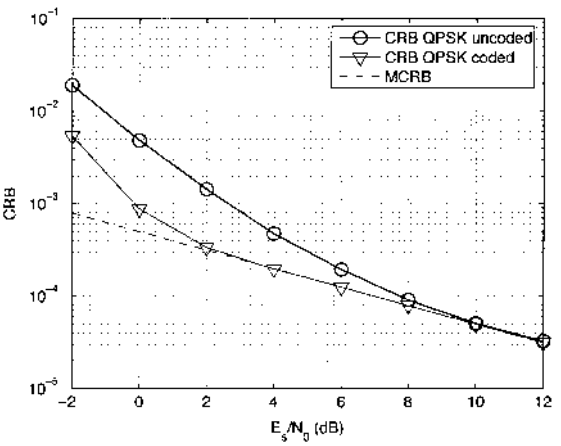


Fig. 6. CRBs versus E_s/N_0-ratio for QPSK uncoded and coded transmissions.

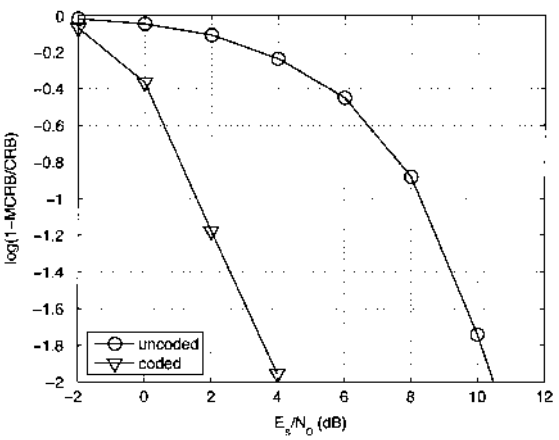


Fig. 7. EM-algorithm mean convergence rate versus E_s/N_0-ratio for QPSK uncoded and coded transmissions.

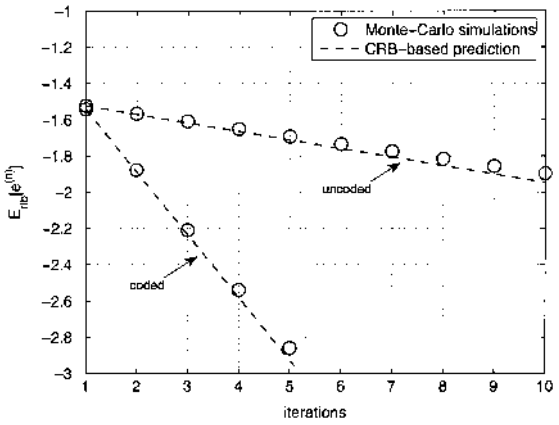


Fig. 8. Mean distance between the ML estimate and the EM-algorithm estimate at a given iteration. Monte Carlo simulations are compared with the performance predicted by (58). The plot illustrates the EM-algorithm speed of convergence for QPSK uncoded and coded transmissions.

c) Effect of Coding: As a last example of the utility of (12) and (58), we consider the effect of an error-correcting code on the speed of convergence of the EM algorithm. Fig. 6 represents the CRB associated to a (particular) coded and an uncoded transmission, respectively. The transmitted frames consist of 1000 QPSK symbols with Gray mapping. For the coded transmission, we used a rate-1/2 convolutional code with constraint length equal to 3. Since the code structure provides some *a priori* information about the transmitted sequence, "coded" CRB is always lower than the "uncoded" one [8]. According to our previous reasoning, this means that the EM algorithm should exhibit a faster convergence for coded transmissions than for uncoded ones. This is illustrated in Figs. 7 and 8. In Fig. 7 we have represented the MCR predicted by (58) in the coded and uncoded cases. We see that, in good accordance with our intuition, the MCR is always lower for the coded than for the uncoded transmission. In Fig. 8, we compare the MCR approximation given in (58) (dashed lines) to actual Monte Carlo simulation results (circles). We see that the EM speed of convergence is much faster in the coded case than in the uncoded one. We can note again the good accordance between the EM performance predicted via (12) and (58) and the simulation results.

V. Conclusion

In this contribution, we addressed the problem of the characterization of the EM algorithm mean speed of convergence. Based on some building assumptions, we showed that the EM-algorithm *asymptotic* mean speed convergence can be related to the CRBs associated to the incomplete and the complete data sets, respectively. Using the derived expression, we showed that we can gain a good intuition into the EM-algorithm behavior: the further the incomplete-data CRB is from the complete-data CRB, the slower the EM-algorithm speed

of convergence. We then particularized our result to the case where the complete data set is defined as the concatenation of the incomplete data set and a nuisance-parameter vector. In such a case, we emphasized that the complete-data CRB is nothing but the well-known modified CRB (MCRB). Finally, we illustrated our derivations by simulation results. In particular, we showed that the performance predicted by our approach is in good accordance with Monte Carlo simulation results when the number of observations is large.

APPENDIX

Here, we show that (56) is valid. The first inequality in (56) directly derives from the definition of the CRB [7], [14]. The second inequality may be proved as follows. Using twice the Bayes rule on $p(\mathbf{z},\mathbf{r}|b)$, we have

$$p(\mathbf{z}\,|\,\mathbf{r},b)\,p(\mathbf{r}\,|\,b) = p(\mathbf{r}|\mathbf{z},b)\,p(\mathbf{z}\,|\,b). \tag{80}$$

Since $\mathbf{z}$ univocally defines $\mathbf{r}$ through the many-to-one mapping $\mathbf{r} = f(\mathbf{z})$ (see Section II) we have

$$p(\mathbf{z}\,|\,\mathbf{r},b)\,p(\mathbf{r}\,|\,b) = \mathbb{I}\{\mathbf{r} = f(\mathbf{z})\}\,p(\mathbf{z}\,|\,b) \tag{81}$$

where $\mathbb{I}\{S\}$ is the indicator function which is equal to 1 if S is true and 0 otherwise. Taking the logarithm and the second derivative of (81), we then get

$$\frac{\partial^2}{\partial b^2}\log p(\mathbf{r}\,|\,b) = -\frac{\partial^2}{\partial b^2}\log p(\mathbf{z}\,|\,\mathbf{r},b) + \frac{\partial^2}{\partial b^2}\log p(\mathbf{z}\,|\,b) \tag{82}$$

since $\mathbb{I}\{\mathbf{r} = f(\mathbf{z})\}$ does not depend on b. Taking the expectation with respect to $p(\mathbf{z}\,|\,\mathbf{r},b)$ of both sides, we finally have

$$\begin{aligned}\frac{\partial^2}{\partial b^2}&\log p(\mathbf{r}\,|\,b)\\ &= -\int_{\mathcal{Z}} p(\mathbf{z}\,|\,\mathbf{r},b)\,\frac{\partial^2}{\partial b^2}\log p(\mathbf{z}\,|\,\mathbf{r},b)\,d\mathbf{z}\\ &\quad + \int_{\mathcal{Z}} p(\mathbf{z}\,|\,\mathbf{r},b)\,\frac{\partial^2}{\partial b^2}\log p(\mathbf{z}\,|\,b)\,d\mathbf{z}\end{aligned} \tag{83}$$

since the LHS does not depend on $\mathbf{z}$. Considering the first term in the RHS of (83), it can readily be shown [7] that

$$\begin{aligned}-\int_{\mathcal{Z}} p(\mathbf{z}\,|\,\mathbf{r},b)&\frac{\partial^2}{\partial b^2}\log p(\mathbf{z}\,|\,\mathbf{r},b)d\mathbf{z}\\ &= \int_{\mathcal{Z}} p(\mathbf{z}\,|\,\mathbf{r},b)\,\left(\frac{\partial}{\partial b}\log p(\mathbf{z}\,|\,\mathbf{r},b)\right)^2 d\mathbf{z} \geq 0.\end{aligned} \tag{84}$$

Hence

$$\frac{\partial^2}{\partial b^2}\log p(\mathbf{r}\,|\,b) \geq \int_{\mathcal{Z}} p(\mathbf{z}\,|\,\mathbf{r},b)\,\frac{\partial^2}{\partial b^2}\log p(\mathbf{z}\,|\,b)\,d\mathbf{z}. \tag{85}$$

Taking the expectation with respect to $p(\mathbf{r}\,|\,b)$ of both sides of (85), we have

$$\begin{aligned}\int_{\mathcal{R}} p(\mathbf{r}\,|\,b)\,&\frac{\partial^2}{\partial b^2}\log p(\mathbf{r}\,|\,b)\,d\mathbf{r}\\ &\geq \int_{\mathcal{Z}} p(\mathbf{z}\,|\,b)\,\frac{\partial^2}{\partial b^2}\log p(\mathbf{z}\,|\,b)\,d\mathbf{z}\end{aligned} \tag{86}$$

where the RHS derives from the fact that

$$\int_{\mathcal{R}} p(\mathbf{z}\,|\,\mathbf{r},b)p(\mathbf{r}\,|\,b)d\mathbf{r} = p(\mathbf{z}\,|\,b).$$

Taking the negative of (86) and inverting, we finally get (56).

REFERENCES

[1] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum-likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc.*, vol. 39, no. 1, pp. 1–38, Jan. 1977.

[2] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. New York: Wiley Series in Probabil. Statist., 1997.

[3] L. Xu and M. I. Jordan, "On convergence properties of the EM algorithm for Gaussian mixtures," *Neural Computat.*, vol. 8, no. 1, pp. 129–151, 1996.

[4] R. Salakhutdinov, S. Roweis, and Z. Ghahramam, On the Convergence of Bound Optimization Algorithms [Online]. Available: citeseer.ist.psu.edu/584732.html URL

[5] C. Herzet and L. Vandenclorpe, "Prediction of the EM-algorithm speed of convergence with Cramer-Rao bounds," in *IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, HI, Apr. 2007.

[6] A. D'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramer-Rao bound and its application to synchronization problems," *IEEE Trans. Commun.*, vol. 42, no. 2/3/4, pp. 1391–1399, Feb./Mar./Apr. 1994.

[7] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing Communications and Control*. Englewood Cliffs, NJ: Prentice-Hall Signal Processing Series, 1995.

[8] N. Noels, H. Steendam, and M. Moeneclaey, "The Cramer-Rao bound for phase estimation from coded linearly modulated signals," *IEEE Commun. Lett.*, vol. 7, no. 5, pp. 207–209, May 2003.

[9] M. Moeneclaey, "On the true and the modified Cramer-Rao bounds for the estimation of a scalar parameter in the presence of nuisance parameters," *IEEE Trans. Commun.*, vol. 46, no. 11, pp. 1536–1544, Nov. 1998.

[10] C. F. J. Wu, "On the convergence properties of the EM algorithm," *Ann. Statist.*, vol. 11, no. 1, pp. 95–103, 1983.

[11] N. Noels, C. Herzet, A. Dejonghe, V. Lottici, H. Steendam, M. Moeneclaey, M. Luise, and L. Vandendorpe, "Turbo-synchronization: An EM algorithm approach," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Anchorage, AK, May 2003, pp. 2933–2937.

[12] W. G. Cowley, "Phase and frequency estimation for PSK packets: Bounds and algorithms," *IEEE Trans. Commun.*, vol. 44, pp. 26–28, Jan. 1996.

[13] A. J. Viterbi and A. M. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *IEEE Trans. Inf. Theory*, vol. 29, pp. 543–551, Jul. 1983.

[14] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

Cédric Herzet was born in Verviers, Belgium, in 1978. He received the electrical engineering and Ph.D. degrees in applied science from the Université Catholique de Louvain (UCL), Louvain-la-Neuve, Belgium, in 2001 and 2006, respectively.

From August 2001 to April 2006, he was a Research Assistant with the Communications and Remote Sensing Laboratory, UCL. From August 2001 to December 2001, he worked on the topic of VDSL systems in collaboration with Alcatel Bell, Antwerp, Belgium. From January 2002 to April 2006, his research concerned synchronization and parameter estimation in iterative receivers. From May to August 2006, he was a Postdoctoral Researcher with the Ecole Normale Supérieure de Cachan, France. From October 2006 to October 2007, he was a Fulbright Postdoctoral Researcher with the University of California, Berkeley. He is currently a permanent researcher with the TEMICS Lab, IRISA/INRIA, Rennes, France. His research interests include detection, parameter estimation, coding theory, graphical models, and message-passing algorithms.

Valéry Ramon was born in Ath, Belgium, in 1978. He received the electrical engineering degree from the Université Catholique de Louvain (UCL), Louvain-la-Neuve, Belgium, in 2001. His graduation thesis concerned multiuser turbo receivers for wideband code division multiple access (WCDMA) transmissions.

Since August 2001, he has been a researcher with the Digital Communications Group (DIGICOM), Communications and Remote Sensing Laboratory, UCL. From August 2001 to December 2001, he worked on the topic of VDSL systems in collaboration with Alcatel Bell, Antwerp, Belgium. From January 2002 to January 2007, his research concerned channel estimation in iterative receivers. He is currently a Senior Researcher with the Wireless Group of the Inter-university Micro-Electronics Center (IMEC), Leuven, Belgium.

Alexandre Renaux was raised in a small French mountain town called "les Vosges." He received the Agrégation, the M.Sc., and the Ph.D. degrees in electrical engineering from the Ecole Normale Supérieure de Cachan, France, in 2002, 2003, and 2006, respectively.

From 2006 to 2007, he was a Postdoctoral Research Associate with the Department of Electrical and Systems Engineering, Washington University, St. Louis, MO. He is currently an Assistant Professor with the Department of Physics and a member of the Laboratory of Signals and Systems (L2S), University Paris 11, France. His research interests include detection and parameter estimation theory in statistical signal and array processing.

Luc Vandendorpe (F'06) was born in Mouscron, Belgium, in 1962. He received the electrical engineering degree (*summa cum laude*) and the Ph.D. degree from the Université Catholique de Louvain (UCL) Louvain-la-Neuve, Belgium, in 1985 and 1991, respectively.

Since 1985, he has been with the Communications and Remote Sensing Laboratory, UCL where he first worked in the field of bit-rate reduction techniques for video coding. From March 1992 to August 1992, he was a Visiting Scientist and Research Fellow with the Telecommunications and Traffic Control Systems Group of the Delft Technical University, The Netherlands, where he worked on Spread Spectrum Techniques for Personal Communications Systems. From October 1992 to August 1997, he was a Senior Research Associate of the Belgian NSF at UCL, and an Invited Assistant Professor. Presently, he is currently with the Université Catholique de Louvain (UCL), Belgium. He is the Belgian Delegate to COST 273 and 289. He is mainly interested in digital communication systems: equalization, joint detection/synchronization for CDMA, OFDM (multicarrier), MIMO and turbo-based communications systems (UMTS, xDSL, WLAN, etc.), and joint source/channel (de)coding.

Dr. Vandendorpe was a corecipient of the Biennal Alcatel-Bell Award from the Belgian NSF for a contribution in the field of image coding in 1990. In 2000, he was a corecipient (with J. Louveaux and F. Deryck) of the Biennal Siemens Award from the Belgian NSF for a contribution about filter bank-based multicarrier transmission. In 2004, he was a co-winner (with J. Czyz) of the Face Authentication Competition, FAC 2004. He is or has been a TPC member for the IEEE VTC Fall 1999, IEEE Globecom 2003 Communications Theory Symposium, the 2003 Turbo Symposium, IEEE VTC Fall 2003, and IEEE SPAWC 2005. He is a Technical Co-Chair (with P. Duhamel) for IEEE ICASSP 2006. He is an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and a member of the Signal Processing Committee for Communications. He was an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS FOR SYNCHRONIZATION AND EQUALIZATION between 2000 and 2002, and chair of the IEEE Benelux joint chapter on Communications and Vehicular Technology between 1999 and 2003.

Annexe B

On the hybrid Cramér-Rao bound and its application to dynamical phase estimation

IEEE Signal Processing Letters, Volume : 15, 2008, pp. 453-456

IEEE SIGNAL PROCESSING LETTERS, VOL. 15, 2008 453

On the Hybrid Cramér Rao Bound and Its Application to Dynamical Phase Estimation

Stéphanie Bay, Benoit Geller, Alexandre Renaux, *Member, IEEE*, Jean-Pierre Barbot, and Jean-Marc Brossier

***Abstract*—This letter deals with the Cramér–Rao bound for the estimation of a hybrid vector with both random and deterministic parameters. We point out the specificity of the case when the deterministic and the random vectors of parameters are statistically dependent. The relevance of this expression is illustrated by studying a practical phase estimation problem in a non-data-aided communication context.**

***Index Terms*—Cramér–Rao bounds, synchronization parameters estimation.**

I. Introduction

A natural problematic when designing an estimator is the evaluation of its performance. Lower bounds on the mean square error (MSE) mainly answer this question and the well-known Cramér–Rao bound (CRB) is widely used by the signal processing community. Depending on assumptions on the parameters, the CRB has different expressions. When the vector of parameters is assumed to be deterministic, we obtain the standard CRB and when the vector of parameters is assumed to be random with an *a priori* probability density function (pdf), we obtain the so-called Bayesian CRB [1].

At the end of the 1980s, an extension combining both the standard and the Bayesian CRBs was proposed [2]. Indeed, in some practical scenarios, it is natural to represent the parameter vector by a deterministic part and by a random part. This bound has thus been called the hybrid CRB (HCRB). Until now, results available in the literature essentially focused on the case where the deterministic part and the random part of the parameter vector are assumed to be statistically independent (see, e.g., [2, eq. (5)], [3, eq. (13)] and [4, eq. (13)]). To the best of our knowledge, a closed-form expression of the HCRB with a statistical dependence between the deterministic and the random parameters has never been reported in the literature. The goal of this letter is then twofold. First, in Section II, we remind the structure of the HCRB and we point out the specificity of the case when the deterministic part and the random part of the parameter vector are statistically dependent. Second, in Section III, motivated by this analysis, we give a closed-form expression of the proposed bound in the practical case of a dynamical phase subject to a linear drift in a non-data-aided communication context.

Manuscript received December 13, 2007; revised February 10, 2008. This work was supported in part by the French ANR (Agence Nationale de la Recherche), LURGA project and in part by the European network of excellence NEWCOM++. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Xiaoli Ma.

S. Bay and J.-P. Barbot are with the Ecole Normale Supérieure de Cachan, SATIE Laboratory, 94235 Cachan, France (e-mail: bay@satie.ens-cachan.fr; barbot@satie.ens-cachan.fr).

B. Geller is with Ecole Nationale Supérieure des Techniques Avancées, Laboratory of Electronics and Computer Engineering (LEI), 75015 Paris, France (e-mail: geller@ensta.fr).

A. Renaux is with University Paris-Sud 11, Laboratory of Signals and Systems, Supélec, 91192 Gif-sur-Yvette cedex, France (e-mail: renaux@lss.supelec.fr).

J.-M. Brossier is with Grenoble Institute of Technology, GIPSA Laboratory, BP 46 38402 Saint Martin d'Hères cedex, France (e-mail: jean-marc.brossier@gipsa-lab.inpg.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2008.921461

II. Hybrid Cramér-Rao Bound

A. Background

Let $\boldsymbol{\mu} = \left(\boldsymbol{\mu}_r^T \; \boldsymbol{\mu}_d^T\right)^T \in \mathbb{R}^n$ be the parameter vector that we have to estimate. This vector is split into two sub-vectors $\boldsymbol{\mu}_d$ and $\boldsymbol{\mu}_r$, where $\boldsymbol{\mu}_d$ is assumed to be a $(n-m)\times 1$ deterministic vector and $\boldsymbol{\mu}_r$ is assumed to be a $m \times 1$ random vector with an *a priori* pdf $p(\boldsymbol{\mu}_r)$. The true value of $\boldsymbol{\mu}_d$ will be denoted $\boldsymbol{\mu}_d^\star$. We consider $\hat{\boldsymbol{\mu}}(\mathbf{y})$ as an estimator of $\boldsymbol{\mu}$, where $\mathbf{y}$ is the observation vector. The HCRB satisfies the following inequality on the MSE:

$$\mathbb{E}_{\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star}\left[(\hat{\boldsymbol{\mu}}(\mathbf{y})-\boldsymbol{\mu})(\hat{\boldsymbol{\mu}}(\mathbf{y})-\boldsymbol{\mu})^T|_{\boldsymbol{\mu}_d^\star}\right] \geq \mathbf{H}^{-1}(\boldsymbol{\mu}_d^\star) \quad (1)$$

where $\mathbf{H}(\boldsymbol{\mu}_d^\star) \in \mathbb{R}^{n\times n}$ is the so-called Hybrid Information Matrix (HIM) defined as [2]

$$\mathbf{H}(\boldsymbol{\mu}_d^\star) = \mathbb{E}_{\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star}\left[-\Delta_{\boldsymbol{\mu}}^{\boldsymbol{\mu}}\log p(\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d)\big|_{\boldsymbol{\mu}_d^\star}\right] \quad (2)$$

where $\left[\Delta_{\boldsymbol{\eta}}^{\boldsymbol{\nu}}\right]_{k,l} = \partial^2/\partial[\boldsymbol{\eta}]_k\,\partial[\boldsymbol{\nu}]_l$.

When the deterministic and the random parts of the parameter vector are assumed to be independent, and after some algebraic manipulations, the HIM can be rewritten as (see [3, eq. (18)])

$$\mathbf{H}(\boldsymbol{\mu}_d^\star) = \mathbb{E}_{\boldsymbol{\mu_r}}[\mathbf{F}(\boldsymbol{\mu}_d^\star,\boldsymbol{\mu}_r)] + \begin{pmatrix} \mathbb{E}_{\boldsymbol{\mu_r}}\left[-\Delta_{\boldsymbol{\mu_r}}^{\boldsymbol{\mu_r}}\log p(\boldsymbol{\mu}_r)\right] & \mathbf{0}_{m\times(n-m)} \\ \mathbf{0}_{(n-m)\times m} & \mathbf{0}_{(n-m)\times(n-m)} \end{pmatrix} \quad (3)$$

where

$$\mathbf{F}(\boldsymbol{\mu}_d^\star,\boldsymbol{\mu}_r) = \mathbb{E}_{\mathbf{y}|\boldsymbol{\mu}_d^\star,\boldsymbol{\mu}_r}\left[-\Delta_{\boldsymbol{\mu}}^{\boldsymbol{\mu}}\log p(\mathbf{y}|\boldsymbol{\mu}_d,\boldsymbol{\mu}_r)\big|_{\boldsymbol{\mu}_d^\star}\right]. \quad (4)$$

With this aforementioned structure, it is straightforward to reobtain the standard and the Bayesian CRBs. Indeed, if $\boldsymbol{\mu}=\boldsymbol{\mu}_d$, we have

$$\mathbf{H}^{-1}(\boldsymbol{\mu}_d^\star) = \left(\mathbb{E}_{\mathbf{y}|\boldsymbol{\mu}_d^\star}\left[-\Delta_{\boldsymbol{\mu}_d}^{\boldsymbol{\mu}_d}\log p(\mathbf{y}|\boldsymbol{\mu}_d)\big|_{\boldsymbol{\mu}_d^\star}\right]\right)^{-1} \quad (5)$$

which is the standard CRB, and, if $\boldsymbol{\mu}=\boldsymbol{\mu}_r$, we have

$$\mathbf{H}^{-1} = \left(\mathbb{E}_{\mathbf{y},\boldsymbol{\mu}_r}\left[-\Delta_{\boldsymbol{\mu}_r}^{\boldsymbol{\mu}_r}\log p(\mathbf{y}|\boldsymbol{\mu}_r)\right] + \mathbb{E}_{\boldsymbol{\mu_r}}\left[-\Delta_{\boldsymbol{\mu_r}}^{\boldsymbol{\mu_r}}\log p(\boldsymbol{\mu}_r)\right]\right)^{-1} \quad (6)$$

which is the Bayesian CRB.

B. Extension When $\boldsymbol{\mu}_r$ and $\boldsymbol{\mu}_d$ Are Statistically Dependent

We now assume a possible statistical dependence between $\boldsymbol{\mu}_r$ and $\boldsymbol{\mu}_d$. In other words, $\boldsymbol{\mu}_r$ is now assumed to be a $m\times 1$ random vector with an *a priori* pdf $p(\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star) \neq p(\boldsymbol{\mu}_r)$.

Based on the HIM definition given by (2) and expending the log-likelihood as $\log p(\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star) = \log p(\mathbf{y}|\boldsymbol{\mu}_d^\star,\boldsymbol{\mu_r}) + \log p(\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star)$, we obtain the following HIM:

$$\mathbf{H}(\boldsymbol{\mu}_d^\star) = \mathbb{E}_{\boldsymbol{\mu_r}|\boldsymbol{\mu}_d^\star}[\mathbf{F}(\boldsymbol{\mu}_d^\star,\boldsymbol{\mu}_r)] + \mathbb{E}_{\boldsymbol{\mu_r}|\boldsymbol{\mu_d}^\star}\left[-\Delta_{\boldsymbol{\mu}}^{\boldsymbol{\mu}}\log p(\boldsymbol{\mu}_r|\boldsymbol{\mu}_d)\big|_{\boldsymbol{\mu}_d^\star}\right] \quad (7)$$

where $\mathbf{F}(\boldsymbol{\mu}_d^\star,\boldsymbol{\mu}_r)$ is given by (4).

In order to explicitly show the modification in comparison with the HIM given by (3), $\mathbf{H}(\boldsymbol{\mu}_d^\star)$ can be rewritten as (8) at the bottom of the page.

Obviously, if we assume $p(\boldsymbol{\mu}_r|\boldsymbol{\mu}_d) = p(\boldsymbol{\mu}_r)$ in this expression, we straightforwardly reobtain (3).

Based on this structure, one now has to prove that there is still an inequality, i.e., a lower bound on the MSE

$$\mathbb{E}_{\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star}\left[(\hat{\boldsymbol{\mu}}(\mathbf{y})-\boldsymbol{\mu})(\hat{\boldsymbol{\mu}}(\mathbf{y})-\boldsymbol{\mu})^T\Big|_{\boldsymbol{\mu}_d^\star}\right] \geq \mathbf{H}^{-1}(\boldsymbol{\mu}_d^\star) \quad (9)$$

when $\mathbf{H}(\boldsymbol{\mu}_d^\star)$ is given by (8).

Proof: Following the idea of [3] to prove the inequality (1), one defines a vector $\mathbf{h}$ such that $\mathbf{h} = \begin{pmatrix} \nabla_{\boldsymbol{\mu}}\log p(\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d)|_{\boldsymbol{\mu}_d^\star} \\ \hat{\boldsymbol{\mu}}(\mathbf{y})-\boldsymbol{\mu}|_{\boldsymbol{\mu}_d^\star}\end{pmatrix}$, where $\nabla_{\boldsymbol{\mu}} = (\partial/\partial[\boldsymbol{\mu}]_1 \quad \dots \quad \partial/\partial[\boldsymbol{\mu}]_n)^T$.

Consequently, the nonnegative definite matrix $\mathbf{G}(\boldsymbol{\mu}_d^\star) = \mathbb{E}_{\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star}[\mathbf{h}\,\mathbf{h}^T]$ can be decomposed as the following block matrix: $\mathbf{G}(\boldsymbol{\mu}_d^\star) = \begin{pmatrix} \mathbf{H}(\boldsymbol{\mu}_d^\star) & \mathbf{L}(\boldsymbol{\mu}_d^\star) \\ \mathbf{L}^T(\boldsymbol{\mu}_d^\star) & \mathbf{R}(\boldsymbol{\mu}_d^\star)\end{pmatrix}$, where $\mathbf{R}(\boldsymbol{\mu}_d^\star)$ is the covariance matrix of $\hat{\boldsymbol{\mu}}(\mathbf{y})$, i.e., $\mathbf{R}(\boldsymbol{\mu}_d^\star) = \mathbb{E}_{\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star}\left[(\hat{\boldsymbol{\mu}}(\mathbf{y})-\boldsymbol{\mu})(\hat{\boldsymbol{\mu}}(\mathbf{y})-\boldsymbol{\mu})^T\Big|_{\boldsymbol{\mu}_d^\star}\right]$ and, where $\mathbf{L}(\boldsymbol{\mu}_d^\star)$ is given by $\mathbf{L}(\boldsymbol{\mu}_d^\star) = \mathbb{E}_{\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star} \times \left[\nabla_{\boldsymbol{\mu}}\log p(\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d)|_{\boldsymbol{\mu}_d^\star}\left(\hat{\boldsymbol{\mu}}(\mathbf{y})-\boldsymbol{\mu}|_{\boldsymbol{\mu}_d^\star}\right)^T\right]$.

Since $\mathbf{G}(\boldsymbol{\mu}_d^\star) \geq 0$, its Schur complement satisfies $\mathbf{R}(\boldsymbol{\mu}_d^\star) \geq \mathbf{L}^T(\boldsymbol{\mu}_d^\star)\mathbf{H}^{-1}(\boldsymbol{\mu}_d^\star)\mathbf{L}(\boldsymbol{\mu}_d^\star)$. ■

It is straightforward to show that, for an unbiased estimator w.r.t. the pdf $p(\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star)$, $\mathbf{L}(\boldsymbol{\mu}_d^\star) = \mathbf{I}_{n\times n}$.

Consequently, the inequality (9) is proved and $\mathbf{H}^{-1}(\boldsymbol{\mu}_d^\star)$ is a lower bound on the MSE.

III. HCRB for a Dynamical Phase Estimation Problem

In [4], we have proposed a closed-form expression of the Bayesian CRB for the estimation of the phase offset for a BPSK transmission in a non-data-aided context. In this section, we extend these previous results by providing a closed-form expression of the HCRB for the estimation of the phase offset and also of the linear drift.

A. Observation and State Models

We consider a linearly modulated signal, obtained by applying to a square-root Nyquist transmit filter an unknown symbol sequence $\mathbf{a} = (a_1\cdots a_K)^T$ taken from a unit energy BPSK constellation. The signal is transmitted over an additive white Gaussian noise channel. The output signal is sampled at the symbol rate which yields to the observations

$$y_k = a_k e^{j\theta_k} + n_k \text{ with } k = 1\dots K \quad (10)$$

where $\{n_k\}$ is a sequence of i.i.d., circular, zero-mean complex Gaussian noise variables with variance σ_n^2. We consider that the system operates in a non-data-aided synchronization mode, i.e., the transmitted symbols are i.i.d. with $P_r(a_k = \pm 1) = 1/2$.

In practice, several sources of distortions affect the phase. An efficient model representing these effects is the so-called Brownian phase with a linear drift widely studied in the literature. The Brownian phase model with a linear drift is given as follows:

$$\theta_k = \theta_{k-1} + \xi + w_k \text{ with } k = 2\dots K \quad (11)$$

where, for any index k, $\{\theta_k\}$ is the sequence of phases to be estimated, ξ represents the deterministic unknown linear drift with true value $\xi^\star$, and where $\{w_k\}$ is an i.i.d. sequence of centered Gaussian random variables with known variance σ_w^2.

The parameter vector of interest is then made up of both random and deterministic parameters $\boldsymbol{\mu} = \left(\boldsymbol{\mu}_r^T\ \mu_d\right)^T$, where $\boldsymbol{\mu}_r = \boldsymbol{\theta} = (\theta_1\dots\theta_K)^T$ and $\mu_d = \xi$. Moreover, from (16), it is clear that $p(\boldsymbol{\theta}|\xi^\star) \neq p(\boldsymbol{\theta})$.

B. Derivation of the HCRB

For notational convenience, we drop the dependence of the different matrices on $\mu_d^\star = \xi^\star$ in the remainder of this letter. From (8), the HIM $\mathbf{H}$ can be rewritten into a block matrix $\mathbf{H} = \begin{pmatrix} \mathbf{H}_{11} & \mathbf{h}_{12} \\ \mathbf{h}_{21} & H_{22}\end{pmatrix}$, where we see (12), shown at the bottom of the next page.

These blocks only depend on the log-likelihoods $\log p(\mathbf{y}|\boldsymbol{\theta},\xi^\star)$ and $\log p(\boldsymbol{\theta}|\xi^\star)$. Let us set $\mathbf{y} = (y_1\cdots y_K)^T$ and assume that the initial phase θ_1 does not depend on ξ, i.e., $p(\theta_1|\xi^\star) = p(\theta_1)$. Using (10) and (11), i.e., the Gaussian nature of the noise and the equiprobability of the symbols, one has to see (13), shown at the bottom of the page.

- Expression of $\mathbf{H}_{11}$: assuming that we have no prior knowledge, i.e., $\mathbb{E}_{\theta_1}\left[\Delta_{\theta_1}^{\theta_1}\log p(\theta_1)\right] = 0$, it is shown in [4] (due

$$\mathbf{H}(\boldsymbol{\mu}_d^\star) = \mathbb{E}_{\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star}[\mathbf{F}(\boldsymbol{\mu}_d^\star,\boldsymbol{\mu}_r)] + \mathbb{E}_{\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star} \times \left[\begin{pmatrix} -\Delta_{\boldsymbol{\mu}_r}^{\boldsymbol{\mu}_r}\log p(\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star) & -\Delta_{\boldsymbol{\mu}_d}^{\boldsymbol{\mu}_r}\log p(\boldsymbol{\mu}_r|\boldsymbol{\mu}_d)\big|_{\boldsymbol{\mu}_d^\star} \\ \left(-\Delta_{\boldsymbol{\mu}_d}^{\boldsymbol{\mu}_r}\log p(\boldsymbol{\mu}_r|\boldsymbol{\mu}_d)\big|_{\boldsymbol{\mu}_d^\star}\right)^T & -\Delta_{\boldsymbol{\mu}_d}^{\boldsymbol{\mu}_d}\log p(\boldsymbol{\mu}_r|\boldsymbol{\mu}_d)\big|_{\boldsymbol{\mu}_d^\star}\end{pmatrix}\right] \quad (8)$$

to the order one Markov structure exhibited by (11)) that $\mathbf{H}_{11}$ takes the following tridiagonal structure:

$$\mathbf{H}_{11} = b \begin{pmatrix} A+1 & 1 & 0 & \dots & 0 \\ 1 & A & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & A & 1 \\ 0 & \dots & 0 & 1 & A+1 \end{pmatrix} \quad (14)$$

where $b = -1/\sigma_w^2$, and where $A = -\sigma_w^2 J_D - 2$ with $J_D = \mathbb{E}_{\mathbf{y},\boldsymbol{\theta}|\xi^\star}\left[-\Delta_{\theta_k}^{\theta_k} \log p\left(y_k|\theta_k,\xi^\star\right)\right]$.

- Expression of $\mathbf{h}_{12}$: since, from (18), $\log p(\mathbf{y}|\boldsymbol{\theta},\xi^\star)$ is independent of $\xi^\star$, $\Delta_\theta^\xi \log p(\mathbf{y}|\boldsymbol{\theta},\xi)\Big|_{\xi^\star} = 0$. Consequently

$$\mathbf{h}_{12} = \mathbb{E}_{\boldsymbol{\theta}|\xi^\star}\left[-\Delta_{\boldsymbol{\theta}}^{\xi} \log p\left(\boldsymbol{\theta}|\xi\right)\Big|_{\xi^\star}\right]. \quad (15)$$

Using the state model, we have

$$\begin{cases} \Delta_\xi^{\theta_1} \log p\left(\boldsymbol{\theta}|\xi\right)\Big|_{\xi^\star} = -\frac{1}{\sigma_w^2} \\ \Delta_\xi^{\theta_K} \log p\left(\boldsymbol{\theta}|\xi\right)\Big|_{\xi^\star} = \frac{1}{\sigma_w^2} \\ \Delta_\xi^{\theta_k} \log p\left(\boldsymbol{\theta}|\xi\right)\Big|_{\xi^\star} = 0 \text{ for } k \in \{2,\dots,K-1\}. \end{cases}$$

Applying the expectation operator $\mathbb{E}_{\boldsymbol{\theta}|\xi^\star}[.]$, we obtain

$$\mathbf{h}_{12} = \left(\begin{array}{ccc} \frac{1}{\sigma_w^2} & \mathbf{0}_{1\times K-2} & -\frac{1}{\sigma_w^2} \end{array}\right)^T. \quad (16)$$

- Expression of H_{22}: since, from (13), $\log p(\mathbf{y}|\boldsymbol{\theta},\xi^\star)$ is independent of $\xi^\star$, $\Delta_\xi^\xi \log p(\mathbf{y}|\boldsymbol{\theta},\xi)\Big|_{\xi^\star} = 0$. Consequently

$$H_{22} = \mathbb{E}_{\boldsymbol{\theta}|\xi^\star}\left[-\Delta_\xi^\xi \log p\left(\boldsymbol{\theta}|\xi\right)\Big|_{\xi^\star}\right] = \frac{K-1}{\sigma_w^2}. \quad (17)$$

- **Expression of the HCRB**: we now give the expression of $\mathbf{H}^{-1}$ which bounds the MSE. Thanks to the block-matrix inversion formula, we have

$$\mathbf{H}^{-1} = \begin{pmatrix} \mathbf{H}_{11}^{-1} + \mathbf{V}_K & -\frac{1}{\lambda}\mathbf{H}_{11}^{-1}\mathbf{h}_{12} \\ -\frac{1}{\lambda}\mathbf{h}_{12}^T\mathbf{H}_{11}^{-1} & \frac{1}{\lambda} \end{pmatrix} \quad (18)$$

where $\lambda = K - 1/\sigma_w^2 - \mathbf{h}_{12}^T\mathbf{H}_{11}^{-1}\mathbf{h}_{12}$ and $\mathbf{V}_K = 1/\lambda \mathbf{H}_{11}^{-1}\mathbf{h}_{12}\mathbf{h}_{12}^T\mathbf{H}_{11}^{-1}$.

We start to compute λ corresponding to the inverse of the minimal bound on the MSE of ξ. Due to the particular structure of matrices $\mathbf{H}_{11}$ and $\mathbf{h}_{12}$ (14), (16), we obtain $\lambda = \frac{K-1}{\sigma_w^2} - \frac{2}{\sigma_w^4}\left(\left[\mathbf{H}_{11}^{-1}\right]_{1,1} - \left[\mathbf{H}_{11}^{-1}\right]_{1,K}\right)$.

From (14), thanks to the cofactor expression in the matrix inversion formula, we have for any index k, $\left[\mathbf{H}_{11}^{-1}\right]_{1,k} = b^{k-1}/|\mathbf{H}_{11}|\,(d_{K-k} + b\,d_{K-k-1})$, where d_k is the determinant of a $k \times k$ matrix $\mathbf{D}_k$. equal to the matrix of (14) without the plus one on each corner.

The sequence $\{d_k\}$ satisfies the following recursive equation: $d_k = A\,b\,d_{k-1} - b^2\,d_{k-2}$ with $d_0 = 1$ and $d_1 = bA$. d_k can thus be written as $d_k = \rho_1 (r_1)^k + \rho_2 (r_2)^k$, where r_1, r_2, ρ_1, and ρ_2 are given by

$$\begin{cases} r_1 = \frac{b}{2}\left(A + \sqrt{A^2-4}\right), & r_2 = \frac{b}{2}\left(A - \sqrt{A^2-4}\right) \\ \rho_1 = \frac{\sqrt{A^2-4}+A}{2\sqrt{A^2-4}}, & \rho_2 = \frac{\sqrt{A^2-4}-A}{2\sqrt{A^2-4}}. \end{cases} \quad (19)$$

Consequently

$$\left[\mathbf{H}_{11}^{-1}\right]_{1,k} = \frac{b^{k-1}}{|\mathbf{H}_{11}|} \times \left(\rho_1 r_1^{K-k-1}(r_1+b) + \rho_2 r_2^{K-k-1}(r_2+b)\right) \quad (20)$$

and $\lambda = \frac{K-1}{\sigma_w^2} - \frac{2}{\sigma_w^4|\mathbf{H}_{11}|} \times \left(\rho_1 r_1^{K-2}(r_1+b) + \rho_2 r_2^{K-2}(r_2+b) - b^{K-1}\right)$.

From the definition of $\mathbf{V}_K$, we have

$$[\mathbf{V}_K]_{k,k} = \frac{1}{\lambda\sigma_w^4}\left(\left[\mathbf{H}_{11}^{-1}\right]_{1,k} - \left[\mathbf{H}_{11}^{-1}\right]_{1,K+1-k}\right)^2. \quad (21)$$

Using (18), (20), and (21), we obtain, for any index k, the analytical expression of the HCRB diagonal elements in (22), shown at the bottom of the next page.

Remark: Note that, if (3) was used instead of (8), the HIM would not be invertible.

C. Simulation Results

We now illustrate the behavior of the HCRB versus the signal-to-noise ratio (SNR) defined by $1/\sigma_n^2$. We consider a block of $K = 40$ BPSK transmitted symbols. For two distinct phase-noise variances ($\sigma_w^2 = 0.1$ rad^2 and $\sigma_w^2 \to 0$ rad^2), Fig. 1 superimposes on one side the HCRB [see (30)], the data-aided HCRB $\left(J_D = 2/\sigma_n^2\right)$, and the BCRB (see [4, eq.

$$\begin{cases} \mathbf{H}_{11} = \mathbb{E}_{\mathbf{y},\boldsymbol{\theta}|\xi^\star}\left[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\mathbf{y}|\boldsymbol{\theta},\xi\right)\Big|_{\xi^\star}\right] + \mathbb{E}_{\boldsymbol{\theta}|\xi^\star}\left[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\boldsymbol{\theta}|\xi^\star\right)\right] \\ \mathbf{h}_{12} = \mathbf{h}_{21}^T = \mathbb{E}_{\mathbf{y},\boldsymbol{\theta}|\xi^\star}\left[-\Delta_{\boldsymbol{\theta}}^{\xi} \log p\left(\mathbf{y}|\boldsymbol{\theta},\xi\right)\Big|_{\xi^\star}\right] + \mathbb{E}_{\boldsymbol{\theta}|\xi^\star}\left[-\Delta_{\boldsymbol{\theta}}^{\xi} \log p\left(\boldsymbol{\theta}|\xi\right)\Big|_{\xi^\star}\right] \\ H_{22} = \mathbb{E}_{\mathbf{y},\boldsymbol{\theta}|\xi^\star}\left[-\Delta_\xi^\xi \log, p\left(\mathbf{y}|\boldsymbol{\theta},\xi\right)\Big|_{\xi^\star}\right] + \mathbb{E}_{\boldsymbol{\theta}|\xi^\star}\left[-\Delta_\xi^\xi \log p\left(\boldsymbol{\theta}|\xi\right)\Big|_{\xi^\star}\right] \end{cases} \quad (12)$$

$$\begin{cases} \log p\left(\mathbf{y}|\boldsymbol{\theta},\xi^\star\right) = \sum_{k=1}^{K}\left(-\log\left(\pi\sigma_n^2\right) - \frac{1+\|y_k\|^2}{\sigma_n^2} + \log\left(\cosh\left(\frac{2}{\sigma_n^2}\Re\left\{y_k e^{-j\theta_k}\right\}\right)\right)\right) \\ \log p\left(\boldsymbol{\theta}|\xi^\star\right) = \log p(\theta_1) + (K-1)\log\left(\frac{1}{\sqrt{2\pi}\sigma_w}\right) - \sum_{k=2}^{K}\frac{(\theta_k - \theta_{k-1} - \xi^\star)^2}{2\sigma_w^2} \end{cases} \quad (13)$$

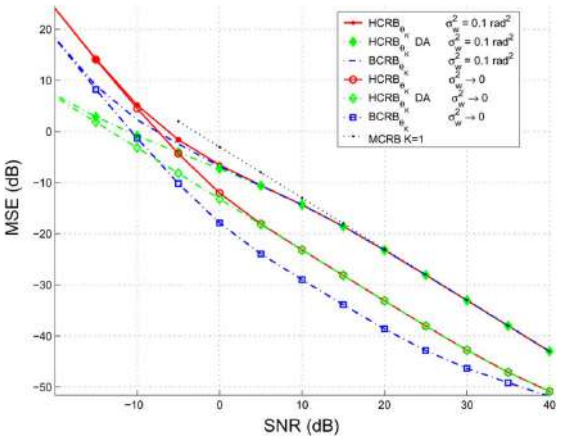


Fig. 1. Bounds on θ_K versus the SNR ($K = 40$ observations, $\sigma_w^2 = 0.1\,\mathrm{rad}^2$, and $\sigma_w^2 \to 0\,\mathrm{rad}^2$, J_D evaluated over 10^8 Monte Carlo trials).

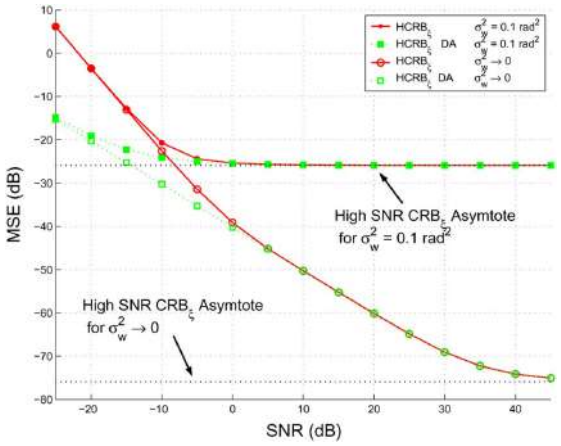


Fig. 2. Bounds on ξ versus the SNR ($K = 40$ observations, $\sigma_w^2 = 0.1\,\mathrm{rad}^2$ and $\sigma_w^2 \to 0\,\mathrm{rad}^2$, J_D evaluated over 10^8 Monte Carlo trials).

(21)]) on θ_K. For the same scenario, Fig. 2 superimposes on one side the HCRB and the data-aided HCRB on ξ.

At high SNR, we first notice that $HCRB_\xi$ converges to its horizontal asymptote given by $\sigma_w^2/K - 1$ which is the standard CRB when $\boldsymbol{\theta}$ is assumed to be known. The observation noise compared to the phase noise is then not significant enough to disturb the estimation of ξ; consequently $HCRB_\xi$ depends only on the phase noise and on the number of observations. Concerning the bounds on θ_K, $HCRB_{\theta_k}$ and $BCRB_{\theta_k}$ both have the same asymptote given by $\sigma_n^2/2$ which is the modified CRB (MCRB) for one observation (see [6]). It means that, at high SNR, the observation y_K is self-sufficient to estimate θ_K, and the error on ξ does not disturb the performance on θ_K. Moreover, the HCRB logically tends to the data-aided HCRB.

For median SNR, $HCRB_{\theta_K}$ and $HCRB_\xi$ leave their respective asymptote. $HCRB_{\theta_K}$ is still lower bounded by the BCRB and upper bounded by the high-SNR asymptote. This stems from the fact that taking into account a block of observations instead of one observation necessarily improves the performance. However, for large σ_w^2 values (e.g., $\sigma_w^2 = 0.1\,\mathrm{rad}^2$), $HCRB_{\theta_K}$ stays close to the MCRB because the correlation between the phase offsets θ_k is less significant than the information brought by the observation y_K. Moreover, when σ_w^2 tends to 0, $HCRB_{\theta_K}$ is above the BCRB because performance is now limited by the accuracy on the parameter ξ.

At low SNR, n_k is preponderant compared to w_k. Both $HCRB_\xi$ and $HCRB_{\theta_K}$ do not depend on σ_w^2: the lack of knowledge on ξ directly affects the estimation on θ_K. As expected, the knowledge of the symbols (data-aided HCRB) leads to a better estimation of $\boldsymbol{\theta}$ and ξ.

IV. Conclusion

In this letter, we have studied the hybrid Cramér–Rao bound when the random and the deterministic parts of the parameter vector are statistically dependent. We have applied this bound in order to evaluate the performance of a dynamical phase estimator where the linear drift is unknown in a non-data-aided context.

References

[1] H. L. Van Trees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968, vol. 1.

[2] Y. Rockah and P. Schultheiss, "Array shape calibration using sources in unknown locations-part I: Far-field sources," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 3, pp. 286–299, Mar. 1987.

[3] H. Messer, "The hybrid Cramér-Rao lower bound—From practice to theory," in *Proc. IEEE Workshop Sensor Array and Multi-Channel Processing (SAM)*, Waltham, MA, Jul. 2006, pp. 304–307.

[4] S. Bay, C. Herzet, J.-M. Brossier, J.-P. Barbot, and B. Geller, "Analytic and asymptotic analysis of Bayesian Cramér-Rao bound for dynamical phase offset estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 61–70, Jan. 2008.

[5] J. A. McNeill, "Jitter in ring oscillators," Ph.D. dissertation, Boston Univ., Boston, MA, 1994.

[6] M. Moeneclaey, "On the true and the modified Cramér-Rao bounds for the estimation of a scalar parameter in the presence of nuisance parameters," *IEEE Trans. Commun.*, vol. 46, no. 11, pp. 1536–1544, Nov. 1998.

$$
\begin{aligned}
\left[\mathbf{H}^{-1}\right]_{k,k} = \frac{1}{|\mathbf{H}_{11}|}\left[\rho_1^2\,(b+r_1)^2\,r_1^{K-3} + \rho_2^2 (b+r_2)^2 r_2^{K-3} - \frac{b^2}{A-2}\left(r_1^{k-2} r_2^{K-k-1} + r_1^{K-k-1} r_2^{k-2}\right)\right] \\
+ \frac{1}{\lambda\sigma_w^4 |\mathbf{H}_{11}|^2}\left[b^{k-1}\left(\rho_1\,(r_1)^{K-k-1}\,(b+r_1) + \rho_2\,(r_2)^{K-k-1}\,(b+r_2)\right)\right. \\
\left. + b^{K-k}\left(\rho_1\,(r_1)^{k-2}\,(b+r_1) + \rho_2\,(r_2)^{k-2}\,(b+r_2)\right)\right]^2
\end{aligned}
\tag{22}
$$

Annexe C

A fresh look at the Bayesian bounds of the Weiss-Weinstein family

IEEE Transactions on Signal Processing, Volume : 56, Issue : 11, Nov. 2008, pp. 5334-5352

5334 IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 56, NO. 11, NOVEMBER 2008

A Fresh Look at the Bayesian Bounds of the Weiss-Weinstein Family

Alexandre Renaux, *Member, IEEE*, Philippe Forster, *Member, IEEE*, Pascal Larzabal, *Member, IEEE*, Christ D. Richmond, *Senior Member, IEEE*, and Arye Nehorai, *Fellow, IEEE*

***Abstract*—Minimal bounds on the mean square error (MSE) are generally used in order to predict the best achievable performance of an estimator for a given observation model. In this paper, we are interested in the Bayesian bound of the Weiss–Weinstein family. Among this family, we have Bayesian Cramér-Rao bound, the Bobrovsky–MayerWolf–Zakaï bound, the Bayesian Bhattacharyya bound, the Bobrovsky–Zakaï bound, the Reuven–Messer bound, and the Weiss–Weinstein bound. We present a unification of all these minimal bounds based on a rewriting of the minimum mean square error estimator (MMSEE) and on a constrained optimization problem. With this approach, we obtain a useful theoretical framework to derive new Bayesian bounds. For that purpose, we propose two bounds. First, we propose a generalization of the Bayesian Bhattacharyya bound extending the works of Bobrovsky, Mayer–Wolf, and Zakaï. Second, we propose a bound based on the Bayesian Bhattacharyya bound and on the Reuven–Messer bound, representing a generalization of these bounds. The proposed bound is the Bayesian extension of the deterministic Abel bound and is found to be tighter than the Bayesian Bhattacharyya bound, the Reuven–Messer bound, the Bobrovsky–Zakaï bound, and the Bayesian Cramér–Rao bound. We propose some closed-form expressions of these bounds for a general Gaussian observation model with parameterized mean. In order to illustrate our results, we present simulation results in the context of a spectral analysis problem.**

***Index Terms*—Bayesian bounds on the MSE, Weiss–Weinstein family.**

Manuscript received April 24, 2007; revised April 18, 2008 First published June 13, 2008; current version published October 15, 2008. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Philippe Loubaton. The work of A. Renaux and A. Nehorai was supported in part by the Department of Defense under the Air Force Office of Scientific Research MURI Grant FA9550-05-1-0443, AFOSR Grant FA9550-05-1-0018, and by the National Science Foundation by Grants CCR-0330342 and CCF-0630734. The material in this paper was presented in part at the IEEE Workshop on Statistical Signal Processing, Bordeaux, France, July 2005 and at the IEEE International Conference on Acoustic, Speech and Signal Processing, Toulouse, France, May 2006.

A. Renaux is with the Laboratoire des Signaux et Systèmes, Laboratory of Signals and Systems, University Paris-Sud 11, 91192 Gif-sur-Yvette cedex, France (e-mail: renaux@lss.supelec.fr).

P. Forster is with the Université de Nanterre, IUT de Ville d'Avray, France. He is also with the SATIE Laboratory (École Normale Supérieure de Cachan), 94235 Cachan, France (e-mail: pforster@u-paris10.fr).

P. Larzabal is with the University Paris-Sud 11, 91192 Gif-sur-Yvette cedex, France. He is also with the SATIE Laboratory (École Normale Supérieure de Cachan), 94235 Cachan, France (e-mail: larzabal@satie.ens-cachan.fr).

C. D. Richmond is with the Advanced Sensor Techniques Group, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA 02420 USA (e-mail: christ@ll.mit.edu).

A. Nehorai is with Washington University, St. Louis, MO 63130 USA (e-mail: nehorai@ese.wustl.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2008.927075

NOTATIONS

The notational convention adopted in this paper is as follows: italic indicates a scalar quantity, as in A; lowercase boldface indicates a vector quantity, as in $\mathbf{a}$; uppercase boldface indicates a matrix quantity, as in $\mathbf{A}$. $\mathrm{Re}\{A\}$ is the real part of A and $\mathrm{Im}\{A\}$ is the imaginary part of A. The complex conjugation of a quantity is indicated by a superscript * as in A^*. The matrix transpose is indicated by a superscript^T as in $\mathbf{A}^T$, and the complex conjugate plus matrix transpose is indicated by a superscript^H as in $\mathbf{A}^H = (\mathbf{A}^T)^*$. The n-th row and m-th column element of the matrix $\mathbf{A}$ is denoted by $\{\mathbf{A}\}_{n,m}$. $\mathbf{I}_N$ denotes the identity matrix of size $N \times N$. $\mathbf{0}_{N\times M}$ is a $N \times M$ matrix of zeros. $||\cdot||$ denotes the norm. $|\cdot|$ denotes the modulus. $\mathrm{abs}(\cdot)$ denotes the absolute value. $\delta(\cdot)$ denotes the Dirac delta function. $\mathbb{E}[\cdot]$ denotes the expectation operator with respect to a density probability function explicitly given by a subscript. Ω is the observation space and Θ is the parameter space.

I. INTRODUCTION

MINIMAL bounds on the mean square error (MSE) provide the ultimate performance that an estimator can expect to achieve for a given observation model. Consequently, they are used as a benchmark in order to evaluate the performance of an estimator and to determine if an improvement is possible. The Cramér–Rao bound [3]–[8] has been the most widely used by the signal processing community and is still under investigation from a theoretical point of view (particularly throughout the differential variety in the Riemannian geometry framework [9]–[14]) as from a practical point of view (see, e.g., [15]–[19]). But the Cramér–Rao bound suffers from some drawbacks when the scenario becomes critical. Indeed, in a nonlinear estimation problem, when the parameters have finite support, there are three distinct MSE areas for an estimator [20, p. 273], [21]. For a large number of observations or for a high signal-to-noise ratio (SNR), the estimator MSE is small and the area is called an asymptotic area. When the scenario becomes critical, i.e., when the number of observations or the SNR decreases, the estimator MSE increases dramatically due to the outlier effect, and the area is called threshold area. Finally, when the number of observations or the SNR is low, the estimator criterion is hugely corrupted by the noise and becomes a quasi-uniform random variable on the parameter support. Since in this last area the observations bring almost no information, it is called no information area. The Cramér–Rao bound is used only in the asymptotic area and is not able to handle the threshold phenomena (i.e., when the performance breaks down).

To fill this lack, a plethora of other minimal bounds tighter than the Cramér–Rao bound has been proposed and studied. All these bounds have been derived by way of several inequalities, such as the Cauchy-Schwartz inequality, the Kotelnikov inequality, the Hölder inequality, the Ibragimov-Hasminskii inequality, the Bhattacharyya inequality, and the Kiefer inequality. Note that due to this diversity, it is sometimes difficult to fully understand the underlying concept and the difference between all these bounds; consequently it is difficult to apply these bounds to a specific estimation problem.

Minimal bounds on the MSE can be divided into two categories: the deterministic bounds for situations in which the true vector of the parameters $\boldsymbol{\theta}_0$ is assumed to be deterministic and the Bayesian bounds for situations in which the vector of parameters $\boldsymbol{\theta}$ is assumed to be random with an *a priori* probability density function $p(\boldsymbol{\theta})$. Among the deterministic bounds, we have the well-known Cramér–Rao bound; the Bhattacharyya bound [22], [23]; the Chapman-Robbins bound [24]–[26], the Barankin bound [27], [28], the Abel bound [29]–[31]; and the Quinlan-Chaumette-Larzabal bound [32]. Bayesian bounds can be subdivided into two categories: the Ziv-Zakaï family, derived from a binary hypothesis testing problem (and more generally from an M-ary hypothesis testing problem), and the Weiss-Weinstein family, derived (as the deterministic bounds) from a covariance inequality principle. The Ziv-Zakaï family contains the Ziv-Zakaï bound [33], the Bellini-Tartara bound [34], the Chazan-Zakaï-Ziv bound [35], the Weinstein bound [36], the Bell-Steinberg-Ephraim-VanTrees bound [37], and the Bell bound [38]. The Weiss-Weinstein family contains the Bayesian Cramér–Rao bound [20, pp. 72 and 84], the Bobrovsky-Mayer-Wolf-Zakaï bound [39], the Bayesian Bhattacharyya bound [20, p. 149], the Bobrovsky-Zakaï bound [40], the Reuven-Messer bound [41], and the Weiss-Weinstein bound [42]. A nice tutorial about both families can be found in the recent book of Van Trees and Bell [43].

The deterministic bounds are used as a lower bound of the *local* MSE in $\boldsymbol{\theta}_0$; i.e.,

$$\mathbf{MSE}_L(\boldsymbol{\theta}_0) = \int_{\Omega} (\hat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta}_0)^T p(\mathbf{y}|\boldsymbol{\theta}_0) d\mathbf{y} \quad (1)$$

where $\mathbf{y} \in \Omega$ is a complex observation vector, $p(\mathbf{y}|\boldsymbol{\theta}_0)$ is the likelihood of the observations parameterized by the true parameter value $\boldsymbol{\theta}_0$, and $\hat{\boldsymbol{\theta}}(\mathbf{y})$ is an estimator of $\boldsymbol{\theta}_0$.

On the other hand, Bayesian bounds are used as a lower bound of the *global* MSE; i.e.,

$$\mathbf{MSE}_G = \int_{\Theta} \int_{\Omega} (\hat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta})^T p(\mathbf{y}, \boldsymbol{\theta}) d\mathbf{y} d\boldsymbol{\theta} \quad (2)$$

where $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ is the random parameter vector with an *a priori* probability density function $p(\boldsymbol{\theta}) = (p(\mathbf{y}, \boldsymbol{\theta})/p(\mathbf{y}|\boldsymbol{\theta}))$ and $p(\mathbf{y}, \boldsymbol{\theta})$ is the joint probability function of the observations and of the parameters.

In the deterministic context, minimal bounds—in particular the Chapman-Robbins bound and the Barankin bound—are generally used in order to predict the aforementioned threshold effect which cannot be handled by the Cramér–Rao bound. The Chapman-Robbins bound and the Barankin bound have already been successfully applied to several estimation problems [28], [31], [44]–[55]. The use of the Abel bound, which can also handle the threshold phenomena, is still marginal [56].

Contrary to the deterministic bounds, the Bayesian bounds take into account the parameter support throughout the *a priori* probability density function $p(\boldsymbol{\theta})$, and they give the ultimate performances of an estimator on the three aforementioned areas of the global MSE. These bounds give the performance of the Bayesian estimator, such as the maximum *a posteriori* (MAP) estimator or the minimum mean square error estimator (MMSEE), and can be used in order to know the global performance of the deterministic estimators such as the maximum likelihood estimator (MLE), since

$$\mathbf{MSE}_G = \int_{\Theta} \mathbf{MSE}_L(\boldsymbol{\theta}) p(\boldsymbol{\theta})\, d\boldsymbol{\theta}. \quad (3)$$

The reader is referred to Xu *et al.* [57]–[59], where the MLE performances are analyzed in the context of an underwater acoustic problem by way of the Ziv-Zakaï and of the Weiss-Weinstein bounds. The Ziv-Zakaï family bounds have been applied in other signal processing areas: time-delay estimation [60]; direction-of-arrival estimation [38], [61], [62]; and digital communication [63]. On the other hand, the Weiss-Weinstein bound has been less investigated: the aforementioned Xu *et al.* works and in the framework of digital communication [64].

This article presents a new unified approach for the establishment of the Weiss-Weinstein family bounds. Note that the unification of the deterministic bounds has already been proposed by [65] and [66] based on a constrained optimization problem. A unification has been proposed by Bell *et al.* in [37] and [38] for the Ziv-Zakaï family.

Concerning the Weiss-Weinstein family unification, a first approach has been given by Weiss and Weinstein in [67]. This approach is based on the following inequality proved by the authors:

$$\mathrm{MSE}_G \geq \frac{\mathbb{E}^2_{\mathbf{y},\theta}[\theta\psi(\mathbf{y},\theta)]}{\mathbb{E}_{\mathbf{y},\theta}[\psi^2(\mathbf{y},\theta)]} \quad (4)$$

where the function $\psi(\mathbf{y},\theta)$ must satisfied

$$\int_{\Theta} \psi(\mathbf{y},\theta) p(\mathbf{y},\theta) d\theta = 0. \quad (5)$$

Weiss and Weinstein gave several functions $\psi(\mathbf{y},\theta)$ satisfying (5) for which they again obtain the Bayesian Cramér–Rao bound, the Bayesian Bhattacharyya bound, the Bobrovsky-Zakaï bound, and the Weiss-Weinstein bound. Moreover, a function $\psi(\mathbf{y},\theta)$ satisfying (5) leading to the Bobrovsky-MayerWolf-Zakaï bound is given in [38]. Unfortunately, there are no general rules to find $\psi(\mathbf{y},\theta)$. In this contribution, the Weiss-Weinstein family unification is based on the best Bayesian bound, i.e., the MSE of the MMSEE. By rewriting the MMSEE and by using a constrained optimization problem similar to one derived for the unification of deterministic bounds [65], [66], we unify the Bayesian Cramér–Rao bound, the Bobrovsky-MayerWolf-Zakaï bound, the Bayesian Bhattacharyya bound, the Bobrovsky-Zakaï

5336 IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 56, NO. 11, NOVEMBER 2008

bound, the Reuven-Messer bound (for which no function $\psi(\mathbf{y},\theta)$ is proposed in the Weiss-Weinstein approach), and the Weiss-Weinstein bound. This approach brings a useful theoretical framework to derive new Bayesian bounds.

For that purpose, we propose two bounds. First, we propose a generalization of the Bayesian Bhattacharyya bound extending the works of Bobrovsky, Mayer-Wolf, and Zakaï. Second, we propose a bound based on the Bayesian Bhattacharyya bound and on the Reuven-Messer bound, one that represents a generalization of these bounds. This bound is found to be tighter than the Bayesian Bhattacharyya bound, the Reuven-Messer bound, the Bobrovsky-Zakaï bound, and the Bayesian Cramér–Rao bound. In order to illustrate our results, we propose some closed-form expressions of the minimal bounds for a Gaussian observation model with parameterized mean widely used in signal processing, and we apply it to a spectral analysis problem for which we present simulation results.

II. MMSE Reformulation

In this section, we start by reformulating the MMSEE as a constrained optimization problem. Then, we rewrite the underlying constraint under three different forms that will be of interest for our proposed unification.

In the Bayesian framework, the minimal global MSE and consequently the best Bayesian bound is the MSE of the MMSEE: $\hat{\theta}(\mathbf{y}) = \int_{\Theta} \theta p(\theta|\mathbf{y}) d\theta$, where $p(\theta|\mathbf{y})$ is the *a posteriori* probability density function of the parameter. Unfortunately, it is generally impossible to obtain a closed-form expression of this MSE. The MMSEE is the solution of the following problem:

$$\min_{\hat{\theta}(\mathbf{y})} \int_{\Theta} \int_{\Omega} (\hat{\theta}(\mathbf{y}) - \theta)^2 p(\mathbf{y},\theta) d\mathbf{y} d\theta. \tag{6}$$

Let $\mathcal{L}_p^2$ be the set of function $v(\mathbf{y},\theta)$ such that $\int_{\Theta}\int_{\Omega} v^2(\mathbf{y},\theta)p(\mathbf{y},\theta)d\mathbf{y}d\theta$ is defined. Let $\mathcal{C}_1 \subset \mathcal{L}_p^2$ be the subset of function satisfying

$$v(\mathbf{y},\theta) = z(\mathbf{y}) - \theta \tag{7}$$

where $z(\mathbf{y})$ is a function only of $\mathbf{y}$.

Consequently, the MMSEE (6) is the solution of the following constrained optimization problem:

$$\begin{cases} \min_v \int_{\Theta}\int_{\Omega} v^2(\mathbf{y},\theta)p(\mathbf{y},\theta)d\mathbf{y}d\theta \\ \text{subject to } v(\mathbf{y},\theta) \in \mathcal{C}_1 \end{cases}. \tag{8}$$

Let $\mathcal{F}$ be the set of functions $f(\mathbf{y},\theta)$ such that $\forall v(\mathbf{y},\theta) \in \mathcal{L}_p^2$

$$\begin{cases} \lim_{\theta\to\pm\infty} v(\mathbf{y},\theta) f(\mathbf{y},\theta) = 0 \\ \lim_{\theta\to\pm\infty} \frac{\partial f(\mathbf{y},\theta)}{\partial\theta} = 0. \end{cases} \tag{9}$$

Let us now introduce the three following subsets of functions belonging to $\mathcal{L}_p^2$:

- Subset $\mathcal{C}_2$ [see (10)–(12) shown at the bottom of the page];
- Subset $\mathcal{C}_3$;
- Subset $\mathcal{C}_4$ with $L(\mathbf{y},\eta,\theta) = (f(\mathbf{y},\eta)/f(\mathbf{y},\theta))$.

Theorem 1 shows that, although these four subsets are generated in a different manner, they are the same.

Theorem 1:

$$\mathcal{C}_1 = \mathcal{C}_2 = \mathcal{C}_3 = \mathcal{C}_4. \tag{13}$$

The Proof of Theorem 1 (13) is given in Appendix A.

Consequently, the MMSEE (6) (best Bayesian bound) is the solution of the following constrained optimization problem: $\forall \mathcal{C}_i i = 1,\ldots,4$

$$\begin{cases} \min_v \int_{\Theta}\int_{\Omega} v^2(\mathbf{y},\theta)p(\mathbf{y},\theta)d\mathbf{y}d\theta \\ \text{subject to } v(\mathbf{y},\theta) \in \mathcal{C}_i. \end{cases} \tag{14}$$

III. Weiss-Weinstein Family Unification

In the light of the previous analysis, it appears a natural manner to introduce Bayesian bound lower than the MMSEE. Indeed, if $\mathcal{P}_i$ is a subset of $\mathcal{C}_i$, the solution of

$$\begin{cases} \min_v \int_{\Theta}\int_{\Omega} v^2(\mathbf{y},\theta)p(\mathbf{y},\theta)d\mathbf{y}d\theta \\ \text{subject to } v(\mathbf{y},\theta) \in \mathcal{P}_i \end{cases} \tag{15}$$

will be also a lower bound of the MMSEE. In this paper, we will first show that an appropriate choice of $\mathcal{P}_i$ leads to the Bayesian bounds of the Weiss-Weinstein family. Second, we will show how this approach can be used in order to build new minimal

$$\mathcal{C}_2 = \left\{ v(\mathbf{y},\theta) \in \mathcal{L}_p^2 \;\middle/\; \forall f \in \mathcal{F}, \int_{\Theta}\int_{\Omega} v(\mathbf{y},\theta)\frac{\partial f(\mathbf{y},\theta)}{\partial\theta} d\mathbf{y}d\theta = \int_{\Theta}\int_{\Omega} f(\mathbf{y},\theta)d\mathbf{y}d\theta \right\} \tag{10}$$

$$\mathcal{C}_3 = \left\{ v(\mathbf{y},\theta) \in \mathcal{L}_p^2 \;\middle/\; \begin{array}{c} \forall f \in \mathcal{F}, \text{ and } \int_{\Theta}\int_{\Omega} f(\mathbf{y},\theta)d\mathbf{y}d\theta = 1, \text{ and } \forall h \text{ such that } \theta + h \in \Theta \\ \int_{\Theta}\int_{\Omega} v(\mathbf{y},\theta)(f(\mathbf{y},\theta+h) - f(\mathbf{y},\theta))d\mathbf{y}d\theta = h \end{array} \right\} \tag{11}$$

$$\mathcal{C}_4 = \left\{ v(\mathbf{y},\theta) \in \mathcal{L}_p^2 \;\middle/\; \begin{array}{c} \forall f \in \mathcal{F}, \text{ and } \int_{\Theta}\int_{\Omega} f(\mathbf{y},\theta)d\mathbf{y}d\theta = 1, \text{ and } \forall h \text{ such that } \theta \pm h \in \Theta, \text{ and } \forall s \in [0,1] \\ \int_{\Theta}\int_{\Omega} v(\mathbf{y},\theta)[L^s(\mathbf{y},\theta+h,\theta) - L^{1-s}(\mathbf{y},\theta-h,\theta)]f(\mathbf{y},\theta)d\mathbf{y}d\theta \\ = h\int_{\Theta}\int_{\Omega} L^{1-s}(\mathbf{y},\theta-h,\theta)f(\mathbf{y},\theta)d\theta d\mathbf{y} \end{array} \right\}. \tag{12}$$

bounds, particularly, by solving the following constrained optimization problem:

$$\begin{cases} \min_v \int\limits_{\Theta}\int\limits_{\Omega} v^2(\mathbf{y},\theta)p(\mathbf{y},\theta)d\mathbf{y}d\theta \\ \text{subject to } v(\mathbf{y},\theta) \in \mathcal{P}_i \cap \mathcal{P}_j \quad i \neq j. \end{cases} \tag{16}$$

In this section, we restrict $\mathcal{C}_2$,$\mathcal{C}_3$, and $\mathcal{C}_4$ in order to obtain a general framework to create minimal bounds. Then, by way of a constrained optimization problem for which we give an explicit solution we unify the bounds of the Weiss-Weinstein family.

A. A General Class of Lower Bounds Based on $\mathcal{C}_2$,$\mathcal{C}_3$, and $\mathcal{C}_4$

Thanks to Theorem 1, we have proposed four equivalent sets of functions $v(\mathbf{y},\theta)$ leading to the MMSEE. Note that this equivalence holds for (17)-(19), shown at the bottom of the page.

Consequently, if we take a finite set of functions $f(\mathbf{y},\theta)$, a finite set of values h, and a finite set of values s, we will find bounds lower than the best Bayesian bounds and consequently a general class of minimal bounds on the MSE.

In this way, $\mathcal{C}_2$,$\mathcal{C}_3$, and $\mathcal{C}_4$ are restricted, respectively, as follows in (20)-(22), shown at the bottom of the page, with $L(\mathbf{y},\eta,\theta) = (f(\mathbf{y},\eta)/f(\mathbf{y},\theta))$.

$\mathcal{P}_2$, $\mathcal{P}_3$, and $\mathcal{P}_4$ define a set of finite constraints, and (15) becomes a classical linear constrained optimization problem

$$\begin{cases} \min_v \int\limits_{\Theta}\int\limits_{\Omega} v^2(\mathbf{y},\theta)p(\mathbf{y},\theta)\, d\mathbf{y}d\theta \\ \text{subject to } \int\limits_{\Theta}\int\limits_{\Omega} v(\mathbf{y},\theta)\tilde{g}_k(\mathbf{y},\theta)\, d\mathbf{y}d\theta = c_k \quad k = 1 \dots K \end{cases} \tag{23}$$

where $g_k(\mathbf{y},\theta)$, and c_k are the functions and the scalars involved in $\mathcal{P}_2$, $\mathcal{P}_3$, and $\mathcal{P}_4$.

For $\mathcal{P}_2$

$$\begin{aligned} \tilde{g}_k(\mathbf{y},\theta) &= \frac{\partial f_k(\mathbf{y},\theta)}{\partial\theta} \\ c_k &= \int\limits_{\Theta}\int\limits_{\Omega} f_k(\mathbf{y},\theta)\, d\mathbf{y}d\theta \quad \text{and} \\ K &= r. \end{aligned} \tag{24}$$

For $\mathcal{P}_3$

$$\begin{aligned} \tilde{g}_k(\mathbf{y},\theta) &= f(\mathbf{y},\theta+h_k) - f(\mathbf{y},\theta) \\ c_k &= h_k \text{ and } K = r. \end{aligned} \tag{25}$$

For $\mathcal{P}_4$, see (26) at the bottom of the page.

$$\forall f(\mathbf{y},\theta) \in \mathcal{F} \text{ in the subset } \mathcal{C}_2 \tag{17}$$

$$\begin{cases} \forall f(\mathbf{y},\theta) \in \mathcal{F} \text{ such that } \int\limits_{\Theta}\int\limits_{\Omega} f(\mathbf{y},\theta)d\mathbf{y}d\theta = 1 \\ \forall h \text{ such that } \theta + h \in \Theta. \end{cases} \text{ in the subset} \mathcal{C}_3 \tag{18}$$

$$\begin{cases} \forall f(\mathbf{y},\theta) \in \mathcal{F} \text{ such that } \int\limits_{\Theta}\int\limits_{\Omega} f(\mathbf{y},\theta)d\mathbf{y}d\theta = 1 \\ \forall h \text{ such that } \theta \pm h \in \Theta, \\ \forall s \in [0,1]. \end{cases} \text{ in the subset } \mathcal{C}_4. \tag{19}$$

$$\mathcal{P}_2 : \begin{cases} \text{for a finite set of functions } f_i(\mathbf{y},\theta) \in \mathcal{F}, \quad i = 1 \cdots r \\ \int\limits_{\Theta}\int\limits_{\Omega} v(\mathbf{y},\theta)\frac{\partial f_i(\mathbf{y},\theta)}{\partial\theta}d\mathbf{y}d\theta = \int\limits_{\Theta}\int\limits_{\Omega} f_i(\mathbf{y},\theta)\, d\mathbf{y}d\theta. \end{cases} \tag{20}$$

$$\mathcal{P}_3 : \begin{cases} \text{for a particular function } f(\mathbf{y},\theta) \in \mathcal{F} \text{ such that } \int\limits_{\Theta}\int\limits_{\Omega} f(\mathbf{y},\theta)\, d\mathbf{y}d\theta = 1 \\ \text{for a finite set of value } h_i \text{ such that } \theta + h_i \in \Theta, \quad i = 1 \cdots r \\ \int\limits_{\Theta}\int\limits_{\Omega} v(\mathbf{y},\theta)(f(\mathbf{y},\theta+h_i) - f(\mathbf{y},\theta))\, d\mathbf{y}d\theta = h_i. \end{cases} \tag{21}$$

$$\mathcal{P}_4 : \begin{cases} \text{for a particular function } f(\mathbf{y},\theta) \in \mathcal{F} \text{ such that } \int\limits_{\Theta}\int\limits_{\Omega} f(\mathbf{y},\theta)\, d\mathbf{y}d\theta = 1 \\ \text{for a finite set of value } h_i \text{ such that } \theta + h_i \in \Theta, \quad i = 1 \cdots r \\ \text{for a finite set of value } s_i \text{ such that } s_i \in [0,1], \quad i = 1 \cdots r \\ \int\limits_{\Theta}\int\limits_{\Omega} v(\mathbf{y},\theta)[L^{s_j}(\mathbf{y},\theta+h_i,\theta) - L^{1-s_j}(\mathbf{y},\theta-h_i,\theta)]f(\mathbf{y},\theta)\, d\mathbf{y}d\theta = h_i \int\limits_{\Theta}\int\limits_{\Omega} L^{1-s_j}(\mathbf{y},\theta-h_i,\theta)f(\mathbf{y},\theta)\, d\theta d\mathbf{y} \end{cases} \tag{22}$$

$$\begin{cases} \tilde{g}_k(\mathbf{y},\theta) = [L^{s_k}(\mathbf{y},\theta+h_k,\theta) - L^{1-s_k}(\mathbf{y},\theta-h_k,\theta)]f(\mathbf{y},\theta) \\ c_k = h_k \int\limits_{\Theta}\int\limits_{\Omega} L^{1-s_k}(\mathbf{y},\theta-h_k,\theta)f(\mathbf{y},\theta)\, d\theta d\mathbf{y} \\ \text{and } K = r. \end{cases} \tag{26}$$

Theorem 2 below gives the solution of (23). Note that this theorem has already been used in the case of a deterministic parameter in [17].

Theorem 2: Let $\mathbf{x} \in \mathbb{R}^N$ be a real vector and $p(\mathbf{x})$ and $q(\mathbf{x})$ be two functions of $\mathbb{R}^N \rightarrow \mathbb{R}$. Let

$$\langle p(\mathbf{x}), q(\mathbf{x}) \rangle = \int_{\mathbb{R}^N} p(\mathbf{x}) q(\mathbf{x}) d\mathbf{x} \tag{27}$$

be an inner product of these two functions and its associate norm $||p(\mathbf{x})||^2 = \langle p(\mathbf{x}), p(\mathbf{x}) \rangle$. Let $u(\mathbf{x})$ and $g_0(\mathbf{x}), \ldots, g_K(\mathbf{x})$ be a set of functions of $\mathbb{R}^N \rightarrow \mathbb{R}$, and let $c_0, c_1, \ldots, c_K$ and $K+1$ be real numbers. Then, the solution of the constrained optimization problem leading to the minimum of $||u(\mathbf{x})||^2$ under the following $K+1$ constraints

$$\langle u(\mathbf{x}), g_k(\mathbf{x}) \rangle = c_k \quad k = 0, \ldots, K \tag{28}$$

is given by

$$\begin{cases} \min_u ||u(\mathbf{x})||^2 = \mathbf{c}^T \mathbf{G}^{-1} \mathbf{c} \\ \text{subject to (28)} \end{cases} \tag{29}$$

with

$$\mathbf{c} = [c_0 \quad c_1 \quad \cdots \quad c_K]^T \tag{30}$$

and

$$\{\mathbf{G}\}_{m,n} = \langle g_m(\mathbf{x}), g_n(\mathbf{x}) \rangle. \tag{31}$$

The Proof of Theorem 2 (29) is given in Appendix B.

B. Application to the Weiss-Weinstein Family

Using (29), $\mathcal{P}_2, \mathcal{P}_3$, and $\mathcal{P}_4$, we have built a general framework to obtain Bayesian minimal bounds on the MSE. In this section, we apply this framework and we revisit the Bayesian bounds of the Weiss-Weinstein family. Let $\mathbf{x} = [\mathbf{y}^T \quad \theta]$ and $u(\mathbf{x}) = v(\mathbf{y}, \theta)\sqrt{p(\mathbf{y}, \theta)}$ (i.e., $\tilde{g}_k(\mathbf{y}, \theta) = \sqrt{p(\mathbf{y}, \theta)} g_k(\mathbf{y}, \theta)$). Note that Theorem 2 still holds for a set of complex observations $\bar{\mathbf{y}}$ by letting $\mathbf{y} = [\mathrm{Re}\{\bar{\mathbf{y}}^T\} \quad \mathrm{Im}\{\bar{\mathbf{y}}^T\}]^T$.

Moreover, due to the restriction at some particular values of $f(\mathbf{y}, \theta), h$, and s, it is still possible to add constraints with our *prior* on the MMSEE in order to achieve tighter bounds. Here we will use the natural constraints of a null bias in terms of the joint probability function; i.e., $\int_{\Theta} \int_{\Omega} v(\mathbf{y}, \theta) p(\mathbf{y}, \theta) d\mathbf{y} d\theta = 0$, where $p(\mathbf{y}, \theta)$ is the joint density of the problem (i.e., $g_0(\mathbf{y}, \theta) = \sqrt{p(\mathbf{y}, \theta)}$ and $c_0 = 0$).

a) Bayesian Cramér–Rao Bound: By using the set $\mathcal{P}_2$ with $K = 1$ and $f_1(\mathbf{y}, \theta) = p(\mathbf{y}, \theta)$ (consequently, $\int_{\Theta} \int_{\Omega} f_1(\mathbf{y}, \theta) d\mathbf{y} d\theta = 1$), we obtain the following set of constraints:

$$\begin{cases} \mathbf{c} = [0 \quad 1]^T \\ g_0(\mathbf{y}, \theta) = \sqrt{p(\mathbf{y}, \theta)} \\ g_1(\mathbf{y}, \theta) = \frac{1}{\sqrt{p(\mathbf{y}, \theta)}} \frac{\partial p(\mathbf{y}, \theta)}{\partial \theta}. \end{cases} \tag{32}$$

Matrix $\mathbf{G}$ involved in Theorem 2 is

$$\mathbf{G} = \begin{pmatrix} 1 & 0 \\ 0 & \int_{\Theta} \int_{\Omega} \left(\frac{\partial \ln p(\mathbf{y}, \theta)}{\partial \theta} \right)^2 p(\mathbf{y}, \theta) \, d\mathbf{y} d\theta \end{pmatrix} \tag{33}$$

since $\int_{\Theta} \int_{\Omega} (\partial p(\mathbf{y}, \theta) / \partial \theta) \, d\mathbf{y} d\theta = 0$.

Finally

$$\begin{aligned} \mathbf{c}^T \mathbf{G}^{-1} \mathbf{c} &= \left(\int_{\Theta} \int_{\Omega} \left(\frac{\partial \ln p(\mathbf{y}, \theta)}{\partial \theta} \right)^2 p(\mathbf{y}, \theta) \, d\mathbf{y} d\theta \right)^{-1} \\ &= \mathrm{BCRB} \end{aligned} \tag{34}$$

which is the Bayesian Cramér–Rao bound [20, pp. 72 and 84].

b) Bayesian Bhattacharyya Bound: By using the set $\mathcal{P}_2$ with $K = r$ and $f_k(\mathbf{y}, \theta) = (\partial^{k-1} p(\mathbf{y}, \theta) / \partial \theta^{k-1})$, we obtain the following set of constraints:

$$\begin{cases} \mathbf{c} = [0 \quad 1 \quad 0 \quad \cdots \quad 0]^T \\ g_0(\mathbf{y}, \theta) = \sqrt{p(\mathbf{y}, \theta)} \\ g_k(\mathbf{y}, \theta) = \frac{1}{\sqrt{p(\mathbf{y}, \theta)}} \frac{\partial^k p(\mathbf{y}, \theta)}{\partial \theta^k} \quad k = 1, \ldots, K. \end{cases} \tag{35}$$

We assume that the joint probability density function is such that $\lim_{\theta \rightarrow \pm\infty} (\partial^{k-1} p(\mathbf{y}, \theta) / \partial \theta^{k-1}) = 0 \quad$ for $k = 3, \ldots, K$. With this assumption and (9), we have

$$\begin{aligned} \mathbf{c}^T \mathbf{G}^{-1} \mathbf{c} &= \{\mathbf{B}^{-1}\}_{1,1} \\ &= \mathrm{Bhatt}B \end{aligned} \tag{36}$$

where

$$\{\mathbf{B}\}_{i,j} = \int_{\Theta} \int_{\Omega} \frac{1}{p(\mathbf{y}, \theta)} \frac{\partial^i p(\mathbf{y}, \theta)}{\partial \theta^i} \frac{\partial^j p(\mathbf{y}, \theta)}{\partial \theta^j} d\mathbf{y} d\theta \tag{37}$$

which is the Bayesian Bhattacharyya bound [20, p. 149].

c) Bobrovsky-MayerWolf-Zakaï Bound: By using the set $\mathcal{P}_2$ with $K = 1$ and $f_1(\mathbf{y}, \theta) = q(\mathbf{y}, \theta) p(\mathbf{y}, \theta)$, where $q(\mathbf{y}, \theta)$ is any function such that $f_1(\mathbf{y}, \theta)$ satisfies (9), we obtain the following set of constraints:

$$\begin{cases} \mathbf{c} = \left[0 \quad \int_{\Theta} \int_{\Omega} q(\mathbf{y}, \theta) p(\mathbf{y}, \theta) d\mathbf{y} d\theta \right]^T \\ g_0(\mathbf{y}, \theta) = \sqrt{p(\mathbf{y}, \theta)} \\ g_1(\mathbf{y}, \theta) = \frac{1}{\sqrt{p(\mathbf{y}, \theta)}} \frac{\partial [p(\mathbf{y}, \theta) q(\mathbf{y}, \theta)]}{\partial \theta}. \end{cases} \tag{38}$$

Due to (9), $\int_{\Theta} \int_{\Omega} (\partial q(\mathbf{y}, \theta) p(\mathbf{y}, \theta) / \partial \theta) \, d\mathbf{y} d\theta = 0$ and the matrix $\mathbf{G}$ involved in Theorem 2 is

$$\mathbf{G} = \begin{pmatrix} 1 & 0 \\ 0 & \int_{\Theta} \int_{\Omega} \frac{1}{\sqrt{p(\mathbf{y}, \theta)}} \frac{\partial [p(\mathbf{y}, \theta) q(\mathbf{y}, \theta)]}{\partial \theta} d\mathbf{y} d\theta \end{pmatrix} \tag{39}$$

Finally

$$\begin{aligned} \mathbf{c}^T \mathbf{G}^{-1} \mathbf{c} &= \frac{\left(\int_{\Theta} \int_{\Omega} q(\mathbf{y}, \theta) p(\mathbf{y}, \theta) d\theta d\mathbf{y} \right)^2}{\int_{\Theta} \int_{\Omega} \frac{1}{p(\mathbf{y}, \theta)} \left(\frac{\partial [p(\mathbf{y}, \theta) q(\mathbf{y}, \theta)]}{\partial \theta} \right)^2 d\mathbf{y} d\theta} \\ &= \mathrm{BMZB}(q(\mathbf{y}, \theta)). \end{aligned} \tag{40}$$

We recognize the Bobrovsky-MayerWolf-Zakaï bound [39], which is an extension of the Bayesian Cramér–Rao bound, since

$$\mathrm{BMZB}(1) = \mathrm{BCRB}. \tag{41}$$

d) Bobrovsky-Zakaï Bound: We choose here that the particular value of $f(\mathbf{y},\theta) = p(\mathbf{y},\theta)$, the joint density probability function of the problem. Consequently, $\int_{\Theta}\int_{\Omega} f(\mathbf{y},\theta)d\mathbf{y}d\theta = 1$.

By using the set $\mathcal{P}_3$ with $K = 1$, we obtain the following set of constraints:

$$\begin{cases} \mathbf{c} = [0 \quad h]^T \\ g_0(\mathbf{y},\theta) = \sqrt{p(\mathbf{y},\theta)} \\ g_1(\mathbf{y},\theta) = \frac{p(\mathbf{y},\theta+h)-p(\mathbf{y},\theta)}{\sqrt{p(\mathbf{y},\theta)}}. \end{cases} \tag{42}$$

Matrix $\mathbf{G}$ involved in Theorem 2 is

$$\mathbf{G} = \begin{pmatrix} 1 & 0 \\ 0 & \int_{\Theta}\int_{\Omega} \frac{(p(\mathbf{y},\theta+h)-p(\mathbf{y},\theta))^2}{p(\mathbf{y},\theta)} d\mathbf{y}d\theta \end{pmatrix}. \tag{43}$$

Finally

$$\begin{aligned} \mathbf{c}^T\mathbf{G}^{-1}\mathbf{c} &= \frac{h^2}{\int_{\Theta}\int_{\Omega} \frac{p^2(\mathbf{y},\theta+h)}{p(\mathbf{y},\theta)} d\mathbf{y}d\theta - 1} \\ &= \mathrm{BZB}(h). \end{aligned} \tag{44}$$

Since h is a parameter left to the user, the highest bound that can be obtained with (44) is given by

$$\mathrm{BZB} = \sup_h \mathrm{BZB}(h) = \sup_h \frac{h^2}{\int_{\Theta}\int_{\Omega} \frac{p^2(\mathbf{y},\theta+h)}{p(\mathbf{y},\theta)} d\mathbf{y}d\theta - 1} \tag{45}$$

which is the Bobrovsky-Zakaï bound [40].

e) Reuven-Messer Bound: We choose here that the particular value of $f(\mathbf{y},\theta) = p(\mathbf{y},\theta)$, the joint density probability function of the problem. Consequently, $\int_{\Theta}\int_{\Omega} f(\mathbf{y},\theta)d\mathbf{y}d\theta = 1$.

In order to obtain a bound tighter than the Bobrovsky-Zakaï bound (i.e., $\mathcal{P}_3 \rightarrow \mathcal{C}_3$), we use the set $\mathcal{P}_3$ with $K = r$. We then obtain the following set of constraints:

$$\begin{cases} \mathbf{c} = [0 \quad \mathbf{h}^T]^T \\ g_0(\mathbf{y},\theta) = \sqrt{p(\mathbf{y},\theta)} \\ g_k(\mathbf{y},\theta) = \frac{p(\mathbf{y},\theta+h_k)-p(\mathbf{y},\theta)}{\sqrt{p(\mathbf{y},\theta)}} \quad k = 1,\ldots,r \end{cases} \tag{46}$$

where $\mathbf{h} = [h_1 \quad \cdots \quad h_r]^T$.

Matrix $\mathbf{G}$ involved in Theorem 2 is

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{D} & \\ 0 & & & \end{pmatrix} \tag{47}$$

where $\mathbf{D}$ $(r \times r)$ is defined as shown in (48) shown at the bottom of the page.

Finally

$$\begin{aligned} \mathbf{c}^T\mathbf{G}^{-1}\mathbf{c} &= \mathbf{h}^T\mathbf{D}^{-1}\mathbf{h} \\ &= \mathrm{RMB}(\mathbf{h}). \end{aligned} \tag{49}$$

As for the Bobrovsky-Zakaï bound, since $\mathbf{h}$ is a parameter vector left to the user, the highest bound that can be obtained with (49) is given by

$$\mathrm{RMB} = \sup_{\mathbf{h}} \mathrm{RMB}(\mathbf{h}) = \sup_{\mathbf{h}} \mathbf{h}^T\mathbf{D}^{-1}\mathbf{h} \tag{50}$$

which is a particular case[1] of the Reuven-Messer bound [41].

f) Weiss-Weinstein Bound: We choose here that the particular value of $f(\mathbf{y},\theta) = p(\mathbf{y},\theta)$, the joint density probability function of the problem. Consequently, $\int_{\Theta}\int_{\Omega} f(\mathbf{y},\theta)d\mathbf{y}d\theta = 1$.

By using the set $\mathcal{P}_4$ with $K = r$, we obtain the following set of constraints: [see (51) at the bottom of the next page].

Let, [see (52)-(54) at the bottom of the next page].

The application of Theorem 2 leads to

$$\begin{aligned} \mathbf{c}^T\mathbf{G}^{-1}\mathbf{c} &= \boldsymbol{\xi}^T\mathbf{W}^{-1}\boldsymbol{\xi} \\ &= \mathrm{WWB}(\mathbf{h},\mathbf{s}) \end{aligned} \tag{55}$$

where

$$\begin{aligned} \{\mathbf{W}\}_{i,j} = \mathbb{E}_{\mathbf{y},\theta}[&(L^{s_i}(\mathbf{y},\theta+h_i,\theta) \\ &- L^{1-s_i}(\mathbf{y},\theta-h_i,\theta))(L^{s_j}(\mathbf{y},\theta+h_j,\theta) \\ &- L^{1-s_j}(\mathbf{y},\theta-h_j,\theta))]. \end{aligned} \tag{56}$$

As for the Bobrovsky-Zakaï bound and the Reuven-Messer bound, since $\mathbf{h}$ and $\mathbf{s}$ are parameter vectors left to the user, the highest bound that can be obtained with (55) is given by

$$\mathrm{WWB} = \sup_{\mathbf{h},\mathbf{s}} \mathrm{WWB}(\mathbf{h},\mathbf{s}) = \sup_{\mathbf{h},\mathbf{s}} \boldsymbol{\xi}^T\mathbf{W}^{-1}\boldsymbol{\xi}. \tag{57}$$

We recognize the Weiss-Weinstein bound [42].

IV. New Minimal Bounds

The framework proposed in the last section allows us to rederive all the bounds of the Weiss-Weinstein family by way of a constrained optimization problem. But this framework is also

[1]In 1997, Reuven and Messer proposed a hybrid minimal bound based on the Barankin bound for both random and nonrandom vector of parameters. Here, only the random case is considered.

$$\begin{aligned} \{\mathbf{D}\}_{i,,j} &= \int_{\Theta}\int_{\Omega} \frac{(p(\mathbf{y},\theta+h_i)-p(\mathbf{y},\theta))(p(\mathbf{y},\theta+h_j)-p(\mathbf{y},\theta))}{p(\mathbf{y},\theta)} d\mathbf{y}d\theta \\ &= \int_{\Theta}\int_{\Omega} \frac{p(\mathbf{y},\theta+h_i)p(\mathbf{y},\theta+h_j)}{p(\mathbf{y},\theta)} d\mathbf{y}d\theta - 1. \end{aligned} \tag{48}$$

useful for deriving new lower bounds. In this section, we propose two lower bounds.

A. Some Global Classes of Bhattacharyya Bounds

In [39], Bobrovsky, Mayer-Wolf, and Zakaï propose an extension of the Bayesian Cramér–Rao bound given by (40). The advantage of this bound is the degree of freedom given by $q(\mathbf{y},\theta)$. Indeed, the authors give some examples for which use of a properly chosen function $q(\mathbf{y},\theta)$ leads to useful bounds. Moreover, when $p(\mathbf{y},\theta)$ does not satisfy the regularity assumption given in [20] (e.g., for uniform random variables), a properly chosen $q(\mathbf{y},\theta)$ can solve the problem. Here we obtain an extension of this bound and of the Bayesian Bhattacharyya bound in a straightforward manner by mixing the constraints of the Bobrovsky-MayerWolf-Zakaï bound and the constraints of the Bayesian Bhattacharyya bound.

By using the set $\mathcal{P}_2$ with $K = r$ and $f_k(\mathbf{y},\theta) = (\partial^{k-1}[q(\mathbf{y},\theta)p(\mathbf{y},\theta)]/\partial\theta^{k-1})$, where $q(\mathbf{y},\theta)$ is any function such that $f_k(\mathbf{y},\theta)$ satisfies (9), we obtain the following set of constraints:

$$\begin{cases} \mathbf{c} = \left[0 \quad \int_{\Theta}\int_{\Omega} q(\mathbf{y},\theta)p(\mathbf{y},\theta)d\mathbf{y}d\theta \quad 0 \quad \cdots \quad 0\right]^T \\ g_0(\mathbf{y},\theta) = \sqrt{p(\mathbf{y},\theta)} \\ g_k(\mathbf{y},\theta) = \frac{1}{\sqrt{p(\mathbf{y},\theta)}}\frac{\partial^k[q(\mathbf{y},\theta)p(\mathbf{y},\theta)]}{\partial\theta^k} \quad k = 1,\ldots,K. \end{cases} \tag{58}$$

We assume that the functions $q(\mathbf{y},\theta)$ and $p(\mathbf{y},\theta)$ are such that $\lim_{\theta\to\pm\infty}(\partial^{k-1}q(\mathbf{y},\theta)p(\mathbf{y},\theta)/\partial\theta^{k-1}) = 0$ for $k = 3,\ldots,K$. With this assumption and (9), we have

$$\mathbf{c}^T\mathbf{G}^{-1}\mathbf{c} = \frac{\left(\int_{\Theta}\int_{\Omega} q(\mathbf{y},\theta)p(\mathbf{y},\theta)d\mathbf{y}d\theta\right)^2}{\{\bar{\mathbf{B}}^{-1}\}_{1,1}} \tag{59}$$

where [see (60) shown at the bottom of the page].

B. The Bayesian Abel Bound

In this section, we propose a new minimal bound on the MSE based on our framework and on the Abel works on deterministic bounds [29], [30]. In the deterministic parameter context, the Cramér–Rao bound and the Bhattacharyya bound account for the *small estimation error* (near the true value of the parameters). The Chapman-Robbins bound and the Barankin bound account for the *large estimation error* generally due to the appearance of outliers which creates the performance breakdown phenomena. In [29] and [30], Abel combined the two kinds of bounds in order to obtain a bound that accounts for both local and large errors. The obtained deterministic Abel bound leads to a generalization of the Cramér–Rao, the Bhattacharyya, the Chapman-Robbins, and the Barankin bounds. As the deterministic bounds, the Bayesian Cramér–Rao bound and the Bayesian Bhattacharyya bound are *small error* bounds, as compared to the Bobrovsky-Zakaï bound and the Reuven-Messer bound which are *large error* bounds. The purpose here is to apply the idea of Abel in the Bayesian context, i.e., to derive a bound that combines the Bayesian small and large error bounds. This application will be accomplished by way of the constrained optimization problem introduced in the last section. Our Bayesian version of the Abel bound is derived by mixing the constraints of the Reuven-Messer bound and the Bayesian Bhattacharyya bound and, thus, represents a generalization of these bounds. Consequently, we are solving the following constrained optimization problem

$$\begin{cases} \min_v \int_{\Theta}\int_{\Omega} v^2(\mathbf{y},\theta)p(\mathbf{y},\theta)d\mathbf{y}d\theta \\ \text{subject to } v(\mathbf{y},\theta) \in \mathcal{P}_2 \cap \mathcal{P}_3. \end{cases} \tag{61}$$

By combining the Bayesian Bhattacharyya constraints (35) and the Reuven-Messer constraints (46), i.e., by concatenating both vectors $\mathbf{g} = [g_0(\mathbf{y},\theta), g_1(\mathbf{y},\theta),\ldots,g_K(\mathbf{y},\theta)]^T$ and $\mathbf{c}$ from the Bayesian Bhattacharyya bound of order m and from

$$\begin{cases} \mathbf{c} = [0 \quad h_1\mathbb{E}_{\mathbf{y},\theta}[L^{1-s_1}(\mathbf{y},\theta-h_1,\theta)] \quad \cdots \quad h_k\mathbb{E}_{\mathbf{y},\theta}[L^{1-s_k}(\mathbf{y},\theta-h_k,\theta)]]^T \\ g_0(\mathbf{y},\theta) = \sqrt{p(\mathbf{y},\theta)} \\ g_k(\mathbf{y},\theta) = \sqrt{p(\mathbf{y},\theta)}(L^{s_k}(\mathbf{y},\theta+h_k,\theta) - L^{1-s_k}(\mathbf{y},\theta-h_k,\theta)) \quad k = 1,\ldots,r. \end{cases} \tag{51}$$

$$\boldsymbol{\xi} = [h_1\mathbb{E}_{\mathbf{y},\theta}[L^{1-s_1}(\mathbf{y},\theta-h_1,\theta)] \quad \cdots \quad h_r\mathbb{E}_{\mathbf{y},\theta}[L^{1-s_r}(\mathbf{y},\theta-h_r,\theta)]]^T \tag{52}$$

$$\mathbf{h} = [h_1 \quad \cdots \quad h_r]^T \tag{53}$$

$$\mathbf{s} = [s_1 \quad \cdots \quad s_r]^T. \tag{54}$$

$$\{\bar{\mathbf{B}}\}_{i,j} = \int_{\Theta}\int_{\Omega} \frac{1}{p(\mathbf{y},\theta)}\frac{\partial^i[q(\mathbf{y},\theta)p(\mathbf{y},\theta)]}{\partial\theta^i}\frac{\partial^j[q(\mathbf{y},\theta)p(\mathbf{y},\theta)]}{\partial\theta^j}d\mathbf{y}d\theta. \tag{60}$$

the Reuven-Messer bound of order r, we obtain the following new set of $K = m + r + 1$ constraints [2]:

$$\mathbf{g} = \frac{1}{\sqrt{p(\mathbf{y},\theta)}} \begin{bmatrix} p(\mathbf{y},\theta) \\ \frac{\partial p(\mathbf{y},\theta)}{\partial \theta} \\ \vdots \\ \frac{\partial^m p(\mathbf{y},\theta)}{\partial \theta^m} \\ - - - - - - \\ p(\mathbf{y},\theta+h_1) - p(\mathbf{y},\theta) \\ \vdots \\ p(\mathbf{y},\theta+h_r) - p(\mathbf{y},\theta) \end{bmatrix}$$

and

$$\mathbf{c} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ - - \\ h_1 \\ \vdots \\ h_r \end{bmatrix}. \tag{62}$$

The calculus are detailed in Appendix C, and the Theorem 2 leads to

$$\mathbf{c}^T\mathbf{G}^{-1}\mathbf{c} = \mathrm{BAB}_{m,r}(\mathbf{h}) = \boldsymbol{\alpha}^T\mathbf{B}^{-1}\boldsymbol{\alpha} + \mathbf{u}^T\mathbf{J}^{-1}\mathbf{u} \tag{63}$$

with

$$\begin{cases} \mathbf{u} = \boldsymbol{\Gamma}\mathbf{B}^{-1}\boldsymbol{\alpha} - \mathbf{h}, r \times 1 \\ \mathbf{J} = \mathbf{D} - \boldsymbol{\Gamma}\mathbf{B}^{-1}\boldsymbol{\Gamma}^T, r \times r \\ \boldsymbol{\alpha} = [1 \quad 0 \quad \cdots \quad 0]^T, m \times 1 \\ \mathbf{h} = [h_1 \quad h_2 \quad \cdots \quad h_r]^T, r \times 1 \\ \{\mathbf{D}\}_{i,,j} = \int\limits_{\Theta}\int\limits_{\Omega} \frac{p(\mathbf{y},\theta+h_i)p(\mathbf{y},\theta+h_j)}{p(\mathbf{y},\theta)} d\mathbf{y}d\theta - 1, r \times r \\ \{\mathbf{B}\}_{i,j} = \int\limits_{\Theta}\int\limits_{\Omega} \frac{1}{p(\mathbf{y},\theta)} \frac{\partial^i p(\mathbf{y},\theta)}{\partial \theta^i} \frac{\partial^j p(\mathbf{y},\theta)}{\partial \theta^j} d\mathbf{y}d\theta, m \times m \\ \{\boldsymbol{\Gamma}\}_{i,j} = \int\limits_{\Theta}\int\limits_{\Omega} \frac{p(\mathbf{y},\theta+h_i)}{p(\mathbf{y},\theta)} \frac{\partial^j p(\mathbf{y},\theta)}{\partial \theta^j} d\mathbf{y}d\theta r \times m. \end{cases} \tag{64}$$

Let us note that the first term on right-hand side (RHS) of (63) is equal to $\mathrm{BAB}_{m,0}$, which is the Bayesian Bhattacharyya bound of order m, and that $\mathrm{BAB}_{0,r}(\mathbf{h})$ is the Reuven-Messer bound of order r. We have previously shown that problem (8) leads to the MMSEE (the best Bayesian bound). Here, from the increase of constraints, it follows that the Bayesian Abel bound is (for r and m fixed) a better approximation of the best Bayesian bound than the Bayesian Bhattacharyya bound of order m and the Reuven-Messer bound of order r.

The Bayesian Abel bound as the Reuven-Messer bound depends on r free parameters $h_1, \ldots, h_r$. Then, a maximization over these parameters is desired to obtain the highest bound. Therefore, the best Bayesian Abel bound is given by

$$\mathrm{BAB}_{m,r} = \sup_{\mathbf{h}}(\boldsymbol{\alpha}^T\mathbf{B}^{-1}\boldsymbol{\alpha} + \mathbf{u}^T\mathbf{J}^{-1}\mathbf{u}). \tag{65}$$

[2] The first constraint of the two bounds is the same.

This multidimensional optimization brings with it a huge computational cost. A possible alternative is given by noting that the Bayesian Cramér–Rao bound is a particular case of the Bayesian Bhattacharyya bound (single derivative) and that the Bobrovsky-Zakaï bound is a particular case of the Reuven-Messer bound (single test point). Therefore, finding a tractable form of the Bayesian Abel bound in the case where $m = 1$ and $r = 1$ could be interesting, since the obtained bound will be tighter than both the Bayesian Cramér–Rao bound and the Bobrovsky-Zakaï bound with a low computational cost. In this case, (65) becomes straightforwardly

$$\mathrm{BAB}_{1,1} = \sup_{h} \frac{\mathrm{BCRB}^{-1} + \mathrm{BZB}^{-1}(h) - 2\phi(h)}{\mathrm{BCRB}^{-1}\mathrm{BZB}^{-1}(h) - \phi^2(h)} \tag{66}$$

where BCRB is the Bayesian Cramér–Rao bound, BZB is the Bobrovsky-Zakaï bound, and

$$\phi(h) = \frac{1}{h}\int\limits_{\Theta}\int\limits_{\Omega} \frac{\partial \ln p(\mathbf{y},\theta)}{\partial \theta} p(\mathbf{y},\theta+h) d\mathbf{y}d\theta. \tag{67}$$

Equation (66) is interesting, since if the Bayesian Cramér–Rao bound and the Bobrovsky-Zakaï bound are available for a given problem, the evaluation of the $\mathrm{BAB}_{1,1}$ requires only the computation of $\phi(h)$.

V. Bayesian Bounds for Signal Processing Problems

In this section, we illustrate our previous analysis through a spectral analysis problem. First, we propose several closed-form expressions for the different bounds of the Weiss-Weinstein family (including the proposed Bayesian Abel bound) for a general Gaussian observation model with parameterized mean widely used in the signal processing literature (see, e.g., [68, p. 35]). Then, we apply these results to the spectral analysis problem. Finally, we give simulation results that compare the different bounds and show the superiority of the Weiss-Weinstein bound.

A. Gaussian Observation Model With Parameterized Mean

We consider the following general observation model:

$$\mathbf{y} = \mathbf{m}(\theta) + \mathbf{n} \tag{68}$$

where $\mathbf{y}$ is the complex observation vector $(N \times 1)$, θ is a real unknown parameter, $\mathbf{m}$ is a complex deterministic vector $(N \times 1)$ depending (nonlinearly) on θ, and $\mathbf{n}$ is the complex vector $(N \times 1)$ of the noise. The noise is assumed to be circular, Gaussian, with zero mean and with covariance matrix $\sigma^2\mathbf{I}_N$. The parameter of interest θ is assumed to have a Gaussian *a priori* probability density function with mean μ and variance σ_θ^2

$$p(\theta) = \frac{1}{\sqrt{2\pi}\sigma_\theta} e^{-\frac{1}{2\sigma_\theta^2}(\theta-\mu)^2}. \tag{69}$$

For this model, the likelihood of the observations is given by

$$p(\mathbf{y}|\theta) = \frac{1}{(\pi\sigma^2)^N} e^{-\frac{1}{\sigma^2}(\mathbf{y}-\mathbf{m}(\theta))^H(\mathbf{y}-\mathbf{m}(\theta))}. \tag{70}$$

To the best of our knowledge, only the Cramér–Rao bound expression is known in this case (see [68]).

The Bayesian Bhattacharyya bound requires the calculation of several derivatives of the joint probability function in order to be significantly tighter than the Cramér–Rao bound, which is generally difficult (see [69, Ch. 4] for an example for which the Bhattacharyya bound of order 2 requires much algebraic effort to finally be equal to the Cramér–Rao bound). Consequently, we will not use this bound here.

The details are given in Appendix D.

1) Bayesian Cramér–Rao Bound:

$$\mathrm{BCRB} = \frac{\sigma_\theta^2}{\frac{2\sigma_\theta^2}{\sigma^2}\mathbb{E}_\theta\left[\left\|\frac{\partial \mathbf{m}(\theta)}{\partial\theta}\right\|^2\right]+1}. \tag{71}$$

2) Bobrovsky-Zakaï Bound:

$$\mathrm{BZB} = \sup_h \frac{h^2}{\int_\Theta \frac{p^2(\theta+h)}{p(\theta)} e^{\frac{2}{\sigma^2}\|\mathbf{m}(\theta+h)-\mathbf{m}(\theta)\|^2} d\theta - 1}. \tag{72}$$

3) Bayesian Abel Bound: $\mathrm{BAB}_{1,1}$ is given by (66):

$$\mathrm{BAB}_{1,1} = \sup_h \frac{\mathrm{BCRB}^{-1} + \mathrm{BZB}^{-1}(h) - 2\phi(h)}{\mathrm{BCRB}^{-1}\mathrm{BZB}^{-1}(h) - \phi^2(h)} \tag{73}$$

where

$$\begin{cases} \mathrm{BCRB} = \frac{\sigma_\theta^2}{\frac{2\sigma_\theta^2}{\sigma^2}\mathbb{E}_\theta\left[\left\|\frac{\partial \mathbf{m}(\theta)}{\partial\theta}\right\|^2\right]+1} \\ \mathrm{BZB}(h) = \frac{h^2}{\int_\Theta \frac{p^2(\theta+h)}{p(\theta)} e^{\frac{2}{\sigma^2}\|\mathbf{m}(\theta+h)-\mathbf{m}(\theta)\|^2} d\theta - 1}, \end{cases} \tag{74}$$

and

$$\phi(h) = \frac{1}{\sigma_\theta^2} + \frac{2}{h\sigma^2}\mathbb{E}_{\theta+h}\left[\mathrm{Re}\left\{\frac{\partial \mathbf{m}^H(\theta)}{\partial\theta}(\mathbf{m}(\theta+h)-\mathbf{m}(\theta))\right\}\right]. \tag{75}$$

4) Weiss-Weinstein Bound: We now consider the Weiss-Weinstein bound with one test point, which can be simplified as follows (see [42, eq. (6)]):

$$\mathrm{WWB} = \sup_{h,s} \frac{h^2 e^{2\eta(s,h)}}{e^{\eta(2s,h)} + e^{\eta(2-2s,-h)} - 2e^{\eta(s,2h)}} \tag{76}$$

where the key point to evaluate this bound is $\eta(\alpha,\beta)$, which is the semi-invariant moment generating function [70], defined as follows:

$$\eta(\alpha,\beta) = \ln \int_\Theta \int_\Omega \frac{p^\alpha(\mathbf{y},\theta+\beta)}{p^{\alpha-1}(\mathbf{y},\theta)} d\mathbf{y} d\theta. \tag{77}$$

This function is given by (78), as shown at the bottom of the page.

B. Spectral Analysis Problem

We now consider the following observation model involved in spectral analysis:

$$y_k = a e^{j2\pi k\theta} + n_k, \quad k = 0,\ldots,N-1 \tag{79}$$

where y_k is the k^{th} complex observation. The observations are assumed to be independent. a is the amplitude of the single cisoïde of frequency θ. $\{n_k\}$ is a sequence of random variables assumed complex, circular, i.i.d, Gaussian, with zero mean and variance σ^2. Consequently the SNR is given by $\mathrm{SNR} = (a^2/\sigma^2)$. The parameter of interest is the frequency $\theta \in \Theta = (-(1/2),(1/2)]$ which is a Gaussian random variable with mean μ and variance σ_θ^2 (69).

This model is a particular case of the model (68), where

$$\mathbf{m}(\theta) = a\mathbf{s}(\theta) \tag{80}$$

with

$$\mathbf{s}(\theta) = [1 \quad e^{j2\pi\theta} \quad \cdots \quad e^{j2\pi(N-1)\theta}]^T. \tag{81}$$

Let $\mathbf{y} = [y_0 \quad \cdots \quad y_{N-1}]^T$. The likelihood of the observation is given by

$$\begin{aligned} p(\mathbf{y}|\theta) &= \prod_{k=0}^{N-1} p(y_k|\theta) \\ &= \frac{1}{(\pi\sigma^2)^N} e^{-\frac{1}{\sigma^2}\left(\|\mathbf{y}\|^2 - 2a\mathrm{Re}\left\{\sum_{k=0}^{N-1} y_k^* e^{j2\pi k\theta}\right\} + Na^2\right)}. \end{aligned} \tag{82}$$

Note that, if θ is assumed to be deterministic and in a digital communications context, some closed-form expressions of deterministic bounds can be found in [56].

The details of the calculus for the Weiss-Weinstein family are given in Appendix E.

1) Cramér–Rao Bound:

$$\mathrm{BCRB} = \frac{\sigma_\theta^2}{\mathrm{SNR}\frac{4\pi^2\sigma_\theta^2}{3}N(2N-1)(N-1)+1}. \tag{83}$$

2) Bobrovsky-Zakaï Bound:

$$\mathrm{BZB} = \sup_h \frac{h^2}{e^{4\mathrm{SNR}\left(N - \sin^2(\pi h N) - \frac{1}{2}\frac{\sin(2\pi h N)}{\tan(\pi h)}\right) + \frac{h^2}{\sigma_\theta^2}} - 1}. \tag{84}$$

3) Bayesian Abel Bound: The $\mathrm{BAB}_{1,1}$ is given by (66)

$$\mathrm{BAB}_{1,1} = \sup_h \frac{\mathrm{BCRB}^{-1} + \mathrm{BZB}^{-1}(h) - 2\phi(h)}{\mathrm{BCRB}^{-1}\mathrm{BZB}^{-1}(h) - \phi^2(h)} \tag{85}$$

$$\eta(\alpha,\beta) = \ln \frac{1}{\sqrt{2\pi}\sigma_\theta} \int_\Theta e^{\frac{\alpha(\alpha-1)}{\sigma^2}\|\mathbf{m}(\theta+\beta)-\mathbf{m}(\theta)\|^2 - \frac{1}{2\sigma_\theta^2}(\theta-(\sqrt{\alpha(\alpha-1)}-\alpha)h-\mu)(\theta+(\sqrt{\alpha(\alpha-1)}+\alpha)h-\mu)} d\theta. \tag{78}$$

where

$$\begin{cases} \mathrm{BCRB} = \frac{1}{\mathrm{SNR}} \frac{\sigma_\theta^2}{\frac{4\pi^2\sigma_\theta^2}{3} N(2N-1)(N-1)+1} \\ \mathrm{BZB}(h) = \frac{h^2}{e^{4\mathrm{SNR}\left(N-\sin^2(\pi hN)-\frac{1}{2}\frac{\sin(2\pi hN)}{\tan(\pi h)}\right)+\frac{h^2}{\sigma_\theta^2}}-1} \end{cases} \tag{86}$$

and

$$\phi(h) = \frac{1}{\sigma_\theta^2} + \frac{2\pi\mathrm{SNR}}{h}\left(N\frac{\cos(2\pi hN)}{\tan(\pi h)} - \sin(2\pi hN)\left(\frac{1}{2\sin(\pi h)} + N\right)\right). \tag{87}$$

4) Weiss-Weinstein Bound: The Weiss-Weinstein bound is given by

$$\mathrm{WWB} = \sup_{h,s} \frac{h^2 e^{2\eta(s,h)}}{e^{\eta(2s,h)} + e^{\eta(2-2s,-h)} - 2e^{\eta(s,2h)}} \tag{88}$$

where $\eta(\alpha,\beta)$ is given by

$$\eta(\alpha,\beta) = \alpha(\alpha-1)\left(2\mathrm{SNR}\left(N - \sin^2(\pi\beta N) - \frac{1}{2}\frac{\sin(2\pi\beta N)}{\tan(\pi\beta)}\right) - \frac{\beta^2}{2\sigma_\theta^2}\right). \tag{89}$$

The Weiss-Weinstein bound needs to be optimized over two continuous parameters, which creates significant computational cost. Here, two methods for reducing the computational cost are presented.

- As previously stated, h is chosen on the parameter support which is approximated by $[-3\sigma_\theta, 3\sigma_\theta]$. This support can be reduced to $[0, 3\sigma_\theta]$, since the function is even with respect to h. Note that this remark holds for the Bayesian Abel bound and the Bobrovsky-Zakaï bound.
- As proposed by Weiss and Weinstein in [42], it is sometimes a good choice to set $s = 1/2$. This approximation is intuitively justified by the fact that the Weiss-Weinstein bound tends to the Bobrovsky-Zakaï bound when s tends to zero or one. Unfortunately, no sound proof that this result is true in general is available in the literature. If we set $s = 1/2, \eta(\alpha,\beta)$ is modified as follows: [see (90) at the bottom of the page] and the modified Weiss-Weinstein bound becomes

$$\overline{\mathrm{WWB}} = \sup_h \frac{h^2}{2} \frac{e^{-\frac{1}{2}\left(2\mathrm{SNR}\left(N-\sin^2(\pi hN)-\frac{1}{2}\frac{\sin(2\pi hN)}{\tan(\pi h)}\right)-\frac{h^2}{2\sigma_\theta^2}\right)}}{1-e^{-\frac{1}{2}\left(\mathrm{SNR}\left(N-\sin^2(2\pi hN)-\frac{1}{2}\frac{\sin(4\pi hN)}{\tan(2\pi h)}\right)-\frac{h^2}{\sigma_\theta^2}\right)}}. \tag{91}$$

The resulting bound has approximaely the same computational cost as the BZB and the BAB.

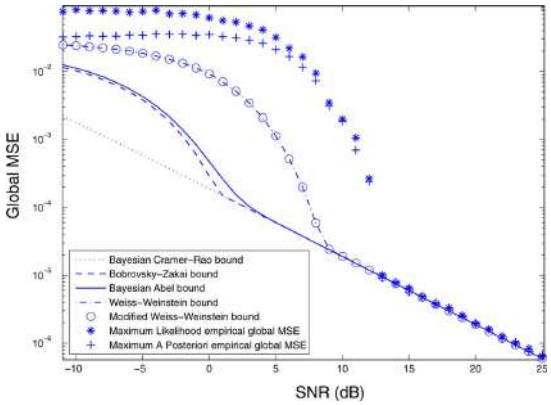


Fig. 1. Comparison of the global MSE of the MLE and of the MAP estimator, the Cramér–Rao bound, the Bobrovsky-Zakaï bound, the Bayesian Abel bound, and the Weiss-Weinstein bound with optimization over s and $s = 1/2$. $N = 15$ observations. $\sigma_\theta^2 = 1/36$ rad^2.

C. Simulations

In order to illustrate our results on the different bounds, we present here a simulation result for the spectral analysis problem.

We consider a scenario with $N = 15$ observations and, without loss of generality, $a = 1$. The estimator will be the maximum likelihood estimator (MLE) given for this model by

$$\hat{\theta}_{\mathrm{ML}} = \arg\min_\theta \left[\|\mathbf{y}\|^2 + Na^2 - 2a\mathrm{Re}\left\{\sum_{k=0}^{N-1} y_k^* e^{j2\pi k\theta}\right\}\right]. \tag{92}$$

We also use the maximum *a posteriori* (MAP) estimator given by

$$\hat{\theta}_{\mathrm{MAP}} = \arg\min_\theta \left[\frac{1}{\sigma^2}\left(\|\mathbf{y}\|^2 + Na^2 - 2a\mathrm{Re}\left\{\sum_{k=0}^{N-1} y_k^* e^{j2\pi k\theta}\right\}\right) + \frac{\theta^2}{2\sigma_\theta^2}\right]. \tag{93}$$

The global MSE will be computed by using (3) and 1000 Monte Carlo runs. For the *a priori* probability density function of the parameter of interest, we choose $\mu = 0$ and $\sigma_\theta^2 = (1/36)$ rad^2.

Fig. 1 superimposes the global MSE of the MLE and of the MAP estimator, the Cramér–Rao bound, the Bobrovsky-Zakaï bound, the Bayesian Abel bound, and the Weiss-Weinstein bound with optimization over s and $s = 1/2$.

$$\begin{cases} \eta\left(\frac{1}{2}, h\right) = -\frac{1}{4}\left(2\mathrm{SNR}\left(N - \sin^2(\pi hN) - \frac{1}{2}\frac{\sin(2\pi hN)}{\tan(\pi h)}\right) - \frac{h^2}{2\sigma_\theta^2}\right) \\ \eta(1,h) = 0 \\ \eta(1,-h) = 0 \\ \eta\left(\frac{1}{2}, 2h\right) = -\frac{1}{2}\left(\mathrm{SNR}\left(N - \sin^2(2\pi hN) - \frac{1}{2}\frac{\sin(4\pi hN)}{\tan(2\pi h)}\right) - \frac{h^2}{\sigma_\theta^2}\right) \end{cases} \tag{90}$$

This figure shows the threshold behavior of both estimators when the SNR decreases. In contrast to the Cramér–Rao bound, the Bobrovsky-Zakaï bound, the Bayesian Abel bound, and the Weiss-Weinstein bound exhibit the threshold phenomena. The Bayesian Abel bound is slightly higher than the Bobrovsky-Zaka ï bound and, consequently, leads to a better prediction of the threshold effect with the same computational cost. The Weiss-Weinstein bounds obtained by numerical evaluation of (88) and (91) are the same; therefore, $\mathrm{s} = 1/2$ seems to be the optimum value in this problem. As expected by the addition of constraints, the Weiss-Weinstein bounds provide a better prediction of the global MSE of the estimators in comparison with the Bobrovsky-Zakaï bound and the Bayesian Abel bound. The Weiss-Weinstein bound threshold value provides a better approximation of the effective SNR at which the estimators experience the threshold behavior.

VI. Conclusion

In this paper, we proposed a framework to study the Bayesian minimal bounds on the MSE of the Weiss-Weinstein family. This framework is based on both the best Bayesian bound (MMSE) and a constrained optimization problem. By rewriting the problem of the MMSEE as a continuous constrained optimization problem and by relaxing these constraints, we reobtain the lower bounds of the Weiss-Weinstein family. Moreover, this framework allows us to propose new minimal bounds. In this way, we propose an extension of the Bayesian Bhattacharyya bound and a Bayesian version of the Abel bound. Additionally, we give some closed-form expressions of several minimal bounds for both a general Gaussian observation model with parameterized mean and a spectral analysis model.

Appendix

A. Proof of Theorem 1

This proof is based on the three following lemmas.

Lemma 1:

$$\mathcal{C}_1 = \mathcal{C}_2. \tag{94}$$

Lemma 2:

$$\mathcal{C}_1 = \mathcal{C}_3. \tag{95}$$

Lemma 3:

$$\mathcal{C}_1 = \mathcal{C}_4. \tag{96}$$

Proof of Lemma 1:

- $C_2 \subset C_1$: we assume that $\mathrm{S}\forall f(\mathbf{y},\theta) \in \mathcal{F}, \int_\Theta \int_\Omega v(\mathbf{y},\theta)(\partial f(\mathbf{y},\theta)/\partial\theta)d\mathbf{y}d\theta = \int_\Theta \int_\Omega f(\mathbf{y},\theta)d\mathbf{y}d\theta.\mathrm{S}$

Since

$$\frac{\partial[v(\mathbf{y},\theta)f(\mathbf{y},\theta)]}{\partial\theta} = v(\mathbf{y},\theta)\frac{\partial f(\mathbf{y},\theta)}{\partial\theta} + \frac{\partial v(\mathbf{y},\theta)}{\partial\theta}f(\mathbf{y},\theta) \tag{97}$$

we have

$$\int_\Theta \int_\Omega \frac{\partial[v(\mathbf{y},\theta)f(\mathbf{y},\theta)]}{\partial\theta} - \frac{\partial v(\mathbf{y},\theta)}{\partial\theta} \times f(\mathbf{y},\theta)d\mathbf{y}d\theta = \int_\Theta \int_\Omega f(\mathbf{y},\theta)\, d\mathbf{y}d\theta \tag{98}$$

$$\Longrightarrow \int_\Theta \int_\Omega \left(1 + \frac{\partial v(\mathbf{y},\theta)}{\partial\theta}\right) f(\mathbf{y},\theta)d\mathbf{y}d\theta = 0. \tag{99}$$

Since (99) holds for any $f(\mathbf{y},\theta)$, if we choose $f(\mathbf{y},\theta) = 1 + (\partial v(\mathbf{y},\theta)/\partial\theta)$, we obtain

$$\begin{aligned}\int_\Theta \int_\Omega \left(1 + \frac{\partial v(\mathbf{y},\theta)}{\partial\theta}\right)^2 d\mathbf{y}d\theta = 0 \\ \Longrightarrow 1 + \frac{\partial v(\mathbf{y},\theta)}{\partial\theta} = 0 \\ \Longrightarrow v(\mathbf{y},\theta) = z(\mathbf{y}) - \theta\end{aligned} \tag{100}$$

where $z(\mathbf{y})$ is a function of $\mathbf{y}$ only.

- $C_1 \subset C_2$: on the other hand, if we assume that $v(\mathbf{y},\theta) = z(\mathbf{y}) - \theta$, then

$$\begin{aligned}&\int_\Theta \int_\Omega v(\mathbf{y},\theta)\frac{\partial f(\mathbf{y},\theta)}{\partial\theta}d\mathbf{y}d\theta \\ &= \int_\Theta \int_\Omega (z(\mathbf{y}) - \theta)\frac{\partial f(\mathbf{y},\theta)}{\partial\theta}d\mathbf{y}d\theta \\ &= \int_\Theta \int_\Omega \frac{\partial[(z(\mathbf{y}) - \theta)f(\mathbf{y},\theta)]}{\partial\theta} + f(\mathbf{y},\theta)d\mathbf{y}d\theta \\ &= \int_\Theta \int_\Omega f(\mathbf{y},\theta)d\mathbf{y}d\theta \quad \forall f(\mathbf{y},\theta) \in \mathcal{F}.\end{aligned} \tag{101}$$

These two items prove Lemma 1. ■

Proof of Lemma 2:

- $C_3 \subset C_1$: we assume that $\forall f(\mathbf{y},\theta) \in \mathcal{F}$ such that $\int_\Theta \int_\Omega f(\mathbf{y},\theta)d\mathbf{y}d\theta = 1$ and $\forall h$ such that $\theta + h \in \Theta$

$$\int_\Theta \int_\Omega v(\mathbf{y},\theta)(f(\mathbf{y},\theta + h) - f(\mathbf{y},\theta))d\mathbf{y}d\theta = h. \tag{102}$$

Then, when $h \to 0$, we have

$$\int_\Theta \int_\Omega v(\mathbf{y},\theta)\frac{\partial f(\mathbf{y},\theta)}{\partial\theta}d\mathbf{y}d\theta = 1 \Longrightarrow v(\mathbf{y},\theta) = z(\mathbf{y}) - \theta \tag{103}$$

thanks to the result of the first item of Lemma 1.

- $C_1 \subset C_3$: on the other hand, if we assume $v(\mathbf{y},\theta) = z(\mathbf{y}) - \theta$, then by setting $\varphi = \theta + h$

$$\begin{aligned}&\int_\Theta\int_\Omega v(\mathbf{y},\theta)f(\mathbf{y},\theta+h)d\mathbf{y}d\theta\\&\quad=\int_\Theta\int_\Omega (z(\mathbf{y})-\theta)f(\mathbf{y},\theta+h)d\mathbf{y}d\theta\\&\quad=\int_\Theta\int_\Omega (z(\mathbf{y})-\varphi+h)f(\mathbf{y},\varphi)d\mathbf{y}d\varphi\\&\quad=\int_\Theta\int_\Omega (z(\mathbf{y})-\varphi)f(\mathbf{y},\varphi)d\mathbf{y}d\varphi+h\end{aligned}\tag{104}$$

leading to

$$\int_\Theta\int_\Omega (z(\mathbf{y})-\theta)(f(\mathbf{y},\theta+h)-f(\mathbf{y},\theta))d\mathbf{y}d\theta = h \tag{105}$$

$\forall f(\mathbf{y},\theta)\in\mathcal{F}$ such that $\int_\Theta\int_\Omega f(\mathbf{y},\theta)d\mathbf{y}d\theta = 1$ and $\forall h$ such that $\theta + h \in \Theta$.

These two items prove Lemma 2. ■

Proof of Lemma 3:

- $C_4 \subset C_1$: let $L(\mathbf{y},\eta,\theta) = (f(\mathbf{y},\eta)/f(\mathbf{y},\theta))$ and assume that $\forall f(\mathbf{y},\theta) \in \mathcal{F}$ such that $\int_\Theta\int_\Omega f(\mathbf{y},\theta)d\mathbf{y}d\theta = 1, \forall h$ such that $\theta \pm h \in \Theta$ and $\forall s \in [0,1]$

$$\begin{aligned}&\int_\Theta\int_\Omega v(\mathbf{y},\theta)[L^s(\mathbf{y},\theta+h,\theta)\\&\quad-L^{1-s}(\mathbf{y},\theta-h,\theta)]f(\mathbf{y},\theta)d\mathbf{y}d\theta\\&\quad=h\int_\Theta\int_\Omega L^{1-s}(\mathbf{y},\theta-h,\theta)f(\mathbf{y},\theta)d\mathbf{y}d\theta.\end{aligned}\tag{106}$$

Then, when $s \to 1$, we obtain

$$\int_\Theta\int_\Omega v(\mathbf{y},\theta)(f(\mathbf{y},\theta+h)-f(\mathbf{y},\theta))d\mathbf{y}d\theta = h$$
$$\Longrightarrow v(\mathbf{y},\theta) = z(\mathbf{y}) - \theta \tag{107}$$

thanks to the result of the first item of Lemma 2.

- $C_1 \subset C_4$: on the other hand, if we assume $v(\mathbf{y},\theta) = z(\mathbf{y}) - \theta$, then by letting $\varphi = \theta + h$

$$\begin{aligned}&\int_\Theta\int_\Omega v(\mathbf{y},\theta)L^s(\mathbf{y},\theta+h,\theta)f(\mathbf{y},\theta)d\theta d\mathbf{y}\\&\quad=\int_\Theta\int_\Omega (z(\mathbf{y})-\theta)L^s(\mathbf{y},\theta+h,\theta)f(\mathbf{y},\theta)d\theta d\mathbf{y}\\&\quad=\int_\Theta\int_\Omega (z(\mathbf{y})-\varphi)L^{1-s}(\mathbf{y},\varphi-h,\varphi)f(\mathbf{y},\varphi)d\varphi d\mathbf{y}\\&\qquad+h\int_\Theta\int_\Omega L^{1-s}(\mathbf{y},\varphi-h,\varphi)f(\mathbf{y},\varphi)\,d\varphi d\mathbf{y}\end{aligned}\tag{108}$$

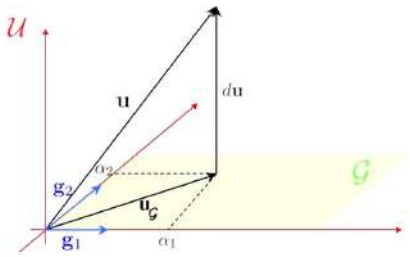


Fig. 2. Graphical representation of the problem.

leading to

$$\begin{aligned}&\int_\Theta\int_\Omega v(\mathbf{y},\theta)[L^s(\mathbf{y},\theta+h,\theta)\\&\quad-L^{1-s}(\mathbf{y},\theta-h,\theta)]\,f(\mathbf{y},\theta)d\theta d\mathbf{y}\\&\qquad=h\int_\Theta\int_\Omega L^{1-s}(\mathbf{y},\theta-h,\theta)f(\mathbf{y},\theta)\,d\theta d\mathbf{y}\end{aligned}\tag{109}$$

$\forall f(\mathbf{y},\theta) \in \mathcal{F}$ such that $\int_\Theta\int_\Omega f(\mathbf{y},\theta)d\mathbf{y}d\theta = 1, \forall h$ such that $\theta \pm h \in \Theta$ and $\forall s \in [0,1]$.

These two items prove Lemma 3. ■

Lemmas 1, 2, and 3 prove Theorem 1. ■

B. Proof of Theorem 2

Let $\mathcal{U}$ be a vector space of any dimension on the field of real numbers $\mathbb{R}$, with an inner product denoted by $\langle \mathbf{u},\mathbf{w}\rangle$, where $\mathbf{u}$ and $\mathbf{w}$ are two vectors of $\mathcal{U}$. Let $\{\mathbf{g}_1,\ldots,\mathbf{g}_K\}$ be a family of K independent vectors of $\mathcal{U}$ and $\mathbf{c} = [c_1 \quad \cdots \quad c_K]^T$ be a vector of $\mathbb{R}^K$. We are interested in the solution of the minimization of $\langle \mathbf{u},\mathbf{u}\rangle$ subject to the following K linear constraints $\langle \mathbf{u},\mathbf{g}_k\rangle = c_k, k \in [1,K]$.

Let $\mathcal{G}$ be the vectorial subspace of dimension K generated by the elements $\{\mathbf{g}_1,\ldots,\mathbf{g}_K\}$. Then, $\forall \mathbf{u} \in \mathcal{U}, \mathbf{u} = \mathbf{u}_\mathcal{G} + d\mathbf{u}$, where $\mathbf{u}_\mathcal{G}$ is the orthogonal projection of $\mathbf{u}$ on $\mathcal{G}$, i.e., the vector $\mathbf{u}_\mathcal{G} \in \mathcal{G}$ such that $\langle \mathbf{u}-\mathbf{u}_\mathcal{G},\mathbf{g}_k\rangle = 0, k \in [1,K]$ (see Fig. 2) for a graphical representation).

Let $\boldsymbol{\alpha} = [\alpha_1 \quad \cdots \quad \alpha_K]^T$ be the coordinates of $\mathbf{u}_\mathcal{G}$ in the basis $\{\mathbf{g}_1,\ldots,\mathbf{g}_K\}$ of $\mathcal{G}$ (i.e., $\mathbf{u}_\mathcal{G} = \sum_{k=1}^K \alpha_k\mathbf{g}_k$). These coordinates satisfy: $\langle \mathbf{u},\mathbf{g}_k\rangle = \langle \mathbf{u}_\mathcal{G},\mathbf{g}_k\rangle, k \in [1,K]$. Moreover, if $\mathbf{u}$ satisfies the K constraints $\langle \mathbf{u},\mathbf{g}_k\rangle = c_k, k \in [1,K]$, then

$$\begin{aligned}\langle \mathbf{u},\mathbf{g}_k\rangle = c_k&\\&\Rightarrow \langle \mathbf{u}_\mathcal{G},\mathbf{g}_k\rangle = c_k\\&\Rightarrow \left\langle \sum_{l=1}^K \alpha_l\mathbf{g}_l,\mathbf{g}_k\right\rangle = c_k\\&\Rightarrow \sum_{l=1}^K \alpha_l\langle \mathbf{g}_l,\mathbf{g}_k\rangle = c_k\end{aligned}\tag{110}$$

i.e., by a matricial rewriting $\mathbf{G}\boldsymbol{\alpha} = \mathbf{c}$, where $\mathbf{G}$ is the Gram matrix associated to the family $\{\mathbf{g}_1,\ldots,\mathbf{g}_K\}$: $G_{k,l} = \langle \mathbf{g}_l,\mathbf{g}_k\rangle$. The equation $\mathbf{G}\boldsymbol{\alpha} = \mathbf{c}$ has for unique solution $\boldsymbol{\alpha} = \mathbf{G}^{-1}\mathbf{c}$. Let

$\mathbf{u}_{\mathcal{G},\mathbf{c}}$ be the vector of $\mathcal{G}$ corresponding to this solution. Then, $\forall \mathbf{u} \in \mathcal{U}$ and for satisfying the K aforementioned constraints we have $\langle \mathbf{u},\mathbf{u}\rangle = \langle \mathbf{u}_{\mathcal{G},\mathbf{c}},\mathbf{u}_{\mathcal{G},\mathbf{c}}\rangle + \langle d\mathbf{u},d\mathbf{u}\rangle \geq \langle \mathbf{u}_{\mathcal{G},\mathbf{c}},\mathbf{u}_{\mathcal{G},\mathbf{c}}\rangle$, and the minimum is achieved for $d\mathbf{u}=0$, which means that $\mathbf{u}_{\mathcal{G},\mathbf{c}}$ is the solution of the problem. The value of the minimal norm is given by

$$\begin{aligned}\langle \mathbf{u}_{\mathcal{G},\mathbf{c}},\mathbf{u}_{\mathcal{G},\mathbf{c}}\rangle &= \left\langle \sum_{k=1}^{K}\alpha_k\mathbf{g}_k, \sum_{l=1}^{K}\alpha_l\mathbf{g}_l \right\rangle \\ &= \sum_{k=1}^{K}\sum_{l=1}^{K}\alpha_k\alpha_l\langle \mathbf{g}_k,\mathbf{g}_l\rangle \\ &= \boldsymbol{\alpha}^T\mathbf{G}\boldsymbol{\alpha} \\ &= (\mathbf{G}^{-1}\mathbf{c})^T\mathbf{G}\mathbf{G}^{-1}\mathbf{c} \\ &= \mathbf{c}^T\mathbf{G}^{-1}\mathbf{c}. \end{aligned} \tag{111}$$

■

C. Derivation of the Bayesian Abel Bound

We have to calculate the quadratic form $\mathbf{c}^T\mathbf{G}^{-1}\mathbf{c}$ (29). Since

$$\int_{\Theta}\int_{\Omega} p(\mathbf{x},\theta+h_i)d\theta d\mathbf{x} = 1 \quad \forall h_i \in \mathbb{R} \tag{112}$$

and, due to (9)

$$\int_{\Theta}\int_{\Omega} \frac{\partial^i p(\mathbf{x},\theta)}{\partial\theta^i} d\theta d\mathbf{x} = 0 \quad \forall i \geq 1 \tag{113}$$

the matrix $\mathbf{G} = \int_{\Theta}\int_{\Omega} \mathbf{g}\mathbf{g}^T d\theta d\mathbf{x}$ can now be written as the following partitioned matrix:

$$\mathbf{G} = \begin{pmatrix} 1 & \mathbf{0}_{1\times m} & \mathbf{0}_{1\times r} \\ \mathbf{0}_{m\times 1} & \mathbf{B} & \boldsymbol{\Gamma}^T \\ \mathbf{0}_{r\times 1} & \boldsymbol{\Gamma} & \mathbf{D} \end{pmatrix} \tag{114}$$

where the elements $\{\mathbf{B}\}_{i,j}$ and $\{\mathbf{D}\}_{i,j}$ of the matrices $\mathbf{B}(m\times m)$ and $\mathbf{D}(r\times r)$ are given by (37) and (48), respectively, and the element $\{\boldsymbol{\Gamma}\}_{i,j}$ of the matrix $\boldsymbol{\Gamma}(r\times m)$ is given by

$$\begin{aligned}\{\boldsymbol{\Gamma}\}_{i,j} &= \int_{\Theta}\int_{\Omega} \frac{p(\mathbf{x},\theta+h_i)-p(\mathbf{x},\theta)}{p(\mathbf{x},\theta)} \frac{\partial^j p(\mathbf{x},\theta)}{\partial\theta^j} d\theta d\mathbf{x} \\ &= \int_{\Theta}\int_{\Omega} \frac{p(\mathbf{x},\theta+h_i)}{p(\mathbf{x},\theta)} \frac{\partial^j p(\mathbf{x},\theta)}{\partial\theta^j} d\theta d\mathbf{x}. \end{aligned} \tag{115}$$

Let $\tilde{\mathbf{G}} = \begin{pmatrix} \mathbf{B} & \boldsymbol{\Gamma}^T \\ \boldsymbol{\Gamma} & \mathbf{D} \end{pmatrix}$ and $\mathbf{c} = [0 \;\; \boldsymbol{\alpha}^T \;\; \mathbf{h}^T]^T$, where $\boldsymbol{\alpha} = [1 \;\; 0 \;\; \cdots \;\; 0]^T$ (size $m\times 1$), and $\mathbf{h} = [h_1 \;\; \cdots \;\; h_r]^T$. Since the first element of $\mathbf{c}$ is null, only the right bottom corner $\tilde{\mathbf{G}}^{-1}$ (size $(m+r)\times(m+r)$) of $\mathbf{G}^{-1}$ is of interest. $\tilde{\mathbf{G}}^{-1}$ is given straightforwardly by

$$\tilde{\mathbf{G}}^{-1} = \begin{pmatrix} \mathbf{B} & \boldsymbol{\Gamma}^T \\ \boldsymbol{\Gamma} & \mathbf{D} \end{pmatrix}^{-1}. \tag{116}$$

Consequently, the Bayesian Abel bound denoted $\mathrm{BAB}_{m,r}$ is then given by

$$\mathrm{BAB}_{m,r} = [\boldsymbol{\alpha}^T \quad \mathbf{h}^T] \begin{pmatrix} \mathbf{B} & \boldsymbol{\Gamma}^T \\ \boldsymbol{\Gamma} & \mathbf{D} \end{pmatrix}^{-1} \begin{bmatrix} \boldsymbol{\alpha} \\ \mathbf{h} \end{bmatrix}. \tag{117}$$

After some algebraic effort, we obtain the final form

$$\mathrm{BAB}_{m,r} = \boldsymbol{\alpha}^T\mathbf{B}^{-1}\boldsymbol{\alpha} + \mathbf{u}^T\mathbf{J}^{-1}\mathbf{u} \tag{118}$$

with

$$\begin{cases} \mathbf{u} = \boldsymbol{\Gamma}\mathbf{B}^{-1}\boldsymbol{\alpha} - \mathbf{h} \\ \mathbf{J} = \mathbf{D} - \boldsymbol{\Gamma}\mathbf{B}^{-1}\boldsymbol{\Gamma}^T. \end{cases} \tag{119}$$

■

D. Minimal Bounds Derivation for the Gaussian Observation Model With Parameterized Mean

1) Bayesian Cramér–Rao Bound: The Bayesian Cramér–Rao bound can be divided into two terms [20]

$$\begin{aligned}&\mathrm{BCRB} \\ &= \left(\int_{\Theta} \mathrm{CRB}^{-1}(\theta)p(\theta)d\theta - \int_{\Theta} \frac{\partial^2 \ln p(\theta)}{\partial\theta^2} p(\theta)d\theta \right)^{-1} \end{aligned} \tag{120}$$

where $\mathrm{CRB}(\theta)$ is the standard (i.e., deterministic) Cramér–Rao bound given by [68]:

$$\mathrm{CRB}(\theta_0) = \frac{\sigma^2}{2\left\| \frac{\partial \mathbf{m}(\theta)}{\partial\theta}\Big|_{\theta_0} \right\|^2} \tag{121}$$

where θ_0 is the true value of the parameter in the deterministic context.

The second term of (120) is

$$\begin{aligned}&\int_{\Theta} \frac{\partial^2 \ln p(\theta)}{\partial\theta^2} p(\theta)d\theta \\ &\quad = -\frac{1}{2\sigma_\theta^2}\int_{\Theta} \frac{\partial^2(\theta-\mu)^2}{\partial\theta^2} p(\theta)d\theta \\ &\quad = -\frac{1}{\sigma_\theta^2}\int_{\Theta} p(\theta)d\theta = -\frac{1}{\sigma_\theta^2}. \end{aligned} \tag{122}$$

Consequently

$$\mathrm{BCRB} = \frac{\sigma_\theta^2}{\frac{2\sigma_\theta^2}{\sigma^2}\mathbb{E}_\theta\left[\left\|\frac{\partial \mathbf{m}(\theta)}{\partial\theta}\right\|^2\right] + 1}. \tag{123}$$

2) Bobrovsky-Zakaï Bound: The Bobrovsky-Zakaï bound is given by

$$\mathrm{BZB} = \sup_h \frac{h^2}{\int_{\Theta}\int_{\Omega} \frac{p^2(\mathbf{y},\theta+h)}{p(\mathbf{y},\theta)} d\mathbf{y}d\theta - 1}. \tag{124}$$

The double integral in the last equation can be rewritten as follows:

$$\int_{\Theta}\int_{\Omega}\frac{p^2(\mathbf{y},\theta+h)}{p(\mathbf{y},\theta)}d\mathbf{y}d\theta = \int_{\Theta}\frac{p^2(\theta+h)}{p(\theta)}\int_{\Omega}\frac{p^2(\mathbf{y}|\theta+h)}{p(\mathbf{y}|\theta)}d\mathbf{y}d\theta. \quad (125)$$

The term $(p^2(\mathbf{y}|\theta+h)/p(\mathbf{y}|\theta))$ becomes (126), shown at the bottom of the page.

Let $\mathbf{x}=\mathbf{y}-2\mathbf{m}(\theta+h)+\mathbf{m}(\theta)$, and note that

$$||\mathbf{x}||^2 = ||\mathbf{y}||^2 + ||2\mathbf{m}(\theta+h)-\mathbf{m}(\theta)||^2 - 2\mathrm{Re}\{\mathbf{y}^H(2\mathbf{m}(\theta+h)-\mathbf{m}(\theta))\}.$$

Consequently [see (127) at the bottom of the page].

The Bobrovsky-Zakaï bound is finally given by

$$\mathrm{BZB}=\sup_h \frac{h^2}{\int_{\Theta}\frac{p^2(\theta+h)}{p(\theta)}e^{\frac{2}{\sigma^2}||\mathbf{m}(\theta+h)-\mathbf{m}(\theta)||^2}d\theta-1}. \quad (128)$$

3) Bayesian Abel Bound: We have to calculate

$$\begin{aligned}\phi(h) &= \frac{1}{h}\int_{\Theta}\int_{\Omega}\frac{\partial \ln p(\mathbf{y},\theta)}{\partial\theta}p(\mathbf{y},\theta+h)d\mathbf{y}d\theta\\ &= \frac{1}{h}\int_{\Theta}p(\theta+h)\int_{\Omega}\left(\frac{\partial \ln p(\mathbf{y}|\theta)+\ln p(\theta)}{\partial\theta}\right)\\ &\quad\times p(\mathbf{y}|\theta+h)d\mathbf{y}d\theta\\ &= \frac{1}{h}\int_{\Theta}p(\theta+h)\int_{\Omega}\frac{\partial \ln p(\mathbf{y}|\theta)}{\partial\theta}p(\mathbf{y}|\theta+h)d\mathbf{y}d\theta\\ &\quad+\frac{1}{h}\int_{\Theta}\frac{\partial \ln p(\theta)}{\partial\theta}p(\theta+h)d\theta. \end{aligned} \quad (129)$$

The first term in (129) is given by

$$\begin{aligned}&\int_{\Omega}\frac{\partial \ln p(\mathbf{y}|\theta)}{\partial\theta}p(\mathbf{y}|\theta+h)d\mathbf{y}\\ &= -\frac{1}{\sigma^2}\int_{\Omega}\frac{\partial(\mathbf{y}-\mathbf{m}(\theta))^H(\mathbf{y}-\mathbf{m}(\theta))}{\partial\theta}p(\mathbf{y}|\theta+h)d\mathbf{y}\\ &= \frac{2}{\sigma^2}\int_{\Omega}\mathrm{Re}\left\{\frac{\partial\mathbf{m}^H(\theta)}{\partial\theta}(\mathbf{y}-\mathbf{m}(\theta))\right\}p(\mathbf{y}|\theta+h)d\mathbf{y}\\ &= \frac{2}{\sigma^2}\mathrm{Re}\left\{\frac{\partial\mathbf{m}^H(\theta)}{\partial\theta}\left(\int_{\Omega}\mathbf{y}p(\mathbf{y}|\theta+h)d\mathbf{y}-\mathbf{m}(\theta)\right)\right\}\\ &= \frac{2}{\sigma^2}\mathrm{Re}\left\{\frac{\partial\mathbf{m}^H(\theta)}{\partial\theta}(\mathbf{m}(\theta+h)-\mathbf{m}(\theta))\right\}. \end{aligned} \quad (130)$$

For the second term in (129), we have

$$\begin{aligned}&\frac{1}{h}\int_{\Theta}p(\theta+h)\frac{\partial \ln p(\theta)}{\partial\theta}d\theta\\ &= -\frac{1}{h\sigma_\theta^2}\int_{\Theta}(\theta-\mu)p(\theta+h)d\theta\\ &= \frac{1}{\sigma_\theta^2}. \end{aligned} \quad (131)$$

$$\begin{aligned}&\frac{p^2(\mathbf{y}|\theta+h)}{p(\mathbf{y}|\theta)}\\ &= \frac{1}{(\pi\sigma^2)^N}e^{-\frac{1}{\sigma^2}(2(\mathbf{y}-\mathbf{m}(\theta+h))^H(\mathbf{y}-\mathbf{m}(\theta+h))-(\mathbf{y}-\mathbf{m}(\theta))^H(\mathbf{y}-\mathbf{m}(\theta)))}\\ &= \frac{1}{(\pi\sigma^2)^N}e^{-\frac{1}{\sigma^2}(||\mathbf{y}||^2+2||\mathbf{m}(\theta+h)||^2-||\mathbf{m}(\theta)||^2-2\mathrm{Re}\{\mathbf{y}^H(2\mathbf{m}(\theta+h)-\mathbf{m}(\theta))\})}. \end{aligned} \quad (126)$$

$$\begin{aligned}&\int_{\Omega}\frac{p^2(\mathbf{y}|\theta+h)}{p(\mathbf{y}|\theta)}d\mathbf{y}\\ &= \frac{1}{(\pi\sigma^2)^N}\int_{\Omega}e^{-\frac{1}{\sigma^2}(||\mathbf{x}||^2+2||\mathbf{m}(\theta+h)||^2-||\mathbf{m}(\theta)||^2-||2\mathbf{m}(\theta+h)-\mathbf{m}(\theta)||^2)}d\mathbf{x}\\ &= \frac{1}{(\pi\sigma^2)^N}e^{-\frac{1}{\sigma^2}(2||\mathbf{m}(\theta+h)||^2+||\mathbf{m}(\theta)||^2-||2\mathbf{m}(\theta+h)-\mathbf{m}(\theta)||^2)}\underbrace{\int_{\Omega}e^{-\frac{1}{\sigma^2}||\mathbf{x}||^2}d\mathbf{x}}_{=(\pi\sigma^2)^N}\\ &= e^{\frac{2}{\sigma^2}||\mathbf{m}(\theta+h)-\mathbf{m}(\theta)||^2}. \end{aligned} \quad (127)$$

Finally

$$\phi(h) = \frac{1}{\sigma_\theta^2} + \frac{2}{h\sigma^2} \times \mathbb{E}_{\theta+h}\left[\mathrm{Re}\left\{\frac{\partial \mathbf{m}^H(\theta)}{\partial \theta}(\mathbf{m}(\theta+h)-\mathbf{m}(\theta))\right\}\right]. \quad (132)$$

4) Weiss-Weinstein Bound: We have to calculate

$$\eta(\alpha,\beta) = \ln \int_\Theta \int_\Omega \frac{p^\alpha(\mathbf{y},\theta+\beta)}{p^{\alpha-1}(\mathbf{y},\theta)} d\mathbf{y} d\theta. \quad (133)$$

This function can be modified as follows:

$$\eta(\alpha,\beta) = \ln \int_\Theta \frac{p^\alpha(\theta+\beta)}{p^{\alpha-1}(\theta)} \int_\Omega \frac{p^\alpha(\mathbf{y}|\theta+\beta)}{p^{\alpha-1}(\mathbf{y}|\theta)} d\mathbf{y} d\theta. \quad (134)$$

Let us first study the term shown in (135) at the bottom of the page.

Let $\mathbf{x} = \mathbf{y} - (\alpha\mathbf{m}(\theta+\beta) - (\alpha-1)\mathbf{m}(\theta))$. Note that

$$\|\mathbf{x}\|^2 = \|\mathbf{y}\|^2 + \|\alpha\mathbf{m}(\theta+\beta) - (\alpha-1)\mathbf{m}(\theta)\|^2 - 2\mathrm{Re}\{\mathbf{y}^H(\alpha\mathbf{m}(\theta+\beta) - (\alpha-1)\mathbf{m}(\theta))\}. \quad (136)$$

Consequently, [see (137) at the bottom of the page].

For the second term

$$\frac{p^\alpha(\theta+\beta)}{p^{\alpha-1}(\theta)} = \frac{1}{\sqrt{2\pi}\sigma_\theta} e^{-\frac{1}{2\sigma_\theta^2}[\alpha(\theta+\beta-\mu)^2 - (\alpha-1)(\theta-\mu)^2]}. \quad (138)$$

Finally, the semi-invariant moment generating function is given by (139), shown at the bottom of the page.

E. Bayesian Bounds Derivation for a Spectral Analysis Problem

1) Cramér–Rao Bound: The Bayesian Cramér–Rao bound is given by (123)

$$\mathrm{BCRB} = \frac{\sigma_\theta^2}{\frac{2\sigma_\theta^2}{\sigma^2}\mathbb{E}_\theta\left[\left\|\frac{\partial \mathbf{m}(\theta)}{\partial \theta}\right\|^2\right] + 1}. \quad (140)$$

The term $\|(\partial \mathbf{m}(\theta)/\partial\theta)\|^2$ can be written

$$\left\|\frac{\partial \mathbf{m}(\theta)}{\partial \theta}\right\|^2 = \left\|a\frac{\partial \mathbf{s}(\theta)}{\partial \theta}\right\|^2 = \sum_{k=0}^{N-1} a^2(j2\pi k e^{j2\pi k\theta})(-j2\pi k e^{-j2\pi k\theta}) = a^2 4\pi^2 \sum_{k=0}^{N-1} k^2 = \frac{2(a\pi)^2}{3} N(2N-1)(N-1) \quad (141)$$

which is independent of θ. Consequently, the Bayesian Cramér–Rao bound is

$$\mathrm{BCRB} = \frac{\sigma_\theta^2}{\mathrm{SNR}\frac{4\pi^2\sigma_\theta^2}{3}N(2N-1)(N-1)+1}. \quad (142)$$

$$\frac{p^\alpha(\mathbf{y}|\theta+\beta)}{p^{\alpha-1}(\mathbf{y}|\theta)} = \frac{1}{(\pi\sigma^2)^N} e^{-\frac{1}{\sigma^2}(\alpha(\mathbf{y}-\mathbf{m}(\theta+\beta))^H(\mathbf{y}-\mathbf{m}(\theta+\beta)) - (\alpha-1)(\mathbf{y}-\mathbf{m}(\theta))^H(\mathbf{y}-\mathbf{m}(\theta)))} = \frac{1}{(\pi\sigma^2)^N} e^{-\frac{1}{\sigma^2}(\|\mathbf{y}\|^2 + \alpha\|\mathbf{m}(\theta+\beta)\|^2 - (\alpha-1)\|\mathbf{m}(\theta)\|^2 - 2\mathrm{Re}\{\mathbf{y}^H(\alpha\mathbf{m}(\theta+\beta)-(\alpha-1)\mathbf{m}(\theta))\})}. \quad (135)$$

$$\int_\Omega \frac{p^\alpha(\mathbf{y}|\theta+\beta)}{p^{\alpha-1}(\mathbf{y}|\theta)} d\mathbf{y} = \frac{1}{(\pi\sigma^2)^N} \int_\Omega e^{-\frac{1}{\sigma^2}(\|\mathbf{x}\|^2 - \|\alpha\mathbf{m}(\theta+\beta)-(\alpha-1)\mathbf{m}(\theta)\|^2 + \alpha\|\mathbf{m}(\theta+\beta)\|^2 - (\alpha-1)\|\mathbf{m}(\theta)\|^2)} d\mathbf{x} = \frac{1}{(\pi\sigma^2)^N} e^{-\frac{1}{\sigma^2}(-\|\alpha\mathbf{m}(\theta+\beta)-(\alpha-1)\mathbf{m}(\theta)\|^2 + \alpha\|\mathbf{m}(\theta+\beta)\|^2 - (\alpha-1)\|\mathbf{m}(\theta)\|^2)} \underbrace{\int_\Omega e^{-\frac{\|x\|^2}{\sigma^2}} d\mathbf{x}}_{=(\pi\sigma^2)^N} = e^{\frac{\alpha(\alpha-1)}{\sigma^2}\|\mathbf{m}(\theta+\beta)-\mathbf{m}(\theta)\|^2}. \quad (137)$$

$$\eta(\alpha,\beta) = \ln \frac{1}{\sqrt{2\pi}\sigma_\theta} \int_\Theta e^{\frac{\alpha(\alpha-1)}{\sigma^2}\|\mathbf{m}(\theta+\beta)-\mathbf{m}(\theta)\|^2 - \frac{1}{2\sigma_\theta^2}(\theta-(\sqrt{\alpha(\alpha-1)}-\alpha)h-\mu)(\theta+(\sqrt{\alpha(\alpha-1)}+\alpha)h-\mu)} d\theta. \quad (139)$$

2) Bobrovsky-Zakaï Bound: The Bobrovsky-Zakaï bound is given by (128)

$$\mathrm{BZB} = \sup_h \frac{h^2}{\int\limits_\Theta \frac{p^2(\theta+h)}{p(\theta)} e^{\frac{2}{\sigma^2}\|\mathbf{m}(\theta+h)-\mathbf{m}(\theta)\|^2} d\theta - 1}. \quad (143)$$

In the case of our specific model (79), the term $\|\mathbf{m}(\theta+h)-\mathbf{m}(\theta)\|^2$ can be written

$$\begin{aligned}
&\|\mathbf{m}(\theta+h)-\mathbf{m}(\theta)\|^2 \\
&\quad = a^2 \sum_{k=0}^{N-1} (e^{j2\pi k(\theta+h)} - e^{j2\pi k\theta})(e^{-j2\pi k(\theta+h)} - e^{-j2\pi k\theta}) \\
&\quad = a^2 \sum_{k=0}^{N-1} (2 - 2\mathrm{Re}\{e^{j2\pi kh}\}) = 2a^2 \sum_{k=0}^{N-1} 1 - \cos(2\pi kh) \\
&\quad = 2a^2 (N - \sin^2(\pi hN) - \frac{1}{2}\frac{\sin(2\pi hN)}{\tan(\pi h)}) \quad (144)
\end{aligned}$$

which is independent of θ. The term $\int\limits_\Theta (p^2(\theta+h)/p(\theta))d\theta$ becomes

$$\begin{aligned}
&\int\limits_\Theta \frac{p^2(\theta+h)}{p(\theta)} d\theta \\
&\quad = \frac{1}{\sqrt{2\pi}\sigma_\theta} e^{-\frac{1}{2\sigma_\theta^2}[2h^2-4h\mu+\mu^2]} \int\limits_\Theta e^{-\frac{1}{2\sigma_\theta^2}[\theta^2+2\theta(2h-\mu)]} d\theta \\
&\quad = e^{-\frac{1}{2\sigma_\theta^2}[2h^2-4h\mu+\mu^2]+\frac{(2h-\mu)^2}{2\sigma_\theta^2}} = e^{\frac{h^2}{\sigma_\theta^2}} \quad (145)
\end{aligned}$$

where the term $\int\limits_\Theta e^{-(1/2\sigma_\theta^2)[\theta^2+2\theta(2h-\mu)]} d\theta$ is given by [71, p. 355, eq. (BI (28) (1))]

$$\int\limits_{-\infty}^{\infty} e^{-p^2x^2 \pm qx} dx = \frac{\sqrt{\pi}}{\mathrm{abs}(p)} e^{\frac{q^2}{4p^2}}. \quad (146)$$

Finally, the Bobrovsky-Zakaï is given by

$$\mathrm{BZB} = \sup_h \frac{h^2}{e^{4\mathrm{SNR}(N-\sin^2(\pi hN)-\frac{1}{2}\frac{\sin(2\pi hN)}{\tan(\pi h)})+\frac{h^2}{\sigma_\theta^2}} - 1}. \quad (147)$$

3) Bayesian Abel Bound: We have to calculate (132)

$$\begin{aligned}
\phi(h) &= \frac{1}{\sigma_\theta^2} + \frac{2}{h\sigma^2} \\
&\quad \mathbb{E}_{\theta+h}\left[\mathrm{Re}\left\{\frac{\partial \mathbf{m}^H(\theta)}{\partial\theta}(\mathbf{m}(\theta+h)-\mathbf{m}(\theta))\right\}\right]. \quad (148)
\end{aligned}$$

The term $\mathrm{Re}\{(\partial \mathbf{m}^H(\theta)/\partial\theta)(\mathbf{m}(\theta+h)-\mathbf{m}(\theta))\}$ can be rewritten as follows:

$$\begin{aligned}
&\mathrm{Re}\left\{\frac{\partial \mathbf{m}^H(\theta)}{\partial\theta}(\mathbf{m}(\theta+h)-\mathbf{m}(\theta))\right\} \\
&\quad = \mathrm{Re}\left\{a^2 \frac{\partial \mathbf{s}^H(\theta)}{\partial\theta}(\mathbf{s}(\theta+h)-\mathbf{s}(\theta))\right\} \\
&\quad = a^2 \mathrm{Re}\left\{\sum_{k=0}^{N-1} (e^{j2\pi k(\theta+h)} - e^{j2\pi k\theta}) \frac{\partial e^{-j2\pi k\theta}}{\partial\theta}\right\} \\
&\quad = -2\pi a^2 \mathrm{Re}\left\{\sum_{k=0}^{N-1} jk(e^{j2\pi kh} - 1)\right\} \\
&\quad = 2\pi a^2 \sum_{k=0}^{N-1} k\sin(2\pi kh) \\
&\quad = \pi a^2 \left(N\frac{\cos(2\pi hN)}{\tan(\pi h)} \right. \\
&\qquad \left. - \sin(2\pi hN)\left(\frac{1}{2\sin(\pi h)} + N\right)\right) \quad (149)
\end{aligned}$$

which is independent of θ. Consequently

$$\begin{aligned}
\phi(h) = \frac{1}{\sigma_\theta^2} + \frac{2\pi\mathrm{SNR}}{h}&\left(N\frac{\cos(2\pi hN)}{\tan(\pi h)} \right. \\
&\left. - \sin(2\pi hN)\left(\frac{1}{2\sin(\pi h)} + N\right)\right). \quad (150)
\end{aligned}$$

4) Weiss-Weinstein Bound: We have to calculate (139)

$$\begin{aligned}
\eta(\alpha,\beta) &= \ln \int\limits_\Theta \frac{p^\alpha(\theta+\beta)}{p^{\alpha-1}(\theta)} e^{\frac{\alpha(\alpha-1)}{\sigma^2}\|\mathbf{m}(\theta+\beta)-\mathbf{m}(\theta)\|^2} d\theta \\
&= \ln\left(e^{2\alpha(\alpha-1)\mathrm{SNR}(N-\sin^2(\pi\beta N)-\frac{1}{2}\frac{\sin(2\pi\beta N)}{\tan(\pi\beta)})} \right. \\
&\qquad \left. \times \int\limits_\Theta \frac{p^\alpha(\theta+\beta)}{p^{\alpha-1}(\theta)} d\theta\right) \quad (151)
\end{aligned}$$

thanks to (144) and to the independence of θ in the term $\|\mathbf{m}(\theta+\beta)-\mathbf{m}(\theta)\|^2$.

The remaining term is given by

$$\begin{aligned}
&\int\limits_\Theta \frac{p^\alpha(\theta+\beta)}{p^{\alpha-1}(\theta)} d\theta \\
&\quad = \frac{1}{\sqrt{2\pi}\sigma_\theta} e^{-\frac{1}{2\sigma_\theta^2}[\alpha\beta^2-2\alpha\beta\mu+\mu^2]} \int\limits_\Theta e^{-\frac{1}{2\sigma_\theta^2}[\theta^2+2\theta(\alpha\beta-\mu)]} d\theta \\
&\quad = e^{-\frac{1}{2\sigma_\theta^2}[\alpha\beta^2-2\alpha\beta\mu+\mu^2]+\frac{(\alpha\beta-\mu)^2}{2\sigma_\theta^2}} \\
&\quad = e^{-\frac{\alpha\beta^2}{2\sigma_\theta^2}(1-\alpha)} \quad (152)
\end{aligned}$$

where $\int\limits_\Theta e^{-(1/2\sigma_\theta^2)[\theta^2+2\theta(\alpha\beta-m)]} d\theta$ is obtained thanks to [71, p. 355, eq. (BI(28)(1))].

Consequently, $\eta(\alpha,\beta)$ is given by

$$\begin{aligned}
\eta(\alpha,\beta) = \alpha(\alpha-1)&\left(2\mathrm{SNR}\left(N \right.\right. \\
&\left.\left. - \sin^2(\pi\beta N) - \frac{1}{2}\frac{\sin(2\pi\beta N)}{\tan(\pi\beta)}\right) - \frac{\beta^2}{2\sigma_\theta^2}\right). \quad (153)
\end{aligned}$$

ACKNOWLEDGMENT

This paper was developed while A. Renaux was a Postdoctoral Research Associate in Prof. Nehorai's research group at the Department of Electrical and Systems Engineering, Washington University, St. Louis.

REFERENCES

[1] A. Renaux, P. Forster, and P. Larzabal, "A new derivation of the Bayesian bounds for parameter estimation," in *Proc. IEEE Statist. Signal Process. Workshop—SSP05*, Bordeaux, France, Jul. 2005, pp. 567–572.

[2] A. Renaux, P. Forster, P. Larzabal, and C. Richmond, "The Bayesian Abel bound on the mean square error," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Toulouse, France, May 2006, vol. 3, pp. 9–12.

[3] R. A. Fisher, "On the mathematical foundations of theoretical statistics," *Phil. Trans. Royal Soc.*, vol. 222, p. 309, 1922.

[4] D. Dugué, "Application des propriétés de la limite au sens du calcul des probabilités à l'étude des diverses questions d'estimation," *Ecol. Poly.*, vol. 3, pp. 305–372, 1937.

[5] M. Frechet, "Sur l'extension de certaines evaluations statistiques an cas de petit echantillons," *Rev. Inst. Int. Statist.*, vol. 11, pp. 182–205, 1943.

[6] G. Darmois, "Sur les lois limites de la dispersion de certaines estimations," *Rev. Inst. Int. Statist.*, vol. 13, pp. 9–15, 1945.

[7] C. R. Rao, "Information and accuracy attainable in the estimation of statistical parameters," *Bull. Calcutta Math. Soc.*, vol. 37, pp. 81–91, 1945.

[8] H. Cramér, *Mathematical Methods of Statistics*. New York: Princeton Univ. Press, 1946.

[9] L. L. Scharf and T. McWhorter, "Geometry of the Cramer Rao bound," *Signal Process.*, vol. 31, pp. 301–311, 1993.

[10] J. Xavier and V. Barroso, "The Riemannian geometry of certain parameter estimation problems with singular Fisher information matrices," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Montreal, Canada, May 2004, vol. 2, pp. 1021–1024.

[11] J. Xavier and V. Barroso, "Intrinsic variance lower bound (IVLB): An extension of the Cramer Rao bound to Riemannian manifolds," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Hong Kong, Mar. 2004, vol. 5, pp. 1033–1036.

[12] A. Manikas, *Differential Geometry in Array Processing*. London, U.K.: Imperial College Press, 2004.

[13] S. T. Smith, "Statistical resolution limits and the complexified Cramer Rao bound," *IEEE Trans. Signal Process.*, vol. 53, pp. 1597–1609, May 2005.

[14] S. T. Smith, "Covariance, subspace, and intrinsic Cramer Rao bounds," *IEEE Trans. Signal Process.*, vol. 53, pp. 1610–1630, May 2005.

[15] J.-P. Delmas and H. Abeida, "Stochastic Cramer-Rao bound for noncircular signals with application to DOA estimation," *IEEE Trans. Signal Process.*, vol. 52, pp. 3192–3199, Nov. 2004.

[16] J.-P. Delmas and H. Abeida, "Cramer Rao bounds of DOA estimates for BPSK and QPSK modulated signals," *IEEE Trans. Signal Process.*, vol. 54, pp. 117–126, Jan. 2005.

[17] E. Chaumette, P. Larzabal, and P. Forster, "On the influence of a detection step on lower bounds for deterministic parameter estimation," *IEEE Trans. Signal Process.*, vol. 53, pp. 4080–4090, Nov. 2005.

[18] Q. Zou, Z. Lin, and R. J. Ober, "The Cramer Rao lower bound for bilinear systems," *IEEE Trans. Signal Process.*, vol. 54, pp. 1666–1680, May 2006.

[19] I. Yetik and A. Nehorai, "Performance bounds on image registration," *IEEE Trans. Signal Process.*, vol. 54, pp. 1737–1736, May 2006.

[20] H. L. Van Trees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968, vol. 1.

[21] D. C. Rife and R. R. Boorstyn, "Single tone parameter estimation from discrete time observations," *IEEE Trans. Inf. Theory*, vol. 20, pp. 591–598, 1974.

[22] A. Bhattacharyya, "On some analogues of the amount of information and their use in statistical estimation," *Sankhya Indian J. Statist.*, vol. 8, pp. 1–14, 201–218, 315–328, 1946.

[23] D. A. S. Fraser and I. Guttman, "Bhattacharyya bounds without regularity assumptions," *Ann. Math. Stat.*, vol. 23, pp. 629–632, 1952.

[24] D. G. Chapman and H. Robbins, "Minimum variance estimation without regularity assumptions," *Ann. Math. Stat.*, vol. 22, pp. 581–586, 1951.

[25] J. Kiefer, "On minimum variance estimators," *Ann. Math. Stat.*, vol. 23, pp. 627–629, 1952.

[26] J. M. Hammersley, "On estimating restricted parametrers," *J. Royal Soc. Ser. B*, vol. 12, pp. 192–240, 1950.

[27] E. W. Barankin, "Locally best unbiased estimates," *Ann. Math. Stat.*, vol. 20, pp. 477–501, 1949.

[28] R. J. McAulay and L. P. Seidman, "A useful form of the Barankin lower bound and its application to ppm threshold analysis," *IEEE Trans. Inf. Theory*, vol. 15, pp. 273–279, Mar. 1969.

[29] J. S. Abel, "A bound on mean square estimate error," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Albuquerque, NM, 1990, vol. 3, pp. 1345–1348.

[30] J. S. Abel, "A bound on mean square estimate error," *IEEE Trans. Inf. Theory*, vol. 39, pp. 1675–1680, Sep. 1993.

[31] R. J. McAulay and E. M. Hofstetter, "Barankin bounds on parameter estimation," *IEEE Trans. Inf. Theory*, vol. 17, pp. 669–676, Nov. 1971.

[32] A. Quinlan, E. Chaumette, and P. Larzabal, "A direct method to generate approximations of the Barankin bound," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Toulouse, France, May 2006, vol. 3, pp. 808–811.

[33] J. Ziv and M. Zakai, "Some lower bounds on signal parameter estimation," *IEEE Trans. Inf. Theory*, vol. 15, pp. 386–391, May 1969.

[34] S. Bellini and G. Tartara, "Bounds on error in signal parameter estimation," *IEEE Trans. Commun.*, vol. 22, pp. 340–342, Mar. 1974.

[35] D. Chazan, M. Zakai, and J. Ziv, "Improved lower bounds on signal parameter estimation," *IEEE Trans. Inf. Theory*, vol. 21, pp. 90–93, Jan. 1975.

[36] E. Weinstein, "Relations between Belini-Tartara, Chazan-Zakai-Ziv, and Wax-Ziv lower bounds," *IEEE Trans. Inf. Theory*, vol. 34, pp. 342–343, Mar. 1988.

[37] K. L. Bell, Y. Steinberg, Y. Ephraim, and H. L. V. Trees, "Extended Ziv Zakai lower bound for vector parameter estimation," *IEEE Trans. Signal Process.*, vol. 43, pp. 624–638, Mar. 1997.

[38] K. L. Bell, "Performance bounds in parameter estimation with application to bearing estimation," Ph.D. dissertation, George Mason Univ., Fairfax, VA, 1995.

[39] B. Z. Bobrovsky, E. Mayer-Wolf, and M. Zakai, "Some classes of global Cramer Rao bounds," *Ann. Statist.*, vol. 15, pp. 1421–1438, 1987.

[40] B. Z. Bobrovsky and M. Zakai, "A lower bound on the estimation error for certain diffusion processes," *IEEE Trans. Inf. Theory*, vol. 22, pp. 45–52, Jan. 1976.

[41] I. Reuven and H. Messer, "A Barankin-type lower bound on the estimation error of a hybrid parameter vector," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 1084–1093, May 1997.

[42] A. J. Weiss and E. Weinstein, "A lower bound on the mean square error in random parameter estimation," *IEEE Trans. Inf. Theory*, vol. 31, pp. 680–682, Sep. 1985.

[43] , H. L. Van Trees and K. L. Bell, Eds., *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. New York: Wiley, 2007.

[44] A. B. Baggeroer, Barankin Bound on the Variance of Estimates of Gaussian Random Process MIT Lincoln Lab., Lexington, MA, Tech. Rep., Jan. 1969.

[45] I. Reuven and H. Messer, "The use of the Barankin bound for determining the threshold SNR in estimating the bearing of a source in the presence of another," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Detroit, MI, May 1995, vol. 3, pp. 1645–1648.

[46] L. Knockaert, "The Barankin bound and threshold behavior in frequency estimation," *IEEE Trans. Signal Process.*, vol. 45, pp. 2398–2401, Sep. 1997.

[47] J. Tabrikian and J. L. Krolik, "Barankin bound for source localization in shallow water," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Munich, Germany, Apr. 1997.

[48] T. L. Marzetta, "Computing the Barankin bound by solving an unconstrained quadratic optimization problem," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Munich, Germany, Apr. 1997, vol. 5, pp. 3829–3832.

[49] I. Reuven and H. Messer, "On the effect of nuisance parameters on the threshold SNR value of the Barankin bound," *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 523–527, Feb. 1999.

[50] J. Tabrikian and J. Krolik, "Barankin bounds for source localization in an uncertain ocean environment," *IEEE Trans. Signal Process.*, vol. 47, pp. 2917–2927, Nov. 1999.

[51] A. Ferrari and J.-Y. Tourneret, "Barankin lower bound for change points in independent sequences," in *Proc. IEEE Statist. Signal Process. Workshop—SSP03*, St. Louis, MO, Sep. 2003, pp. 557–560.

[52] P. Ciblat, M. Ghogho, P. Forster, and P. Larzabal, "Harmonic retrieval in the presence of non-circular Gaussian multiplicative noise: Performance bounds," *Signal Process.*, vol. 85, pp. 737–749, Apr. 2005.

[53] P. Ciblat, P. Forster, and P. Larzabal, "Harmonic retrieval in noncircular complex-valued multiplicative noise: Barankin bound," in *Proc. EUSIPCO*, Vienne, Australia, Sep. 2004, pp. 2151–2154.

[54] L. Atallah, J. P. Barbot, and P. Larzabal, "SNR threshold indicator in data aided frequency synchronization," *IEEE Signal Process. Lett.*, vol. 11, pp. 652–654, Aug. 2004.

[55] L. Atallah, J. P. Barbot, and P. Larzabal, "From Chapman Robbins bound towards Barankin bound in threshold behavior prediction," *Electron. Lett.*, vol. 40, pp. 279–280, Feb. 2004.
[56] A. Renaux, L. N. Atallah, P. Forster, and P. Larzabal, "A useful form of the Abel bound and its application to estimator threshold prediction," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2365–2369, May 2007.
[57] W. Xu, "Performances bounds on matched-field methods for source localization and estimation of ocean environmental parameters," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, Jun. 2001.
[58] W. Xu, A. B. Baggeroer, and K. Bell, "A bound on mean-square estimation with background parameter mismatch," *IEEE Trans. Inf. Theory*, vol. 50, pp. 621–632, Apr. 2004.
[59] W. Xu, A. B. Baggeroer, and C. D. Richmond, "Bayesian bounds for matched-field parameter estimation," *IEEE Trans. Signal Process.*, vol. 52, pp. 3293–3305, Dec. 2004.
[60] A. J. Weiss and E. Weinstein, "Fundamental limitation in passive time delay estimation Part 1: Narrowband systems," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, pp. 472–486, Apr. 1983.
[61] K. L. Bell, Y. Ephraim, and H. L. V. Trees, "Ziv Zakai lower bounds in bearing estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Detroit, MI, 1995, vol. 5, pp. 2852–2855.
[62] K. L. Bell, Y. Ephraim, and H. L. V. Trees, "Explicit Ziv Zakai lower bound for bearing estimation," *IEEE Trans. Signal Process.*, vol. 44, pp. 2810–2824, Nov. 1996.
[63] P. Ciblat and M. Ghogho, "Ziv Zakai bound for harmonic retrieval in multiplicative and additive Gaussian noise," in *Proc. IEEE Statist. Signal Process. Workshop—SSP05*, Bordeaux, France, Jul. 2005, pp. 561–566.
[64] A. Renaux, "Weiss-Weinstein bound for data aided carrier estimation," *IEEE Signal Process. Lett.*, vol. 14, no. 4, pp. 283–286, Apr. 2007.
[65] F. E. Glave, "A new look at the Barankin lower bound," *IEEE Trans. Inf. Theory*, vol. 18, no. 3, pp. 349–356, May 1972.
[66] P. Forster and P. Larzabal, "On lower bounds for deterministic parameter estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Orlando, FL, 2002, vol. 2, pp. 1137–1140.
[67] E. Weinstein and A. J. Weiss, "A general class of lower bounds in parameter estimation," *IEEE Trans. Inf. Theory*, vol. 34, pp. 338–342, Mar. 1988.
[68] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1993, vol. I.
[69] A. Renaux, "Contribution a panalyse des performances d'estimaiion en traitement statisticue du signal" Ph.D. dissertation, Ecole Normale Superieure de Cachan, Cachan, France, Jul. 2006 [Online]. Available: http://www.satie.ens-c;achan.fr/ts/These_Alex.pdf
[70] H. L. Van Trees, *Detection, Estimation and Modulation Theory: Radar-Sonar Signal Processing and Gaussian Signals in Noise*. New York: Wiley, 2001, vol. 3.
[71] S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. San Diego, CA: Academic, 1994.

Alexandre Renaux (S'06–M'08) comes from a small French mountain called "les Vosges." He received the Agrégation, the M.Sc., and the Ph.D. degrees in electrical engineering from the École Normale Supérieure de Cachan, France, in 2002, 2003, and 2006, respectively.

From 2006 to 2007, he was a Postdoctoral Research Associate with the Department of Electrical and Systems Engineering, Washington University, St. Louis, MO. He is currently an Assistant Professor with the Department of Physics and a Member of the Laboratory of Signals and Systems (L2S), University Paris 11, France. His research interests include detection and parameter estimation theory in statistical signal and array processing.

Philippe Forster (M'89) was born in Brest, France, in 1960. He received the Agrégation de Physique Appliquée degree from the École Normale Supérieure de Cachan, France, in 1983, and the Ph.D. degree in electrical engineering in 1988 from the University of Rennes, France.

He is currently a Professor of Electrical Engineering with the Université de Nanterre, IUT de Ville d'Avray, France, and member of the SATIE Laboratory, École Normale Supérieure de Cachan, France. His research interests are in the field of statistical signal processing and array processing.

Pascal Larzabal (M'93) was born in the Basque country in the south of France in 1962. He received the Agrégation in Electrical Engineering in 1988, the Ph.D. degree in 1992, and the "habilitation à diriger les recherches" in 1998, all from the École Normale Supérieure de Cachan, France.

He is now Professor of Electrical Engineering with the University of Paris-Sud 11, France. He teaches electronics, signal processing, control, and mathematics. From 1998 to 2003, he was the Director of IUP GEII, University of Paris-Sud 11. From March 2003 to March 2007, he was Head of the Electrical Engineering Department, Institut Universitaire de Technologie of Cachan. Since January 2007, he has been the Director of the Laboratory SATIE (UMR CNRS 8029, École Normale Supérieure de Cachan). Since 1993, he has been Head of the Signal Processing Group, Laboratory SATIE. His research concerns estimation in array processing and spectral analysis for wave-front identification, radars, communications, tomography, and medical imaging. His recent works concern estimator performances and associated minimal bounds.

Christ D. Richmond (M'99–SM'05) received the B.S. degree in electrical engineering from the University of Maryland, College Park, and the B.S. degree in mathematics from Bowie State University, Bowie, MD. He received the S.M., E.E., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge.

He is currently a member of the Technical Research Staff at MIT Lincoln Laboratory, Lexington. His research interests include detection and parameter estimation theory, sensor array and multichannel signal processing, statistical signal processing, random matrix theory, radar/sonar signal processing, multivariate statistical analysis, information theory, multiuser detection, multiinput/multioutput (MIMO) systems, and wireless communications.

Dr. Richmond is the recipient of the Office of Naval Research Graduate Fellowship Award 1990–1994, the Alan Berman Research Publications Award March 1994 (Naval Research Laboratory), and the IEEE Signal Processing Society 1999 Young Author Best Paper Award in area of Sensor Array and Multi-channel (SAM) Signal Processing. He served as the Technical Chairman of the Adaptive Sensor Array Processing (ASAP) Workshop, MIT Lincoln Laboratory, 2007, 2006, and 1998, and served as a member the IEEE Technical Committee on SAM Signal Processing. He was an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2002 to 2005. He was an invited reviewer for the book *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, by Prof. H. L. Van Trees and Prof. K. Bell of George Mason University, Eds. (Piscataway, NJ: IEEE, 2007).

Arye Nehorai (S'80–M'83–SM'90–F'94) received the B.Sc. and M.Sc. degrees in electrical engineering from the Technion, Israel, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA.

From 1985 to 1995, he was a faculty member with the Department of Electrical Engineering, Yale University, New Haven, CT. In 1995, he joined the Department of Electrical Engineering and Computer Science, The University of Illinois at Chicago (UIC) as Full Professor. From 2000 to 2001, he was Chair of the department's Electrical and Computer Engineering (ECE) Division, which then became a new department. In 2001, he was named University Scholar of the University of Illinois. In 2006, he became Chairman of the Department of Electrical and Systems Engineering, Washington University, St. Louis (WUSTL), MO. He is the inaugural holder of the Eugene and Martha Lohman Professorship and the Director of the Center for Sensor Signal and Information Processing (CSSIP) at WUSTL since 2006.

Dr. Nehorai was Editor-in-Chief of the IEEE TRANSACTIONS ON SIGNAL PROCESSING during 2000 to 2002. During 2003–2005, he was Vice President (Publications) of the IEEE Signal Processing Society, Chair of the Publications Board, member of the Board of Governors, and member of the Executive Committee of this Society. From 2003 to 2006, he was the Founding Editor of the Special Columns on Leadership Reflections in the IEEE SIGNAL PROCESSING MAGAZINE. He was corecipient of the IEEE SPS 1989 Senior Award for Best Paper with P. Stoica, coauthor of the 2003 Young Author Best Paper Award, and corecipient of the 2004 Magazine Paper Award with A. Dogandzic. He was elected Distinguished Lecturer of the IEEE SPS for the term 2004 to 2005 and received the 2006 IEEE SPS Technical Achievement Award. He is the Principal Investigator of the new multidisciplinary university research initiative (MURI) project entitled Adaptive Waveform Diversity for Full Spectral Dominance. He has been a Fellow of the Royal Statistical Society since 1996.

Annexe D

Statistical analysis of the covariance matrix MLE in K-distributed clutter

Elsevier Signal Processing, Volume : 90, Issue : 4, Apr. 2010, pp. 1165-1175

Signal Processing 90 (2010) 1165–1175

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Statistical analysis of the covariance matrix MLE in K-distributed clutter

Frédéric Pascal [a,*], Alexandre Renaux [b]

[a] SONDRA/Supelec, 3 rue Joliot-Curie, F-91192 Gif-sur-Yvette Cedex, France
[b] Laboratoire des Signaux et Systemes (L2S), Université Paris-Sud XI (UPS), CNRS, SUPELEC, Gif-Sur-Yvette, France

ARTICLE INFO

Article history:
Received 22 January 2009
Received in revised form
13 July 2009
Accepted 26 September 2009
Available online 17 October 2009

Keywords:
Covariance matrix estimation
Cramér–Rao bound
Maximum likelihood
K-distribution
Spherically invariant random vector

ABSTRACT

In the context of radar detection, the clutter covariance matrix estimation is an important point to design optimal detectors. While the Gaussian clutter case has been extensively studied, the new advances in radar technology show that non-Gaussian clutter models have to be considered. Among these models, the *spherically invariant random vector* modelling is particularly interesting since it includes the K-distributed clutter model, known to fit very well with experimental data. This is why recent results in the literature focus on this distribution. More precisely, the maximum likelihood estimator of a K-distributed clutter covariance matrix has already been derived. This paper proposes a complete statistical performance analysis of this estimator through its consistency and its unbiasedness at finite number of samples. Moreover, the closed-form expression of the true Cramér–Rao bound is derived for the K-distribution covariance matrix and the efficiency of the maximum likelihood estimator is emphasized by simulations.

Notations

The notational convention adopted is as follows: italic indicates a scalar quantity, as in A; lower case boldface indicates a vector quantity, as in $\mathbf{a}$; upper case boldface indicates a matrix quantity, as in $\mathbf{A}$. $\mathcal{R}e\{A\}$ and $\mathcal{I}m\{A\}$ are the real and the imaginary parts of A, respectively. The complex conjugation, the matrix transpose operator, and the conjugate transpose operator are indicated by *, ^{T}, and ^{H}, respectively. The j th element of a vector $\mathbf{a}$ is denoted $a^{(j)}$. The n th row and m th column element of the matrix $\mathbf{A}$ will be denoted by $A_{n,m}$. $|\mathbf{A}|$ and $\mathrm{Tr}(\mathbf{A})$ are the determinant and the trace of the matrix $\mathbf{A}$, respectively. $\otimes$ denotes the Kronecker product. $\|\cdot\|$ denotes any matrix norm. The operator vec($\mathbf{A}$) stacks the columns of the matrix $\mathbf{A}$ one under another into a single column vector. The operator vech($\mathbf{A}$), where $\mathbf{A}$ is a symmetric matrix, does the same things as vec($\mathbf{A}$) with the upper triangular portion excluded. The operator veck($\mathbf{A}$) of a skew-symmetric matrix (i.e., $\mathbf{A}^T = -\mathbf{A}$) does the same thing as vech($\mathbf{A}$) by omitting the diagonal elements. The identity matrix, with appropriate dimensions, is denoted $\mathbf{I}$ and the zero matrix is denoted $\mathbf{0}$. $E[\cdot]$ denotes the expectation operator. $\xrightarrow{a.s.}$ stands for the almost sure convergence and $\xrightarrow{Pr}$ stands for the convergence in probability. A zero-mean complex circular Gaussian distribution with covariance matrix $\mathbf{A}$ is denoted $\mathcal{CN}(\mathbf{0},\mathbf{A})$. A gamma distribution with shape parameter k and scale parameter θ is denoted $\mathcal{G}(k,\theta)$. A complex m-variate K-distribution with parameters k, θ, and covariance matrix $\mathbf{A}$ is denoted $\mathcal{K}_m(k,\theta,\mathbf{A})$. A central chi-square distribution with k degrees of freedom is denoted $\chi^2(k)$. A uniform distribution with boundaries a and b is denoted $\mathcal{U}_{[a,b]}$.

1. Introduction

The Gaussian assumption makes sense in many applications, e.g., sources localization in passive sonar,

* Corresponding author. Tel.: +33 147405320; fax: +33 147402199.
E-mail addresses: frederic.pascal@supelec.fr (F. Pascal), alexandre.renaux@lss.supelec.fr (A. Renaux).

0165-1684/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2009.09.029

radar detection where thermal noise and clutter are generally modelled as Gaussian processes. In these contexts, Gaussian models have been thoroughly investigated in the framework of statistical estimation and detection theory (see, e.g., [1–3]). They have led to attractive algorithms such as the stochastic maximum likelihood method [4,5] or Bayesian estimators.

However, the assumption of Gaussian noise is not always valid. For instance, due to the recent evolution of radar technology, one can cite the area of space time adaptive processing-high resolution (STAP-HR) where the resolution is such that the central limit theorem cannot be applied anymore since the number of backscatters is too small. Equivalently, it is known that reflected signals can be very impulsive when they are collected by a low grazing angle radar [6,7]. This is why, in the last decades, the radar community has been very interested in problems dealing with non-Gaussian clutter modelling (see, e.g., [8–11]).

One of the most general non-Gaussian noise model is provided by spherically invariant random vectors (SIRV) which are a compound processes [12–14]. More precisely, an SIRV is the product of a Gaussian random vector (the so-called speckle) with the square root of a non-negative random scalar variable (the so-called texture). In other words, a noise modelled as an SIRV is a non-homogeneous Gaussian process with random power. Thus, these kind of processes are fully characterized by the texture and the unknown covariance matrix of the speckle. One of the major challenging difficulties in SIRV modelling is to estimate these two unknown quantities [15]. These problems have been investigated in [16] for the texture estimation while [17,18] have proposed different estimation procedures for the covariance matrix. Moreover, the knowledge of these estimates accuracy is essential in radar detection since the covariance matrix and the texture are required to design the different detection schemes.

In this context, this paper focuses on parameters estimation performance where the clutter is modelled by a K-distribution. A K-distribution is an SIRV, with a gamma distributed texture depending on two real positive parameters α and β. Consequently, a K-distribution depends on α, β and on the covariance matrix $\mathbf{M}$. This model choice is justified by the fact that a lot of operational data experimentations have shown the good agreement between real data and the K-distribution model (see [7,19–22] and references herein).

This K-distribution model has been extensively studied in the literature. First, concerning the parameters estimation problem, [23,24] have estimated the gamma distribution parameters assuming that $\mathbf{M}$ is equal to the identity matrix, [17] has proposed a recursive algorithm for the covariance matrix $\mathbf{M}$ estimation assuming α and β known and [25] has used a parameter-expanded expectation-maximization (PX-EM) algorithm for the covariance matrix $\mathbf{M}$ estimation and for a parameter ν assuming $\nu=\alpha=1/\beta$. Note also that estimation schemes in K-distribution context can be found in [26,27] and references herein. Second, concerning the statistical performance of these estimators, it has been proved in [28] that the recursive scheme proposed by [17] converges and has a unique solution which is the maximum likelihood (ML) estimator. Consequently, this estimator has become very attractive. In order to evaluate the ultimate performance in terms of mean square error, [23] has derived the true Cramér–Rao Bound (CRB) for the parameters of the gamma texture (namely, α and β) assuming $\mathbf{M}$ equal to the identity matrix. Gini [29] has derived the modified CRB on the one-lag correlation coefficient of $\mathbf{M}$ where the parameters of the gamma texture are assumed to be nuisance parameters. Concerning the covariance matrix $\mathbf{M}$, a first approach for the true CRB study, which is known to be tighter than the modified one, has been proposed in [25] whatever the texture distribution. However, note that, for the particular case of a gamma distributed texture, the analysis of [25] involves several numerical integrations and no useful information concerning the structure of the Fisher information matrix (FIM) is given. Finally, classical covariance matrix estimators are compared in [30] in the more general context of SIRV.

The knowledge of an accurate covariance matrix estimate is of the utmost interest in context of radar detection since this matrix is always involved in the detector expression [30]. Therefore, the goal of this contribution is twofold. First, the covariance matrix ML estimate statistical analysis is provided in terms of consistency and bias. Second, the closed-form expression of the true CRB for the covariance matrix $\mathbf{M}$ is given and is analyzed. Finally, through a discussion and simulation results, classical estimation procedures in Gaussian and SIRV contexts are compared.

The paper is organized as follows. Section 2 presents the problem formulation while Sections 3 and 4 contain the main results of this paper: the ML estimate statistical performance in terms of consistency and bias and the derivation of the true CRB. Finally, Section 5 gives simulations which validate theoretical results.

2. Problem formulation

In radar detection, the basic problem consists in detecting if a known signal corrupted by an additive clutter is present or not. In order to estimate the clutter parameters before detection, it is generally assumed that K signal-free independent measurements, traditionally called the secondary data $\mathbf{c}_k$, $k=1,\ldots,K$ are available.

As stated in the introduction, one considers a clutter modelled thanks to a K-distribution denoted

$$\mathbf{c}_k \sim \mathcal{K}_m(\alpha,(2/\beta)^2,\mathbf{M}). \tag{1}$$

From the SIRV definition, $\mathbf{c}_k$ can be written as

$$\mathbf{c}_k=\sqrt{\tau_k}\mathbf{x}_k, \tag{2}$$

where τ_k is gamma distributed with parameters α and $(2/\beta)^2$, i.e., $\tau_k\sim\mathcal{G}(\alpha,(2/\beta)^2)$ and, where $\mathbf{x}_k$ is a complex circular zero-mean m-dimensional Gaussian vector with covariance matrix $E[\mathbf{x}_k\mathbf{x}_k^H]=\mathbf{M}$ independent of τ_k. For identifiability considerations, $\mathbf{M}$ is normalized according to $\mathrm{Tr}(\mathbf{M})=m$ (see [17]). Note that the parameter α represents the spikiness of the clutter. Indeed, when α is

high the clutter tends to be Gaussian and, when α is small, the tail of the clutter becomes heavy.

The probability density function (PDF) of a random variable τ_k distributed according to $\mathcal{G}(\alpha,(2/\beta)^2)$ is given by

$$p(\tau_k)=\left(\frac{\beta^2}{4}\right)^{\alpha}\frac{\tau_k^{\alpha-1}}{\Gamma(\alpha)}\exp\left(-\frac{\beta^2}{4}\tau_k\right), \tag{3}$$

where $\Gamma(\alpha)$ is the gamma function defined by

$$\Gamma(\alpha)=\int_0^{+\infty}x^{\alpha-1}\exp(-x)\,dx. \tag{4}$$

From Eq. (2), the PDF of $\mathbf{c}_k$ can be written

$$p(\mathbf{c}_k;\mathbf{M},\alpha,\beta)=\int_0^{+\infty}\frac{1}{\tau_k^m\pi^m|\mathbf{M}|}\exp\left(-\frac{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k}{\tau_k}\right)p(\tau_k)\,d\tau_k, \tag{5}$$

which is equal to

$$p(\mathbf{c}_k;\mathbf{M},\alpha,\beta)=\frac{\beta^{\alpha+m}(\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k)^{(\alpha-m)/2}}{2^{\alpha+m-1}\pi^m|\mathbf{M}|\Gamma(\alpha)}K_{m-\alpha}\left(\beta\sqrt{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k}\right), \tag{6}$$

where $K_\nu(\cdot)$ is the modified Bessel function of the second kind of order ν [31].

Gini et al. have derived the ML estimator as the solution of the following equation [17]:

$$\hat{\mathbf{M}}_{ML}=\frac{1}{K}\sum_{k=1}^{K}c_m(\mathbf{c}_k^H\hat{\mathbf{M}}_{ML}^{-1}\mathbf{c}_k)\mathbf{c}_k\mathbf{c}_k^H, \tag{7}$$

where the function $c_m(q)$ is defined as

$$c_m(q)=\frac{\beta}{2\sqrt{q}}\frac{K_{\alpha-m-1}(\beta\sqrt{q})}{K_{\alpha-m}(\beta\sqrt{q})}. \tag{8}$$

Note that the ML estimate $\hat{\mathbf{M}}_{ML}$ has to be normalized as $\mathbf{M}:\mathrm{Tr}(\hat{\mathbf{M}}_{ML})=m$. Finally, it has been shown in [28] that the solution to Eq. (7) exists and is unique for the aforementioned normalization.

3. Statistical analysis of $\hat{\mathbf{M}}_{ML}$

This section is devoted to the statistical analysis of $\hat{\mathbf{M}}_{ML}$ in terms of consistency and bias.

3.1. Consistency

An estimator $\hat{\mathbf{M}}$ of $\mathbf{M}$ is said to be consistent if

$$\|\hat{\mathbf{M}}-\mathbf{M}\|\underset{K\to+\infty}{\overset{Pr}{\to}}0,$$

where K is the number of secondary data $\mathbf{c}_k$'s used to estimate $\mathbf{M}$.

Theorem 3.1 ($\hat{\mathbf{M}}_{ML}$ *consistency*). *$\hat{\mathbf{M}}_{ML}$ is a consistent estimate of $\mathbf{M}$.*

Proof. In the sequel, $\hat{\mathbf{M}}_{ML}$ will be denoted $\hat{\mathbf{M}}(K)$ to show explicitly the dependence between $\hat{\mathbf{M}}_{ML}$ and the number K of $\mathbf{x}_k'$s. Let us define the function $f_{K,\mathbf{M}}$ such that

$$f_{K,\mathbf{M}}:\begin{cases}\mathcal{D}\longrightarrow\mathcal{D},\\ \mathbf{A}\longrightarrow\dfrac{1}{K}\displaystyle\sum_{k=1}^{K}c_m(\mathbf{c}_k^H\mathbf{A}^{-1}\mathbf{c}_k)\mathbf{c}_k\mathbf{c}_k^H,\end{cases} \tag{9}$$

where $c_m(\cdot)$ is defined by Eq. (8), where $\mathcal{D}=\{\mathbf{A}\in M_m(\mathbb{C})|\mathbf{A}^H=\mathbf{A},\mathbf{A}$ positive definite matrix$\}$ with $M_m(\mathbb{C})=\{m\times m$ matrices with elements in $\mathbb{C}\}$, and where $\mathbb{C}$ is the set of complex scalar. As $\hat{\mathbf{M}}(K)$ is a fixed point of function $f_{K,\mathbf{M}}$, it is the unique zero, which respects the constraint $\mathrm{Tr}(\hat{\mathbf{M}}(K))=m$, of the following function:

$$g_K:\begin{cases}\mathcal{D}\longrightarrow\mathcal{D},\\ \mathbf{A}\longrightarrow g_K(\mathbf{A})=\mathbf{A}-f_{K,\mathbf{M}}(\mathbf{A}).\end{cases}$$

To prove the consistency of $\hat{\mathbf{M}}(K)$, Theorem 5.9 of [32, p. 46] will be used. First, the strong law of large numbers (SLLN) gives

$$\forall\mathbf{A}\in\mathcal{D},\quad g_K(\mathbf{A})\underset{K\to+\infty}{\overset{a.s.}{\to}}g(\mathbf{A}),$$

where

$$\forall\mathbf{A}\in\mathcal{D},\quad g(\mathbf{A})=\mathbf{A}-E[c_m(\mathbf{c}^H\mathbf{A}^{-1}\mathbf{c})\mathbf{c}\mathbf{c}^H] \tag{10}$$

for $\mathbf{c}\sim\mathcal{K}_m(\alpha,(2/\beta)^2,\mathbf{M})$.

Let us now apply the change of variable $\mathbf{y}=\mathbf{A}^{-1/2}\mathbf{c}$. We obtain

$$\mathbf{y}\sim\mathcal{K}_m(\alpha,(2/\beta)^2,\mathbf{A}^{-1/2}\mathbf{M}\mathbf{A}^{-1/2})$$

and

$$\forall\mathbf{A}\in\mathcal{D},\quad g(\mathbf{A})=\mathbf{A}^{1/2}(\mathbf{I}-E[c_m(\mathbf{y}^H\mathbf{y})\mathbf{y}\mathbf{y}^H])\mathbf{A}^{1/2}$$

and

$$\forall\mathbf{A}\in\mathcal{D},\quad g_K(\mathbf{A})=\mathbf{A}^{1/2}\left(\mathbf{I}-\frac{1}{K}\sum_{k=1}^{K}c_m(\mathbf{y}_k^H\mathbf{y}_k)\mathbf{y}_k\mathbf{y}_k^H\right)\mathbf{A}^{1/2}.$$

Let us verify the hypothesis of Theorem 5.9 of [32, p. 46]. We have to prove that for every $\varepsilon>0$,

$$(\mathrm{H}_1):\sup_{\mathbf{A}\in\mathcal{D}}\{\|g_K(\mathbf{A})-g(\mathbf{A})\|\}\underset{K\to+\infty}{\overset{Pr}{\to}}0,$$

$$(\mathrm{H}_2):\inf_{\mathbf{A}:\|\mathbf{A}-\mathbf{M}\|\geq\varepsilon}\{\|g(\mathbf{A})\|\}>0=g(\mathbf{M}).$$

For every $\mathbf{A}\in\mathcal{D}$, we have

$$\|g_K(\mathbf{A})-g(\mathbf{A})\|=\|\mathbf{A}\|\left\|\frac{1}{K}\sum_{k=1}^{K}(c_m(\mathbf{y}_k^H\mathbf{y}_k)\mathbf{y}_k\mathbf{y}_k^H-\mathbb{E}[c_m(\mathbf{y}^H\mathbf{y})\mathbf{y}\mathbf{y}^H])\right\|.$$

Since $E[c_m(\mathbf{y}^H\mathbf{y})\mathbf{y}\mathbf{y}^H]<+\infty$, one can apply the SLLN to the K i.i.d. variables $c_m(\mathbf{y}_k^H\mathbf{y}_k)\mathbf{y}_k\mathbf{y}_k^H$, with same first order moment. This ensures (H_1).

Moreover, the function $c_m(\mathbf{c}^H\mathbf{A}^{-1}\mathbf{c})$ is strictly decreasing w.r.t. $\mathbf{A}$. Consequently, $E[c_m(\mathbf{c}^H\mathbf{A}^{-1}\mathbf{c})\mathbf{c}\mathbf{c}^H]$ too. This implies that $E[c_m(\mathbf{c}^H\mathbf{A}^{-1}\mathbf{c})\mathbf{c}\mathbf{c}^H]\neq\mathbf{A}$, except for $\mathbf{A}=\mathbf{M}$. This ensures (H_2).

Finally, Theorem 5.9 of [32, p. 46] concludes the proof and $\hat{\mathbf{M}}_{ML}\underset{K\to+\infty}{\overset{Pr}{\to}}\mathbf{M}$. □

3.2. Bias

This subsection provides an analysis of the bias B defined by $B(\hat{\mathbf{M}}_{ML})=E[\hat{\mathbf{M}}_{ML}]-\mathbf{M}$.

Theorem 3.2 (*Unbiasedness of $\hat{\mathbf{M}}_{ML}$*). *$\hat{\mathbf{M}}_{ML}$ is an unbiased estimate of $\mathbf{M}$ at finite distance* (i.e., *at finite number K*).

Proof. For the sake of simplicity, $\hat{\mathbf{M}}_{ML}$ will be denoted $\hat{\mathbf{M}}$ in this part. By applying the following change of variable, $\mathbf{y}_k = \mathbf{M}^{-1/2}\mathbf{c}_k$, to Eq. (7), one has

$$\hat{\mathbf{M}} = \frac{1}{K}\sum_{k=1}^{K} c_m(\mathbf{y}_k^H\hat{\mathbf{T}}^{-1}\mathbf{y}_k)\mathbf{M}^{1/2}\mathbf{y}_k\mathbf{y}_k^H\mathbf{M}^{1/2},$$

where

$$\hat{\mathbf{T}} = \mathbf{M}^{-1/2}\hat{\mathbf{M}}\mathbf{M}^{-1/2}.$$

Therefore,

$$\hat{\mathbf{T}} = \frac{1}{K}\sum_{k=1}^{K} c_m(\mathbf{y}_k^H\hat{\mathbf{T}}^{-1}\mathbf{y}_k)\mathbf{y}_k\mathbf{y}_k^H.$$

$\hat{\mathbf{T}}$ is thus the unique estimate (see [28, Theorem III.1]) of the identity matrix, with $\mathrm{Tr}(\hat{\mathbf{T}}) = m$. Its statistic is clearly independent of $\mathbf{M}$ since the $\mathbf{y}_k$'s are i.i.d. SIRVs with a gamma distributed texture and identity matrix for the Gaussian covariance matrix. In other words, $\mathbf{y}_k \sim \mathcal{K}_m(\alpha, (2/\beta)^2, \mathbf{I})$.

Moreover, for any unitary matrix $\mathbf{U}$,

$$\mathbf{U}\hat{\mathbf{T}}\mathbf{U}^H = \frac{1}{K}\sum_{k=1}^{K} c_m(\mathbf{z}_k^H(\mathbf{U}\hat{\mathbf{T}}\mathbf{U}^H)^{-1}\mathbf{z}_k)\mathbf{z}_k\mathbf{z}_k^H,$$

where $\mathbf{z}_k = \mathbf{U}\mathbf{y}_k$ are also i.i.d. and distributed as $\mathcal{K}_m(\alpha, (2/\beta)^2, \mathbf{I})$ and $\mathbf{U}\hat{\mathbf{T}}\mathbf{U}^H$ has the same distribution as $\hat{\mathbf{T}}$. Consequently,

$$E[\hat{\mathbf{T}}] = \mathbf{U}E[\hat{\mathbf{T}}]\mathbf{U}^H \quad \text{for any unitary matrix } \mathbf{U}.$$

Since $E[\hat{\mathbf{T}}]$ is different from $\mathbf{0}$, Lemma A.1 of [30] ensures that $E[\hat{\mathbf{T}}] = \gamma\mathbf{I}$ for $\gamma \in \mathbb{R}$. Remind that $\hat{\mathbf{T}} = \mathbf{M}^{-1/2}\hat{\mathbf{M}}\mathbf{M}^{-1/2}$, then $E[\hat{\mathbf{M}}] = \gamma\mathbf{M}$. Moreover, since $\mathrm{Tr}(\hat{\mathbf{M}}) = \mathrm{Tr}(\mathbf{M}) = m$, one has

$$m = E(\mathrm{Tr}(\hat{\mathbf{M}})) = \mathrm{Tr}(E(\hat{\mathbf{M}})) = \gamma\mathrm{Tr}(\mathbf{M}) = \gamma m, \tag{11}$$

which implies that $\gamma = 1$.

In conclusion, $\hat{\mathbf{M}}$ is an unbiased estimate of $\mathbf{M}$, for any number K of secondary data. □

3.3. Comments

Theorems 3.1 and 3.2 show the attractiveness of the estimator (7) in terms of statistical properties, i.e., consistency and unbiasedness. Note also that this estimator is robust since the unbiasedness property is at finite number of samples. In the next section, the Cramér–Rao bound is derived for the observation model (2).

4. Cramér–Rao bound

In this Section, the Cramér–Rao bound w.r.t. $\mathbf{M}$ is derived. The CRB gives the best variance that an unbiased estimator can achieve. The proposed bound will be compared to the mean square error of the previously studied ML estimator (Eq. (7)) in the next section.

The CRB for a parameter vector $\boldsymbol{\theta}$ is given by

$$\mathbf{CRB}_{\boldsymbol{\theta}} = \left(-E\left[\frac{\partial^2 \ln p(\mathbf{c}_1,\ldots,\mathbf{c}_K;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}^T}\right]\right)^{-1}, \tag{12}$$

where $p(\mathbf{c}_1,\ldots,\mathbf{c}_K;\boldsymbol{\theta})$ is the likelihood function of the observations $\mathbf{c}_k$, $k = 1,\ldots,K$. Concerning our model, the parameter vector is

$$\boldsymbol{\theta} = (\mathrm{vech}^T(\mathcal{R}e\{\mathbf{M}\})\ \mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\}))^T, \quad m^2 \times 1. \tag{13}$$

With the parametrization of Eq. (13), the structure of CRB becomes

$$\mathbf{CRB}_{\boldsymbol{\theta}} = \begin{pmatrix} \mathbf{F}_{1,1} & \mathbf{F}_{1,2} \\ \mathbf{F}_{2,1} & \mathbf{F}_{2,2} \end{pmatrix}^{-1}, \tag{14}$$

where the $\mathbf{F}_{i,j}$'s are the elements of the FIM given by

$$\mathbf{F}_{1,1} = -E\left[\frac{\partial^2 \ln p(\mathbf{c}_1,\ldots,\mathbf{c}_K;\boldsymbol{\theta})}{\partial\mathrm{vech}(\mathcal{R}e\{\mathbf{M}\})\partial\mathrm{vech}^T(\mathcal{R}e\{\mathbf{M}\})}\right], \quad \frac{m(m+1)}{2}\times\frac{m(m+1)}{2}, \tag{15}$$

$$\mathbf{F}_{1,2} = \mathbf{F}_{2,1}^T = -E\left[\frac{\partial^2 \ln p(\mathbf{c}_1,\ldots,\mathbf{c}_K;\boldsymbol{\theta})}{\partial\mathrm{vech}(\mathcal{R}e\{\mathbf{M}\})\partial\mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})}\right], \quad \frac{m(m+1)}{2}\times\frac{m(m-1)}{2}, \tag{16}$$

$$\mathbf{F}_{2,2} = -E\left[\frac{\partial^2 \ln p(\mathbf{c}_1,\ldots,\mathbf{c}_K;\boldsymbol{\theta})}{\partial\mathrm{veck}(\mathcal{I}m\{\mathbf{M}\})\partial\mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})}\right], \quad \frac{m(m-1)}{2}\times\frac{m(m-1)}{2}. \tag{17}$$

Since the $\mathbf{c}_k$'s are i.i.d. random vectors, one have from Eq. (6),

$$p(\mathbf{c}_1,\ldots,\mathbf{c}_K;\boldsymbol{\theta}) = \prod_{k=1}^{K}\frac{\beta^{\alpha+m}(\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k)^{(\alpha-m)/2}}{2^{\alpha+m-1}\pi^m|\mathbf{M}|\Gamma(\alpha)}K_{m-\alpha}(\beta\sqrt{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k}). \tag{18}$$

Consequently, the log-likelihood function can be written as

$$\ln p(\mathbf{c}_1,\ldots,\mathbf{c}_K|\boldsymbol{\theta}) = K\ln\left(\frac{\beta^{\alpha+m}}{2^{\alpha+m-1}\pi^m\Gamma(\alpha)}\right) - K\ln(|\mathbf{M}|) + \sum_{k=1}^{K}\ln((\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k)^{(\alpha-m)/2}K_{m-\alpha}(\beta\sqrt{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k})). \tag{19}$$

4.1. Result

The next subsections will show that the CRB w.r.t. $\boldsymbol{\theta}$, in this context, is given by

$$\mathbf{CRB}_{\boldsymbol{\theta}} = \frac{1}{K}\begin{pmatrix} (\mathbf{H}^T\mathbf{F}\mathbf{H})^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{K}^T\mathbf{F}\mathbf{K})^{-1} \end{pmatrix}, \tag{20}$$

where the matrices $\mathbf{H}$ and $\mathbf{K}$ are constant transformation matrices filled with ones and zeros such that $\mathrm{vec}(\mathbf{A}) = \mathbf{H}\mathrm{vech}(\mathbf{A})$ and $\mathrm{vec}(\mathbf{A}) = \mathbf{K}\mathrm{veck}(\mathbf{A})$, with $\mathbf{A}$ a skew-symmetric matrix, and where

$$\mathbf{F} = (\mathbf{M}^T\otimes\mathbf{M})^{-1} - \left(\frac{2(\alpha+1)}{m+1} - \frac{\varphi(\alpha,m)}{8}\right) \times(\mathbf{M}^T\otimes\mathbf{M})^{-1/2}(\mathbf{I}+\mathrm{vec}(\mathbf{I})\mathrm{vec}^T(\mathbf{I}))(\mathbf{M}^T\otimes\mathbf{M})^{-1/2}, \tag{21}$$

where $\varphi(\alpha,m)$ is given by Eq. (F.3).

Remarks.

- $\varphi(\alpha,m)$ is a constant which does not depend on β since $\beta^2\tau_{k'} \sim \mathcal{G}(\alpha,4)$ (see Eq. (F.3)).

- The first term of the right hand side of Eq. (21) is the Gaussian FIM (i.e., when $\tau_k = 1\ \forall k$). Indeed, using Eqs. (24), (49) and (30),[1] the Gaussian CRB, denoted $\mathbf{GCRB}_{\boldsymbol{\theta}}$, is straightforwardly obtained as

$$\mathbf{GCRB}_{\boldsymbol{\theta}} = \frac{1}{K}\begin{pmatrix} (\mathbf{H}^T(\mathbf{M}^T\otimes\mathbf{M})^{-1}\mathbf{H})^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{K}^T(\mathbf{M}^T\otimes\mathbf{M})^{-1}\mathbf{K})^{-1}\end{pmatrix}. \tag{22}$$

By identification with Eq. (21), it means that

$$\lim_{\alpha\to\infty}\left(\frac{2(\alpha+1)}{m+1}-\frac{\gamma}{8}\right)=0.$$

Consequently, due to the structure of $\mathbf{I}+\mathrm{vec}(\mathbf{I})\mathrm{vec}^T(\mathbf{I})$, the FIM for K-distributed observations is given by the Gaussian FIM minus a sparse matrix depending on α, $\mathbf{M}$ and m.

4.2. Outline of the proof

To make the reading easier, only the outline of the CRB derivation is given below. All the details are reported into the different appendices.

4.2.1. Analysis of $\mathbf{F}_{1,1}$

With Eq. (19) one has

$$\begin{aligned}&\frac{\partial^2\ln p(\mathbf{c}_1,\ldots,\mathbf{c}_K;\boldsymbol{\theta})}{\partial\mathrm{vech}(\mathcal{R}e\{\mathbf{M}\})\partial\mathrm{vech}^T(\mathcal{R}e\{\mathbf{M}\})}\\ &\quad=-K\frac{\partial^2\ln(|\mathbf{M}|)}{\partial\mathrm{vech}(\mathcal{R}e\{\mathbf{M}\})\partial\mathrm{vech}^T(\mathcal{R}e\{\mathbf{M}\})}\\ &\qquad+\sum_{k=1}^{K}\frac{\partial^2\ln((\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k)^{(\alpha-m)/2}K_{m-\alpha}(\beta\sqrt{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k}))}{\partial\mathrm{vech}(\mathcal{R}e\{\mathbf{M}\})\partial\mathrm{vech}^T(\mathcal{R}e\{\mathbf{M}\})}.\end{aligned} \tag{23}$$

The first part of the right hand side of Eq. (23) is given by

$$\begin{aligned}&-K\frac{\partial^2\ln(|\mathbf{M}|)}{\partial\mathrm{vech}(\mathcal{R}e\{\mathbf{M}\})\partial\mathrm{vech}^T(\mathcal{R}e\{\mathbf{M}\})}\\ &\quad=K\mathbf{H}^T(\mathbf{M}^T\otimes\mathbf{M})^{-1}\mathbf{H}\end{aligned} \tag{24}$$

thanks to Appendix A, Eq. (A.3), and thanks to the fact that

$$\frac{\partial\mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})}{\partial\mathrm{vech}(\mathcal{R}e\{\mathbf{M}\})}=\mathbf{0} \tag{25}$$

and

$$\frac{\partial\mathrm{vech}^T(\mathcal{R}e\{\mathbf{M}\})}{\partial\mathrm{vech}(\mathcal{R}e\{\mathbf{M}\})}=\mathbf{I}. \tag{26}$$

Through the remain of the paper, let us set $z=\beta^2\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k$ and $f(z)=\ln((z/\beta^2)^{(\alpha-m)/2}K_{m-\alpha}(\sqrt{z}))$. The second term (inside the sum) of the right-hand side of Eq. (23) is given by

$$\begin{aligned}&\frac{\partial^2\ln((\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k)^{(\alpha-m)/2}K_{m-\alpha}(\beta\sqrt{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k}))}{\partial\mathrm{vech}(\mathcal{R}\rfloor\{\mathbf{M}\})\partial\mathrm{vech}^T(\mathcal{R}\rfloor\{\mathbf{M}\})}\\ &\quad=\frac{\partial^2 f(z)}{\partial\mathrm{vech}(\mathcal{R}e\{\mathbf{M}\})\partial\mathrm{vech}^T(\mathcal{R}e\{\mathbf{M}\})}.\end{aligned} \tag{27}$$

Note that

$$\partial^2 f(z)=\frac{\partial}{\partial z}\left(\frac{\partial f(z)}{\partial z}\partial z\right)\partial z=\frac{\partial^2 f(z)}{\partial z^2}\partial z\,\partial z+\frac{\partial f(z)}{\partial z}\partial^2 z, \tag{28}$$

with (details are given in Appendix B)

$$\begin{aligned}\partial z=&-\beta^2\partial\mathrm{vech}^T(\mathcal{R}\rfloor\{\mathbf{M}\})\mathbf{H}^T(\mathbf{M}^T\otimes\mathbf{M})^{-1}\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\\ &+i\beta^2\partial\mathrm{veck}^T(\mathcal{I}\Updownarrow\{\mathbf{M}\})\mathbf{K}^T(\mathbf{M}^T\otimes\mathbf{M})^{-1}\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\end{aligned} \tag{29}$$

and

$$\begin{aligned}\partial^2 z=&2\beta^2\partial\mathrm{vech}^T(\mathcal{R}e\{\mathbf{M}\})\mathbf{H}^T((\mathbf{M}^{-T}\mathbf{c}_k^*\mathbf{c}_k^T\mathbf{M}^{-T})\\ &\otimes\mathbf{M}^{-1})\mathbf{H}\partial\mathrm{vech}(\mathcal{R}e\{\mathbf{M}\})\\ &+2\beta^2\partial\mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})\mathbf{K}^T((\mathbf{M}^{-T}\mathbf{c}_k^*\mathbf{c}_k^T\mathbf{M}^{-T})\\ &\otimes\mathbf{M}^{-1})\mathbf{K}\partial\mathrm{veck}(\mathcal{I}m\{\mathbf{M}\})\end{aligned} \tag{30}$$

and (details are given in Appendix C)

$$d_1(z)=\frac{\partial f(z)}{\partial z}=-\frac{1}{2\sqrt{z}}\frac{K_{m-\alpha+1}(\sqrt{z})}{K_{m-\alpha}(\sqrt{z})}) \tag{31}$$

and

$$\begin{aligned}d_2(z)=&\frac{\partial^2 f(z)}{\partial z^2}\\ =&\frac{1}{4z}\left(1+\frac{K_{m-\alpha+1}(\sqrt{z})}{K_{m-\alpha}(\sqrt{z})})\left(\frac{2}{\sqrt{z}}-\frac{K_{m-\alpha-1}(\sqrt{z})}{K_{m-\alpha}(\sqrt{z})})\right)\right).\end{aligned} \tag{32}$$

Consequently, the structure of Eq. (28) becomes

$$\begin{aligned}\partial^2 f(z)=&2\beta^2 d_1(z)\partial\mathrm{vech}^T(\mathcal{R}e\{\mathbf{M}\})\mathbf{H}^T\mathbf{P}_1\mathbf{H}\partial\mathrm{vech}(\mathcal{R}e\{\mathbf{M}\})\\ &+2\beta^2 d_1(z)\partial\mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})\mathbf{K}^T\mathbf{P}_1\mathbf{K}\partial\mathrm{veck}(\mathcal{I}m\{\mathbf{M}\})\\ &+\beta^4 d_2(z)\partial\mathrm{vech}^T(\mathcal{R}e\{\mathbf{M}\})\mathbf{H}^T\mathbf{P}_2\mathbf{H}\partial\mathrm{vech}(\mathcal{R}e\{\mathbf{M}\})\\ &+\beta^4 d_2(z)\partial\mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})\mathbf{K}^T\mathbf{P}_2\mathbf{K}\partial\mathrm{veck}(\mathcal{I}m\{\mathbf{M}\}),\end{aligned} \tag{33}$$

where

$$\mathbf{P}_1=(\mathbf{M}^{-T}\mathbf{c}_k^*\mathbf{c}_k^T\mathbf{M}^{-T})\otimes\mathbf{M}^{-1}, \tag{34}$$

$$\mathbf{P}_2=(\mathbf{M}^T\otimes\mathbf{M})^{-1}\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathrm{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)(\mathbf{M}^T\otimes\mathbf{M})^{-1}. \tag{35}$$

Therefore, Eq. (27) is given by

$$\begin{aligned}&\frac{\partial^2 f(z)}{\partial\mathrm{vech}(\mathcal{R}e\{\mathbf{M}\})\partial\mathrm{vech}^T(\mathcal{R}e\{\mathbf{M}\})}\\ &\quad=2\beta^2 d_1(z)\mathbf{H}^T\mathbf{P}_1\mathbf{H}+\beta^4 d_2(z)\mathbf{H}^T\mathbf{P}_2\mathbf{H}\end{aligned} \tag{36}$$

due to Eqs. (25) and (26).

Using Eqs. (23), (24), and (36), $\mathbf{F}_{1,1}$ is given by

$$\begin{aligned}\mathbf{F}_{1,1}=&-K\mathbf{H}^T((\mathbf{M}^T\otimes\mathbf{M})^{-1}+2\beta^2 E[d_1(z)\mathbf{P}_1]\\ &+\beta^4 E[d_2(z)\mathbf{P}_2])\mathbf{H},\end{aligned} \tag{37}$$

where $d_1(z)$ and $d_2(z)$ are defined by Eqs. (31) and (32), respectively.

[1] With $\beta^2=1$, $z=\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k$ which is the term to be derived w.r.t. $\boldsymbol{\theta}$ in the exponential term of the multivariate Gaussian distribution.

The two expectation operators involved in the previous equation can be detailed as follows:

$$E[d_1(z)\mathbf{P}_1] = -\tfrac{1}{2}(\mathbf{M}^{-T}\boldsymbol{\Gamma}\mathbf{M}^{-T}) \otimes \mathbf{M}^{-1} \tag{38}$$

and

$$E[d_2(z)\mathbf{P}_2] = \tfrac{1}{4}(\mathbf{M}^T \otimes \mathbf{M})^{-1}(\boldsymbol{\Psi} + \boldsymbol{\Xi} - \boldsymbol{\Upsilon})(\mathbf{M}^T \otimes \mathbf{M})^{-1}, \tag{39}$$

where

$$\begin{cases} \boldsymbol{\Gamma} = E\left[\dfrac{1}{\sqrt{z}}\dfrac{K_{m-\alpha+1}(\sqrt{z})}{K_{m-\alpha}(\sqrt{z})}\mathbf{c}_k^*\mathbf{c}_k^T\right], \\ \boldsymbol{\Psi} = E\left[\dfrac{1}{z}\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathrm{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)\right], \\ \boldsymbol{\Xi} = E\left[\dfrac{2}{z^{3/2}}\dfrac{K_{m-\alpha+1}(\sqrt{z})}{K_{m-\alpha}(\sqrt{z})}\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathrm{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)\right], \\ \boldsymbol{\Upsilon} = E\left[\dfrac{1}{z}\dfrac{K_{m-\alpha-1}(\sqrt{z})K_{m-\alpha+1}(\sqrt{z})}{K_{m-\alpha}^2(\sqrt{z})}\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathrm{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)\right]. \end{cases} \tag{40}$$

After some calculus detailed in Appendix D, one finds

$$\boldsymbol{\Gamma} = \frac{2}{\beta^2}\mathbf{M}^T. \tag{41}$$

Concerning $\boldsymbol{\Psi}$, one has

$$\boldsymbol{\Psi} = \frac{1}{\beta^2}E\left[\frac{\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathrm{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)}{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k}\right], \tag{42}$$

where the expectation is taken under a complex K-distribution $\mathcal{K}_m(\alpha,(2/\beta)^2,\mathbf{M})$.

Concerning $\boldsymbol{\Xi}$, one has

$$\boldsymbol{\Xi} = \frac{1}{(\alpha-1)\beta^2}E\left[\frac{\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathrm{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)}{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k}\right], \tag{43}$$

where the expectation is taken under a complex K-distribution $\mathcal{K}_m(\alpha-1,(2/\beta)^2,\mathbf{M})$.

Concerning $\boldsymbol{\Upsilon}$, one has

$$\boldsymbol{\Upsilon} = \frac{1}{\beta^2}E\left[\frac{K_{m-\alpha-1}(\sqrt{z})K_{m-\alpha+1}(\sqrt{z})}{K_{m-\alpha}^2(\sqrt{z})}\frac{\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathrm{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)}{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k}\right], \tag{44}$$

where the expectation is taken under a complex K-distribution $\mathcal{K}_m(\alpha,(2/\beta)^2,\mathbf{M})$.

The closed-form expression of $E[\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathrm{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)/\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k]$ under a complex K-distribution is given in Appendix E. One finds

$$\boldsymbol{\Psi} = \frac{8}{\beta^4}\frac{\alpha}{m+1}(\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2})(\mathbf{I}+\mathrm{vec}(\mathbf{I})\mathrm{vec}^T(\mathbf{I}))(\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2}) \tag{45}$$

and

$$\boldsymbol{\Xi} = \frac{8}{\beta^4}\frac{1}{m+1}(\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2})(\mathbf{I}+\mathrm{vec}(\mathbf{I})\mathrm{vec}^T(\mathbf{I}))(\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2}). \tag{46}$$

The structure of $\boldsymbol{\Upsilon}$ is analyzed in Appendix F.

Consequently, Eq. (37) is reduced to

$$\mathbf{F}_{1,1} = K\mathbf{H}^T\left((\mathbf{M}^T\otimes\mathbf{M})^{-1} - \left(\frac{2(\alpha+1)}{m+1} - \frac{\varphi(\alpha,m)}{8}\right)(\mathbf{M}^T\otimes\mathbf{M})^{-1/2}(\mathbf{I}+\mathrm{vec}(\mathbf{I})\mathrm{vec}^T(\mathbf{I}))(\mathbf{M}^T\otimes\mathbf{M})^{-1/2}\right)\mathbf{H}, \tag{47}$$

where $\varphi(\alpha,m)$ is given by Eq. (F.3).

4.2.2. Analysis of $\mathbf{F}_{2,2}$

The analysis of $\mathbf{F}_{2,2}$ is similar to the one used for $\mathbf{F}_{1,1}$. Indeed, one has to calculate

$$\begin{aligned} &\frac{\partial^2 \ln p(\mathbf{c}_1,\ldots,\mathbf{c}_K;\boldsymbol{\theta})}{\partial\mathrm{veck}(\mathcal{I}m\{\mathbf{M}\})\partial\mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})} \\ &= -K\frac{\partial^2\ln(|\mathbf{M}|)}{\partial\mathrm{veck}(\mathcal{I}m\{\mathbf{M}\})\partial\mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})} \\ &\quad + \sum_{k=1}^{K}\frac{\partial^2\ln((\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k)^{(\alpha-m)/2}K_{m-\alpha}(\beta\sqrt{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k}))}{\partial\mathrm{veck}(\mathcal{I}m\{\mathbf{M}\})\partial\mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})}. \end{aligned} \tag{48}$$

Using Eq. (A.3), one has

$$\begin{aligned} &-K\frac{\partial^2\ln(|\mathbf{M}|)}{\partial\mathrm{veck}(\mathcal{I}m\{\mathbf{M}\})\partial\mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})} \\ &= K\mathbf{K}^T(\mathbf{M}^T\otimes\mathbf{M})^{-1}\mathbf{K} \end{aligned} \tag{49}$$

due to Eq. (25) and due to

$$\frac{\partial\mathrm{veck}(\mathcal{I}m\{\mathbf{M}\})}{\partial\mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})} = \mathbf{I}. \tag{50}$$

By using the same notation as for the derivation of $\mathbf{F}_{1,1}$ and by using Eq. (33), one obtains for the second term on the right hand side of Eq. (48) as

$$\begin{aligned} &\frac{\partial^2\ln((\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k)^{(\alpha-m)/2}K_{m-\alpha}(\beta\sqrt{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k}))}{\partial\mathrm{veck}(\mathcal{I}m\{\mathbf{M}\})\partial\mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})} \\ &= 2\beta^2 d_1(z)\mathbf{K}^T\mathbf{P}_1\mathbf{K} + \beta^4 d_2(z)\mathbf{K}^T\mathbf{P}_2\mathbf{K}, \end{aligned} \tag{51}$$

where $\mathbf{P}_1$, $\mathbf{P}_2$, $d_1(z)$ and $d_2(z)$ are defined by Eqs. (34), (35), (31), and (32), respectively. Therefore, the structure of $\mathbf{F}_{2,2}$ is the same as the structure of $\mathbf{F}_{1,1}$ except that one replaces the matrix $\mathbf{H}$ by the matrix $\mathbf{K}$.

4.2.3. Analysis of $\mathbf{F}_{1,2} = \mathbf{F}_{2,1}^T$

Due to the structure of Eq. (A.3) and Eq. (33), and since the derivation is w.r.t. $\partial\mathrm{vech}(\mathcal{R}\lfloor\{\mathbf{M}\})$ and $\partial\mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})$, it is clear that

$$\mathbf{F}_{1,2} = \mathbf{F}_{2,1}^T = \mathbf{0} \tag{52}$$

by using Eq. (25).

This concludes the proof of the CRB derivation.

5. Simulation results

In this section, some simulations are provided in order to illustrate the proposed previous results in terms of consistency, bias, and variance analysis (throughout the CRB). While no mathematical proof is given in other sections, we show the efficiency of the MLE meaning that the CRB is achieved by the variance of the MLE.

The results presented are obtained for complex K-distributed clutter with covariance matrix $\mathbf{M}$ randomly chosen.

All the results are presented different values of parameter α and $\beta = 2\sqrt{\alpha}$ following the scenario of [17]. The size of each vector $\mathbf{c}_k$ is $m = 3$. Remember that the parameter α represents the spikiness of the clutter (when α is high the clutter tends to be Gaussian and, when α is small, the tail of the clutter becomes heavy). The norm used for consistency and bias is the L^2 norm.

5.1. Consistency

Fig. 1 presents results of MLE consistency for 1000 Monte Carlo runs per each value of K. For that purpose, a plot of $D(\hat{\mathbf{M}}, K) = \|\hat{\mathbf{M}} - \mathbf{M}\|$ versus the number K of $\mathbf{c}_k$'s is presented for each estimate. It can be noticed that the above criterion $D(\hat{\mathbf{M}}, K)$ tends to 0 when K tends to ∞ for

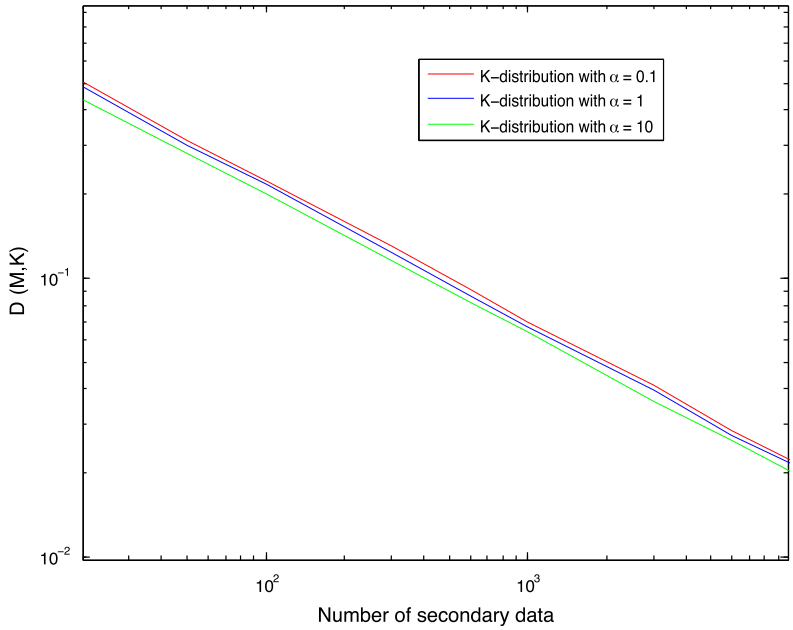


Fig. 1. $D(\hat{\mathbf{M}}, K)$ versus the number of secondary data for different values of α.

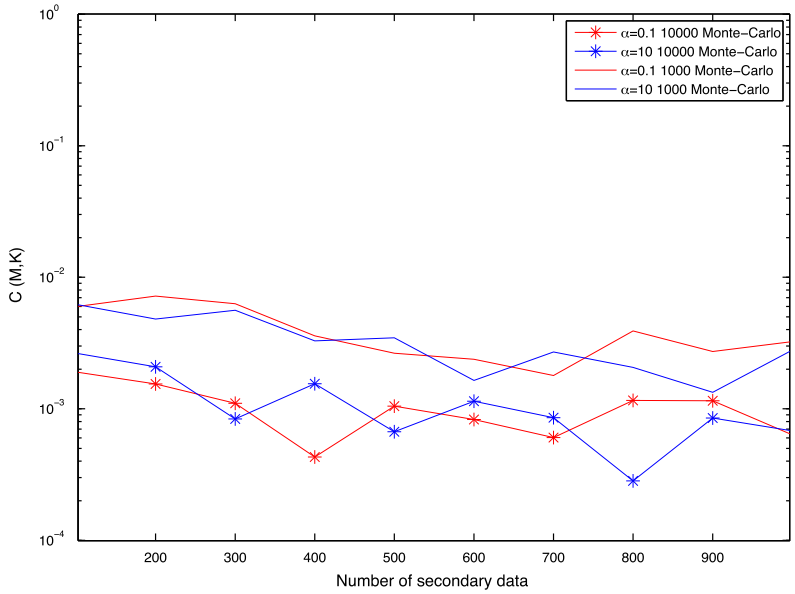


Fig. 2. $C(\hat{\mathbf{M}}, K)$ versus the number of secondary data for different values of α.

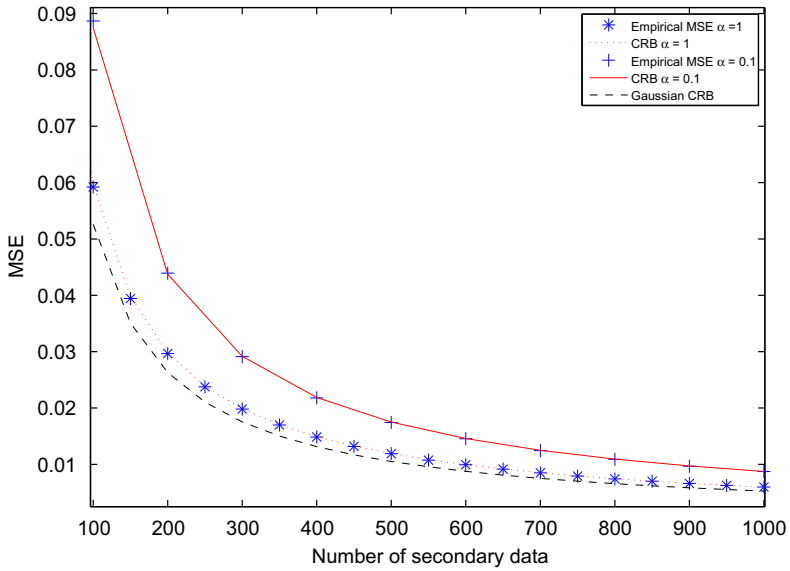


Fig. 3. GCRB, CRB and empirical variance of the MLE for different values of α.

each estimate. Moreover, note that the parameter α has very few influence on the convergence speed which highlights the robustness of the MLE.

5.2. Bias

Fig. 2 shows the bias of each estimate for the different values of α. The number of Monte Carlo runs is given in the legend of the figure. For that purpose, a plot of the criterion $C(\hat{\mathbf{M}}, K) = \|\bar{\hat{\mathbf{M}}} - \mathbf{M}\|$ versus the number K of $\mathbf{c}_k$'s is presented for each estimate. $\bar{\hat{\mathbf{M}}}$ is defined as the empirical mean of the quantities $\hat{\mathbf{M}}(i)$ obtained from I Monte Carlo runs. For each iteration i, a new set of K secondary data $\mathbf{c}_k$ is generated to compute $\hat{\mathbf{M}}(i)$. It can be noticed that, as enlightened by the previous theoretical analysis, the bias of $\hat{\mathbf{M}}$ tends to 0 whatever the value of K. Furthermore, one sees again the weak influence of the parameter α on the unbiasedness of the MLE.

5.3. CRB and MSE

The CRB and empirical variance of the MLE for 10 000 Monte Carlo runs are plotted in Fig. 3. For comparison, we also plot the Gaussian CRB (i.e., when $\tau_k = 1\ \forall k$) given by Eq. (22). Although it is not mathematically proved in this paper, one observes, in Fig. 3, the efficiency of the MLE even for impulsive noise (α small).

6. Conclusion

In this paper, a statistical analysis of the maximum likelihood estimator of the covariance matrix of a complex multivariate K-distributed process has been proposed. More particularly, the consistency and the unbiasedness (at finite number of samples) have been proved. In order to analyze the variance of the estimator, the Cramér–Rao lower bound is derived. The Fisher information matrix in this case is simply the Fisher information matrix of the Gaussian case plus a term depending on the tail of the K-distribution. Simulation results have been proposed to illustrate these theoretical analyses. These results have shown the efficiency of the estimator and the weak influence of the spikiness parameter in terms of consistency and bias.

Appendix A. Derivation of $\partial^2 \ln\{|\mathbf{M}|\}$

To find this term and several other, we will use the following results [33]:

$$\partial \mathrm{Tr}(\mathbf{X}) = \mathrm{Tr}(\partial \mathbf{X}), \tag{A.1a}$$

$$\partial \mathrm{vec}(\mathbf{X}) = \mathrm{vec}(\partial \mathbf{X}), \tag{A.1b}$$

$$\partial \mathbf{A}^{-1} = -\mathbf{A}^{-1} \partial \mathbf{A} \mathbf{A}^{-1}, \tag{A.1c}$$

$$\partial |\mathbf{A}| = |\mathbf{A}| \,\mathrm{Tr}(\mathbf{A}^{-1} \partial \mathbf{A}), \tag{A.1d}$$

$$\partial(\mathbf{A} \quad \mathbf{B}) = \partial(\mathbf{A}) \quad \mathbf{B} + \mathbf{A} \quad \partial(\mathbf{B}) \quad \text{where} \quad = \times \text{ or } \otimes \tag{A.1e}$$

$$\partial \ln(|\mathbf{M}|) = \mathrm{Tr}(\mathbf{M}^{-1} \partial \mathbf{M}), \tag{A.1f}$$

$$\partial(\partial \mathbf{M}) = \mathbf{0}, \tag{A.1g}$$

$$\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA}) \tag{A.1h}$$

$$\mathbf{M}^{-1} \otimes \mathbf{M}^{-1} = (\mathbf{M} \otimes \mathbf{M})^{-1}, \tag{A.1i}$$

$$\mathrm{Tr}(\mathbf{A}^H \mathbf{B}) = \mathrm{vec}^H(\mathbf{A}) \mathrm{vec}(\mathbf{B}), \tag{A.1j}$$

$$\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B}). \quad \text{(A.1k)}$$

By using these properties, one has

$$\partial \ln(|\mathbf{M}|) = \mathrm{Tr}(\mathbf{M}^{-1}(\partial \mathbf{M})) \quad \text{from (A.1f)}$$

$$\begin{aligned}\partial^2 \ln(|\mathbf{M}|) &= -\mathrm{Tr}(\mathbf{M}^{-1}(\partial \mathbf{M})\mathbf{M}^{-1}\partial \mathbf{M}) \\ &\quad \text{from (A.1a)(A.1e)(A.1g)(A.1c)} \\ &= -\mathrm{vec}^H(\mathbf{M}^{-1}(\partial \mathbf{M})\mathbf{M}^{-1})\mathrm{vec}(\partial \mathbf{M}) \quad \text{from (A.1j)} \\ &= -\mathrm{vec}^H(\partial \mathbf{M})(\mathbf{M}^{-T} \otimes \mathbf{M}^{-1})^H \mathrm{vec}(\partial \mathbf{M}) \quad \text{from (A.1k)} \\ &= -\partial \mathrm{vec}^H(\mathbf{M})(\mathbf{M}^T \otimes \mathbf{M})^{-1}\partial \mathrm{vec}(\mathbf{M}) \\ &\quad \text{from (A.1b)(A.1i).}\end{aligned}$$

By letting $\mathbf{M} = \mathcal{R}e\{\mathbf{M}\} + i\mathcal{I}m\{\mathbf{M}\}$ in Eq. (A.3), one has

$$\begin{aligned}\partial^2 \ln(|\mathbf{M}|) &= -\partial \mathrm{vec}^H(\mathcal{R}e\{\mathbf{M}\} + i\mathcal{I}m\{\mathbf{M}\}) \\ &\quad \times(\mathbf{M}^T \otimes \mathbf{M})^{-1}\partial \mathrm{vec}(\mathcal{R}e\{\mathbf{M}\} + i\mathcal{I}m\{\mathbf{M}\}) \\ &= -\partial \mathrm{vec}^H(\mathcal{R}e\{\mathbf{M}\})(\mathbf{M}^T \otimes \mathbf{M})^{-1}\partial \mathrm{vec}(\mathcal{R}e\{\mathbf{M}\}) \\ &\quad -\partial \mathrm{vec}^H(i\mathcal{I}m\{\mathbf{M}\})(\mathbf{M}^T \otimes \mathbf{M})^{-1}\partial \mathrm{vec}(\mathcal{R}e\{\mathbf{M}\}) \\ &\quad -\partial \mathrm{vec}^H(\mathcal{R}e\{\mathbf{M}\})(\mathbf{M}^T \otimes \mathbf{M})^{-1}\partial \mathrm{vec}(i\mathcal{I}m\{\mathbf{M}\}) \\ &\quad -\partial \mathrm{vec}^H(i\mathcal{I}m\{\mathbf{M}\})(\mathbf{M}^T \otimes \mathbf{M})^{-1}\partial \mathrm{vec}(i\mathcal{I}m\{\mathbf{M}\}).\end{aligned} \quad \text{(A.2)}$$

Since $\mathrm{vec}^H(i\mathcal{I}m\{\mathbf{M}\}) = -i\mathrm{vec}^T(\mathcal{I}m\{\mathbf{M}\})$ and $\mathrm{vec}^H(\mathcal{R}e\{\mathbf{M}\}) = \mathrm{vec}^T(\mathcal{R}e\{\mathbf{M}\})$, $\partial^2 \ln(|\mathbf{M}|)$ is reduced to

$$\begin{aligned}\partial^2 \ln(|\mathbf{M}|) &= -\partial \mathrm{vec}^T(\mathcal{R}e\{\mathbf{M}\})(\mathbf{M}^T \otimes \mathbf{M})^{-1}\partial \mathrm{vec}(\mathcal{R}e\{\mathbf{M}\}) \\ &\quad -\partial \mathrm{vec}^T(\mathcal{I}m\{\mathbf{M}\})(\mathbf{M}^T \otimes \mathbf{M})^{-1}\partial \mathrm{vec}(\mathcal{I}m\{\mathbf{M}\}) \\ &= -\partial \mathrm{vech}^T(\mathcal{R}e\{\mathbf{M}\})\mathbf{H}^T \\ &\quad \times(\mathbf{M}^T \otimes \mathbf{M})^{-1}\mathbf{H}\partial \mathrm{vech}(\mathcal{R}e\{\mathbf{M}\}) \\ &\quad -\partial \mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})\mathbf{K}^T(\mathbf{M}^T \otimes \mathbf{M})^{-1}\mathbf{K}\partial \mathrm{veck}(\mathcal{I}m\{\mathbf{M}\}),\end{aligned} \quad \text{(A.3)}$$

where the matrix $\mathbf{H}$ and $\mathbf{K}$ are constant transformation matrices filled with ones and zeros such that $\mathrm{vec}(\mathbf{A}) = \mathbf{H}\mathrm{vech}(\mathbf{A})$ and $\mathrm{vec}(\mathbf{A}) = \mathbf{K}\mathrm{veck}(\mathbf{A})$.

Appendix B. Derivation of ∂z and $\partial^2 z$

By using the properties from Eq. (A.1a) to (A.1k), one has for ∂z

$$\begin{aligned}\partial z &= \beta^2 \partial(\mathbf{c}_k^H \mathbf{M}^{-1}\mathbf{c}_k) = \beta^2 \partial \mathrm{Tr}(\mathbf{c}_k^H \mathbf{M}^{-1}\mathbf{c}_k) \\ &= \beta^2 \mathrm{Tr}(\mathbf{c}_k^H \partial(\mathbf{M}^{-1})\mathbf{c}_k) \quad \text{from (A.1a)} \\ &= -\beta^2 \mathrm{Tr}(\mathbf{c}_k^H \mathbf{M}^{-1}\partial \mathbf{M}\mathbf{M}^{-1}\mathbf{c}_k) \quad \text{from (A.1c)} \\ &= -\beta^2 \mathrm{Tr}(\partial \mathbf{M}\mathbf{M}^{-1}\mathbf{c}_k\mathbf{c}_k^H\mathbf{M}^{-1}) \quad \text{from (A.1h)} \\ &= -\beta^2 \mathrm{vec}^H(\partial \mathbf{M})\mathrm{vec}(\mathbf{M}^{-1}\mathbf{c}_k\mathbf{c}_k^H\mathbf{M}^{-1}) \quad \text{from (A.1j)} \\ &= -\beta^2 \partial \mathrm{vec}^H(\mathbf{M}) \\ &\quad \times(\mathbf{M}^T \otimes \mathbf{M})^{-1}\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H) \quad \text{from (A.1b)(A.1k)(A.1i).}\end{aligned}$$

By letting $\mathbf{M} = \mathcal{R}e\{\mathbf{M}\} + i\mathcal{I}m\{\mathbf{M}\}$, one obtains

$$\begin{aligned}\partial z &= -\beta^2 \partial \mathrm{vec}^T(\mathcal{R}e\{\mathbf{M}\})(\mathbf{M}^T \otimes \mathbf{M})^{-1}\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H) \\ &\quad + i\beta^2 \partial \mathrm{vec}^T(\mathcal{I}m\{\mathbf{M}\})(\mathbf{M}^T \otimes \mathbf{M})^{-1}\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H) \\ &= -\beta^2 \partial \mathrm{vech}^T(\mathcal{R}e\{\mathbf{M}\})\mathbf{H}^T(\mathbf{M}^T \otimes \mathbf{M})^{-1}\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H) \\ &\quad + i\beta^2 \partial \mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})\mathbf{K}^T(\mathbf{M}^T \otimes \mathbf{M})^{-1}\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H).\end{aligned} \quad \text{(B.1)}$$

Concerning $\partial^2 z$, one has

$$\partial z = -\beta^2 \mathrm{Tr}(\mathbf{c}_k^H \mathbf{M}^{-1}\partial \mathbf{M}\mathbf{M}^{-1}\mathbf{c}_k) \quad \text{(B.2)}$$

$$\begin{aligned}\partial^2 z &= -\beta^2 \mathrm{Tr}(\mathbf{c}_k^H \partial(\mathbf{M}^{-1}\partial \mathbf{M}\mathbf{M}^{-1})\mathbf{c}_k) \quad \text{from (A.1a)} \\ &= -\beta^2 \mathrm{Tr}(\mathbf{c}_k^H \partial(\mathbf{M}^{-1})\partial \mathbf{M}\mathbf{M}^{-1}\mathbf{c}_k \\ &\quad + \mathbf{c}_k^H \mathbf{M}^{-1}\partial \mathbf{M}\partial(\mathbf{M}^{-1})\mathbf{c}_k) \quad \text{from (A.1e)(A.1g)} \\ &= 2\beta^2 \mathrm{Tr}(\partial \mathbf{M}\mathbf{M}^{-1}\partial \mathbf{M}\mathbf{M}^{-1}\mathbf{c}_k\mathbf{c}_k^H\mathbf{M}^{-1}) \quad \text{from (A.1c)(A.1h)} \\ &= 2\beta^2 \partial \mathrm{vec}^H(\mathbf{M})((\mathbf{M}^{-T}\mathbf{c}_k^*\mathbf{c}_k^T\mathbf{M}^{-T}) \\ &\quad \otimes \mathbf{M}^{-1})\partial \mathrm{vec}(\mathbf{M}) \quad \text{from (A.1j)(A.1k)(A.1b)}\end{aligned}$$

By letting $\mathbf{M} = \mathcal{R}e\{\mathbf{M}\} + i\mathcal{I}m\{\mathbf{M}\}$, one obtains

$$\begin{aligned}\partial^2 z &= 2\beta^2 \partial \mathrm{vech}^T(\mathcal{R}e\{\mathbf{M}\})\mathbf{H}^T((\mathbf{M}^{-T}\mathbf{c}_k^*\mathbf{c}_k^T\mathbf{M}^{-T})\mathbf{M}^{-1})\mathbf{H}\partial \mathrm{vech}(\mathcal{R}e\{\mathbf{M}\}) \\ &\quad +2\beta^2 \partial \mathrm{veck}^T(\mathcal{I}m\{\mathbf{M}\})\mathbf{K}^T((\mathbf{M}^{-T}\mathbf{c}_k^*\mathbf{c}_k^T\mathbf{M}^{-T}) \\ &\quad \otimes \mathbf{M}^{-1})\mathbf{K}\partial \mathrm{veck}(\mathcal{I}m\{\mathbf{M}\}).\end{aligned} \quad \text{(B.3)}$$

Appendix C. Derivation of $\partial f(z)/\partial z$ and $\partial^2 f(z)/\partial z^2$

Concerning the first derivative of $f(z)$, one has

$$\begin{aligned}\frac{\partial f(z)}{\partial z} &= \frac{\partial}{\partial z}\ln((z/\beta^2)^{(\alpha-m)/2}K_{m-\alpha}(\sqrt{z})) \\ &= \frac{\alpha - m}{2z} + \frac{1}{K_{m-\alpha}(\sqrt{z})}\frac{\partial K_{m-\alpha}(\sqrt{z})}{\partial z}.\end{aligned} \quad \text{(C.1)}$$

Since $\partial K_\nu(y)/\partial y = -K_{\nu+1}(y) + (\nu/y)K_\nu(y)$ [34]. It follows that

$$\frac{\partial K_{m-\alpha}(\sqrt{z})}{\partial z} = -\frac{1}{2\sqrt{z}}K_{m-\alpha+1}(\sqrt{z}) + \frac{m-\alpha}{2z}K_{m-\alpha}(\sqrt{z}). \quad \text{(C.2)}$$

Plugging Eq. (C.2) in Eq. (C.1), one obtains

$$\frac{\partial f(z)}{\partial z} = -\frac{1}{2\sqrt{z}}\frac{K_{m-\alpha+1}(\sqrt{z})}{K_{m-\alpha}(\sqrt{z})}. \quad \text{(C.3)}$$

Concerning the second derivative of $f(z)$, one has

$$\begin{aligned}\frac{\partial^2 f(z)}{\partial z^2} &= -\frac{1}{2}\frac{\partial}{\partial z}\left(\frac{1}{\sqrt{z}}\frac{K_{m-\alpha+1}(\sqrt{z})}{K_{m-\alpha}(\sqrt{z}}\right)\Big) \\ &= \frac{1}{4z^{3/2}}\frac{K_{m-\alpha+1}(\sqrt{z})}{K_{m-\alpha}(\sqrt{z})} - \frac{1}{4z}\left(\frac{\partial}{\partial y}\frac{K_{m-\alpha+1}(y)}{K_{m-\alpha}(y)}\Big|_{y=\sqrt{z}}\right).\end{aligned} \quad \text{(C.4)}$$

Since $\partial K_\nu(y)/\partial y = -K_{\nu-1}(y) - (\nu/y)K_\nu(y)$ [34]. It follows that

$$\begin{aligned}&\frac{\partial}{\partial y}\frac{K_{m-\alpha+1}(y)}{K_{m-\alpha}(y)}\Big|_{y=\sqrt{z}} \\ &= \frac{-(K_{m-\alpha}(\sqrt{z}) + \dfrac{m-\alpha+1}{\sqrt{z}}K_{m-\alpha+1}(\sqrt{z}))K_{m-\alpha}(\sqrt{z})}{K_{m-\alpha}^2(\sqrt{z})} \\ &+ \frac{K_{m-\alpha+1}(\sqrt{z})(K_{m-\alpha-1}(\sqrt{z}) + \dfrac{m-\alpha}{\sqrt{z}}K_{m-\alpha}(\sqrt{z}))}{K_{m-\alpha}^2(\sqrt{z})}.\end{aligned} \quad \text{(C.5)}$$

Plugging Eq. (C.5) in Eq. (C.4), one obtains

$$\frac{\partial^2 f(z)}{\partial z^2} = \frac{1}{4z} + \frac{1}{2z^{3/2}}\frac{K_{m-\alpha+1}(\sqrt{z})}{K_{m-\alpha}(\sqrt{z})} - \frac{1}{4z}\frac{K_{m-\alpha+1}(\sqrt{z})K_{m-\alpha-1}(\sqrt{z})}{K_{m-\alpha}^2(\sqrt{z})}. \quad \text{(C.6)}$$

Appendix D. Derivation of matrix $\mathbf{\Gamma}$

The matrix $\mathbf{\Gamma}$ is given by

$$\mathbf{\Gamma} = E\left[\frac{1}{\sqrt{z}}\frac{K_{m-\alpha+1}(\sqrt{z})}{K_{m-\alpha}(\sqrt{z})}\mathbf{c}_k^*\mathbf{c}_k^T\right] = \frac{\Gamma(\alpha-1)}{2\Gamma(\alpha)}E^T[\mathbf{c}_k\mathbf{c}_k^H], \tag{D.1}$$

where the last expectation is taken under the distribution

$$p(\mathbf{c}_k) = \frac{\beta^{\alpha+m-1}|\mathbf{M}|^{-1}}{2^{\alpha+m-2}\pi^m\Gamma(\alpha-1)}(\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k)^{(\alpha-m-1)/2} \times K_{m-\alpha+1}(\beta\sqrt{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k}), \tag{D.2}$$

which is a complex K-distribution $\mathcal{K}_m(\alpha-1,(2/\beta)^2,\mathbf{M})$. Then

$$\mathbf{\Gamma} = \frac{1}{2(\alpha-1)}E^T[\mathbf{c}_k\mathbf{c}_k^H] = \frac{1}{2(\alpha-1)}E^T[\tau_k\mathbf{x}_k\mathbf{x}_k^H] = \frac{1}{2(\alpha-1)}E[\tau_k]E^T[\mathbf{x}_k\mathbf{x}_k^H], \tag{D.3}$$

where $E[\tau_k]=(\alpha-1)(2/\beta)^2$ since τ_k follows a Gamma distribution $\mathcal{G}(\alpha-1,(2/\beta)^2)$ and $E^T[\mathbf{x}_k\mathbf{x}_k^H]=\mathbf{M}^T$ since $\mathbf{x}_k$ is a complex normal random vector (independent of τ_k) with zero mean and covariance matrix $\mathbf{M}$. Consequently, $\mathbf{\Gamma}=(2/\beta^2)\mathbf{M}^T$.

Appendix E. Derivation of $E[\mathbf{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathbf{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)/\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k]$

In this appendix, we derive the expression of $E[\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathrm{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)/\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k]$ where $\mathbf{c}_k\sim\mathcal{K}_m(\alpha,(2/\beta)^2,\mathbf{M})$. The case where $\mathbf{c}_k\sim\mathcal{K}_m(\alpha-1,(2/\beta)^2,\mathbf{M})$ will be, of course, straightforward. Let us set the following change of variable: $\mathbf{c}_k=\mathbf{M}^{1/2}\mathbf{y}_k$. One obtains from Eq. (A.1k)

$$E\left[\frac{\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathrm{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)}{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k}\right] = (\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2})$$

$$E\left[\frac{\mathrm{vec}(\mathbf{y}_k\mathbf{y}_k^H)\mathrm{vec}^H(\mathbf{y}_k\mathbf{y}_k^H)}{\mathbf{y}_k^H\mathbf{y}_k}\right](\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2}), \tag{E.1}$$

where $\mathbf{y}_k\sim\mathcal{K}_m(\alpha,(2/\beta)^2,\mathbf{I})$. Since $\mathbf{y}_k=\sqrt{\tau_k}\mathbf{x}_k$, where $\tau_k\sim\mathcal{G}(\alpha,(2/\beta)^2)$ is independent of $\mathbf{x}_k\sim\mathcal{CN}(\mathbf{0},\mathbf{I})$, one has

$$E\left[\frac{\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathrm{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)}{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k}\right] = (\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2})E[\tau_k]\mathbf{R} \times(\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2}), \tag{E.2}$$

where $E[\tau_k]=\alpha(2/\beta)^2$ and where

$$\mathbf{R} = E\left[\frac{\mathrm{vec}(\mathbf{x}_k\mathbf{x}_k^H)\mathrm{vec}^H(\mathbf{x}_k\mathbf{x}_k^H)}{\mathbf{x}_k^H\mathbf{x}_k}\right]. \tag{E.3}$$

Let us set $x_k^{(j)}=\sqrt{\rho_j^2}\exp(i\theta_j)$ for $j=1,\ldots,m$. Note that $\rho_j^2\sim\chi^2(2)$ is independent of $\theta_j\sim\mathcal{U}_{[0,2\pi]}$. Consequently, the elements of the matrix $\mathbf{R}$ can be rewritten as

$$R_{k,l} = E\left[\frac{\sqrt{\rho_p^2\rho_q^2\rho_{p'}^2\rho_{q'}^2}}{\sum_{j=1}^m\rho_j^2}\right]E[\exp(i(\theta_p-\theta_{p'}+\theta_{q'}-\theta_q))] \tag{E.4}$$

since

$$\mathbf{x}_k^H\mathbf{x}_k=\sum_{j=1}^m\rho_j^2 \text{ and } [\mathrm{vec}(\mathbf{x}_k\mathbf{x}_k^H)]_n = \sqrt{\rho_p^2\rho_{p'}^2}\exp(i(\theta_p-\theta_{p'})). \tag{E.5}$$

Note that $E[\exp(i(\theta_p-\theta_{p'}+\theta_{q'}-\theta_q))]\neq 0$ if and only if

(1) $p=p'=q=q'$, i.e., $l=k=p+m(p-1)$,
(2) $p=p'$, $q=q'$ and $p\neq q$, i.e., $l=k=p+m(q-1)$,
(3) $p=q$, $p'=q'$ and $p\neq p'$, i.e., $l=p'+m(p'-1)$ and $k=p+m(p-1)$.

Consequently, the non-zero elements of $R_{k,l}$ are given by

(1) $R_{p+m(p-1),p+m(p-1)}=E[(\rho_p^2)^2/\sum_{j=1}^m\rho_j^2]=4/(m+1)$,
(2) $R_{p+m(q-1),p+m(q-1)}=E[\rho_p^2\rho_{p'}^2/\sum_{j=1}^m\rho_j^2]=2/(m+1)$,
(3) $R_{p+m(p-1),p'+m(p'-1)}=E[\rho_p^2\rho_q^2/\sum_{j=1}^m\rho_j^2]=2/(m+1)$,

and the matrix $E[\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathrm{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)/\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k]$ can be written as

$$E\left[\frac{\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathrm{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)}{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k}\right] = \frac{2\alpha}{m+1}\left(\frac{2}{\beta}\right)^2 \times(\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2}) \times(\mathbf{I}+\mathrm{vec}(\mathbf{I})\mathrm{vec}^T(\mathbf{I}))(\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2}). \tag{E.6}$$

In the same way, the expression of $E[\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathrm{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)/\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k]$ where $\mathbf{c}_k\sim\mathcal{K}_m(\alpha-1,(2/\beta)^2,\mathbf{M})$ is given by

$$E\left[\frac{\mathrm{vec}(\mathbf{c}_k\mathbf{c}_k^H)\mathrm{vec}^H(\mathbf{c}_k\mathbf{c}_k^H)}{\mathbf{c}_k^H\mathbf{M}^{-1}\mathbf{c}_k}\right] = \frac{2(\alpha-1)}{m+1}\left(\frac{2}{\beta}\right)^2 \times(\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2}) \times(\mathbf{I}+\mathrm{vec}(\mathbf{I})\mathrm{vec}^T(\mathbf{I}))(\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2}). \tag{E.7}$$

Appendix F. Analysis of $\mathbf{\Upsilon}$

Let us set the following change of variable: $\mathbf{c}_k=\mathbf{M}^{1/2}\sqrt{\tau_k}\mathbf{x}_k$, where $\tau_k\sim\mathcal{G}(\alpha,(2/\beta)^2)$ is independent of $\mathbf{x}_k\sim\mathcal{CN}(\mathbf{0},\mathbf{I})$, one has

$$\mathbf{\Upsilon} = \frac{1}{\beta^2}(\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2})\tilde{\mathbf{\Upsilon}}(\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2}),$$

where

$$\tilde{\mathbf{\Upsilon}} = E\left[\frac{\tau_k K_{m-\alpha-1}(\beta\sqrt{\tau_k\mathbf{x}_k^H\mathbf{x}_k})K_{m-\alpha+1}(\beta\sqrt{\tau_k\mathbf{x}_k^H\mathbf{x}_k})}{K_{m-\alpha}^2(\beta\sqrt{\tau_k\mathbf{x}_k^H\mathbf{x}_k})}\frac{\mathrm{vec}(\mathbf{x}_k\mathbf{x}_k^H)\mathrm{vec}^H(\mathbf{x}_k\mathbf{x}_k^H)}{\mathbf{x}_k^H\mathbf{x}_k}\right]. \tag{F.1}$$

Let us set $x_k^{(j)}=\sqrt{\rho_j^2}\exp(i\theta_j)$ for $j=1,\ldots,m$. $\rho_j^2\sim\chi^2(2)$ is independent of $\theta_j\sim\mathcal{U}_{[0,2\pi]}$. Consequently, due to Eq. (E.5), the elements of the matrix $\tilde{\mathbf{\Upsilon}}$ can be rewritten as

$$\tilde{\Upsilon}_{k,l} = E\left[\tau_k\frac{K_{m-\alpha-1}\left(\beta\sqrt{\tau_k\sum_{j=1}^m\rho_j^2}\right)K_{m-\alpha+1}\left(\beta\sqrt{\tau_k\sum_{j=1}^m\rho_j^2}\right)}{K_{m-\alpha}^2\left(\beta\sqrt{\tau_k\sum_{j=1}^m\rho_j^2}\right)}\frac{\sqrt{\rho_p^2\rho_q^2\rho_{p'}^2\rho_{q'}^2}}{\sum_{j=1}^m\rho_j^2}\right] \times E[\exp(i(\theta_p-\theta_{p'}+\theta_{q'}-\theta_q))]. \tag{F.2}$$

As before, $E[\exp(i(\theta_p-\theta_{p'}+\theta_{q'}-\theta_q))]\neq 0$ if and only if

(1) $p=p'=q=q'$, i.e., $l=k=p+m(p-1)$,
(2) $p=p'$, $q=q'$ and $p\neq q$, i.e., $l=k=p+m(q-1)$,
(3) $p=q$, $p'=q'$ and $p\neq p'$, i.e., $l=p'+m(p'-1)$ and $k=p+m(p-1)$.

Consequently, the non-zero elements of $\tilde{\Upsilon}_{k,l}$ are given by

$$(1)\quad \tilde{\Upsilon}_{p+m(p-1),p+m(p-1)} = E\left[\tau_{k'}\frac{K_{m-\alpha-1}\left(\beta\sqrt{\tau_{k'}\sum_{j=1}^m\rho_j^2}\right)K_{m-\alpha+1}\left(\beta\sqrt{\tau_{k'}\sum_{j=1}^m\rho_j^2}\right)}{K_{m-\alpha}^2\left(\beta\sqrt{\tau_{k'}\sum_{j=1}^m\rho_j^2}\right)}\frac{(\rho_p^2)^2}{\sum_{j=1}^m\rho_j^2}\right],$$

$$(2)\quad \tilde{\Upsilon}_{p+m(q-1),p+m(q-1)} = E\left[\tau_{k'}\frac{K_{m-\alpha-1}\left(\beta\sqrt{\tau_{k'}\sum_{j=1}^m\rho_j^2}\right)K_{m-\alpha+1}\left(\beta\sqrt{\tau_{k'}\sum_{j=1}^m\rho_j^2}\right)}{K_{m-\alpha}^2\left(\beta\sqrt{\tau_{k'}\sum_{j=1}^m\rho_j^2}\right)}\frac{\rho_p^2\rho_{p'}^2}{\sum_{j=1}^m\rho_j^2}\right],$$

$$(3)\quad \tilde{\Upsilon}_{p+m(p-1),p'+m(p'-1)} = E\left[\tau_{k'}\frac{K_{m-\alpha-1}\left(\beta\sqrt{\tau_{k'}\sum_{j=1}^m\rho_j^2}\right)K_{m-\alpha+1}\left(\beta\sqrt{\tau_{k'}\sum_{j=1}^m\rho_j^2}\right)}{K_{m-\alpha}^2\left(\beta\sqrt{\tau_{k'}\sum_{j=1}^m\rho_j^2}\right)}\frac{\rho_p^2\rho_q^2}{\sum_{j=1}^m\rho_j^2}\right].$$

Note that $\tilde{\Upsilon}_{p+m(q-1),p+m(q-1)}=\tilde{\Upsilon}_{p+m(p-1),p'+m(p'-1)}=\frac{1}{2}\tilde{\Upsilon}_{p+m(p-1),p+m(p-1)}\ \forall p\ \forall p'\ \forall q\ \forall q'$. Consequently, only

$$\varphi(\alpha,m) = \beta^2 E\left[\tau_{k'}\frac{K_{m-\alpha-1}\left(\beta\sqrt{\tau_{k'}\sum_{j=1}^m\rho_j^2}\right)K_{m-\alpha+1}(\beta\sqrt{\tau_{k'}\sum_{j=1}^m\rho_j^2})}{K_{m-\alpha}^2(\beta\sqrt{\tau_{k'}\sum_{j=1}^m\rho_j^2})}\frac{(\rho_p^2)^2}{\sum_{j=1}^m\rho_j^2}\right] \tag{F.3}$$

has to be computed and the matrix Υ can be written as

$$\Upsilon=\frac{\varphi(\alpha,m)}{2\beta^4}(\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2})(\mathbf{I}+\mathrm{vec}(\mathbf{I})\mathrm{vec}^T(\mathbf{I}))\times(\mathbf{M}^{T/2}\otimes\mathbf{M}^{1/2}). \tag{F.4}$$

Note that $\varphi(\alpha,m)$ is independent of β since, $\beta^2\tau_{k'}\sim\mathcal{G}(\alpha,4)$.

References

[1] S.M. Kay, Fundamentals of Statistical Signal Processing—Detection Theory, vol. 2, Prentice-Hall, PTR, Englewood Cliffs, New Jersey, 1998.
[2] H.L. Van Trees, Detection Estimation and Modulation Theory, Part I, Wiley, New York, 1968.
[3] L.L. Scharf, D.W. Lytle, Signal detection in Gaussian noise of unknown level: an invariance application, IEEE Trans. Inf. Theory 17 (1971) 404–411.
[4] S. Haykin, Array Signal Processing, Prentice-Hall Signal Processing Series, Prentice-Hall, Englewood Cliffs, New Jersey, 1985.
[5] H.L. Van Trees, Detection, Estimation and Modulation Theory, Part IV: Optimum Array Processing, Wiley, New York, 2002.
[6] J.B. Billingsley, Ground clutter measurements for surface-sited radar, Technical Report 780, MIT, February 1993.
[7] A. Farina, F. Gini, M.V. Greco, L. Verrazzani, High resolution sea clutter data: a statistical analysis of recorded live data, IEE Proc. Part F 144 (3) (1997) 121–130.
[8] E. Conte, M. Longo, Characterization of radar clutter as a spherically invariant random process, IEE Proc. Part F 134 (2) (1987) 191–197.
[9] K.J. Sangston, K.R. Gerlach, Coherent detection of radar targets in a non-Gaussian background, IEEE Trans. Aerosp. Electron. Syst. 30 (1994) 330–340.
[10] F. Gini, Sub-optimum coherent radar detection in a mixture of K-distributed and Gaussian clutter, IEE Proc. Radar Sonar Navigation 144 (1) (1997) 39–48.
[11] F. Gini, M.V. Greco, M. Diani, L. Verrazzani, Performance analysis of two adaptive radar detectors against non-Gaussian real sea clutter data, IEEE Trans. Aerosp. Electron. Syst. 36 (4) (2000) 1429–1439.
[12] B. Picinbono, Spherically invariant and compound Gaussian stochastic processes, IEEE Trans. Inf. Theory (1970) 77–79.
[13] K. Yao, A representation theorem and its applications to spherically invariant random processes, IEEE Trans. Inf. Theory 19 (5) (1973) 600–608.
[14] M. Rangaswamy, D.D. Weiner, A. Ozturk, Non-Gaussian vector identification using spherically invariant random processes, IEEE Trans. Aerosp. Electron. Syst. 29 (1) (1993) 111–124.
[15] M. Rangaswamy, Statistical analysis of the nonhomogeneity detector for non-Gaussian interference backgrounds, IEEE Trans. Signal Process. 53 (6) (2005) 2101–2111.
[16] E. Jay, J.-P. Ovarlez, D. Declercq, P. Duvaut, Bord: Bayesian optimum radar detector, Signal Process. 83 (6) (2003) 1151–1162.
[17] F. Gini, M.V. Greco, Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter, Signal Process. 82 (12) (2002) 1847–1859 (special section on SP with Heavy Tailed Distributions).
[18] E. Conte, A. DeMaio, G. Ricci, Recursive estimation of the covariance matrix of a compound-Gaussian process and its application to adaptive CFAR detection, IEEE Trans. Signal Process. 50 (8) (2002) 1908–1915.
[19] S. Watts, Radar detection prediction in sea clutter using the compound K-distribution model, IEE Proc. Part F 132 (7) (1985) 613–620.
[20] T. Nohara, S. Haykin, Canada east coast trials and the K-distribution, IEE Proc. Part F 138 (2) (1991) 82–88.
[21] S. Watts, Radar performance in K-distributed sea clutter, in: RTO SET Symposium on "Low Grazing Angle Clutter: Its Characterisation, Measurement and Application", RTO MP-60, 2000, pp. 372–379.
[22] I. Antipov, Statistical analysis of northern Australian coastline sea clutter data, Technical Report, Defence Science and Technology Organisation, DSTO, Available at: ⟨http://www.dsto.defence.gov.au/corporate/reports/DSTO-TR-1236.pdf⟩, November 2001.
[23] D. Blacknell, Comparison of parameter estimators for K-distribution, IEE Proc. Radar Sonar Navigation 141 (1) (1994) 45–52.
[24] W.J.J. Roberts, S. Furui, Maximum likelihood estimation of K-distribution parameters via the expectation-maximization algorithm, IEEE Trans. Signal Process. 48 (12) (2000) 3303–3306.
[25] J. Wang, A. Dogandzic, A. Nehorai, Maximum likelihood estimation of compound-Gaussian clutter and target parameters, IEEE Trans. Signal Process. 54 (10) (2006) 3884–3898.
[26] D.R. Raghavan, R.S., N.B. Pulsone, A generalization of the adaptive matched filter receiver for array detection in a class of a non-gaussian interference, in: Proceedings of the ASAP Workshop, Lexinton, MA, 1996, pp. 499–517.
[27] N.B. Pulsone, Adaptive signal detection in non-gaussian interference, Ph.D. Thesis, Northeastern University, Boston, MA, May 1997.
[28] Y. Chitour, F. Pascal, Exact maximum likelihood estimates for SIRV covariance matrix: existence and algorithm analysis, IEEE Trans. Signal Process. 56 (10) (2008) 4563–4573.
[29] F. Gini, A radar application of a modified Cramér–Rao bound: parameter estimation in non-Gaussian clutter, IEEE Trans. Signal Process. 46 (7) (1998) 1945–1953.
[30] F. Pascal, P. Forster, J.-P. Ovarlez, P. Larzabal, Performance analysis of covariance matrix estimates in impulsive noise, IEEE Trans. Signal Process. 56 (6) (2008) 2206–2217.
[31] M. Abramowitz, I. Stegun, Handbook of Mathematical Functions, National Bureau of Standard, AMS 55, 1964.
[32] A.W. van der Vaart, Asymptotic Statistics, Cambridge University Press, Cambridge, 1998.
[33] K.B. Petersen, M.S. Pedersen, The matrix cookbook, Available at: ⟨http://matrixcookbook.com⟩, February 2008.
[34] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables, Dover, New York, 1972.

Annexe E

Barankin-type lower bound on multiple change-point estimation

IEEE Transactions on Signal Processing, Volume : 58, Issue : 11, Nov. 2010, pp. 5534-5549

5534 IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 58, NO. 11, NOVEMBER 2010

Barankin-Type Lower Bound on Multiple Change-Point Estimation

Patricio S. La Rosa, *Member, IEEE*, Alexandre Renaux, *Member, IEEE*, Carlos H. Muravchik, *Senior Member, IEEE*, and Arye Nehorai, *Fellow, IEEE*

***Abstract*—We compute lower bounds on the mean-square error of multiple change-point estimation. In this context, the parameters are discrete and the Cramér-Rao bound is not applicable. Consequently, we focus on computing the Barankin bound (BB), the greatest lower bound on the covariance of any unbiased estimator, which is still valid for discrete parameters. In particular, we compute the multi-parameter version of the Hammersley–Chapman–Robbins, which is a Barankin-type lower bound. We first give the structure of the so-called Barankin information matrix (BIM) and derive a simplified form of the BB. We show that the particular case of two change points is fundamental to finding the inverse of this matrix. Several closed-form expressions of the elements of BIM are given for changes in the parameters of Gaussian and Poisson distributions. The computation of the BB requires finding the supremum of a finite set of positive definite matrices with respect to the *Loewner partial ordering*. Although each matrix in this set of candidates is a lower bound on the covariance matrix of the estimator, the existence of a unique supremum w.r.t. to this set, i.e., the tightest bound, might not be guaranteed. To overcome this problem, we compute a suitable minimal-upper bound to this set given by the matrix associated with the Loewner-John Ellipsoid of the set of hyper-ellipsoids associated to the set of candidate lower-bound matrices. Finally, we present some numerical examples to compare the proposed approximated BB with the performance achieved by the maximum likelihood estimator.**

***Index Terms*—Barankin bound, multiple change-point estimation, performance analysis.**

I. Introduction

ESTIMATION of changes in time series is an important and active research area with several applications, for example, in fault detection, medical imaging, genetics, and econometrics. The literature is abundant concerning estimation algorithms for change-point estimation (see, e.g., [1]–[3]). However, less work has been done concerning the ultimate performance of such algorithms in terms of mean-square error (MSE). Indeed, if an estimator is available, the evaluation of its performance depends on knowing whether it is optimal or if further improvement is still possible. Note that some other criteria of performance in the context of sequential detection of a change-point are available in the literature; see, e.g., [4], [5], and references therein.

The classic way to analyze the performance of an estimator in terms of MSE is to compute the well-known Cramér–Rao bound (CRB) [6]. Unfortunately, for discrete time-measurement models the change-point location parameter is discrete; therefore the CRB, which is a function of the derivative of the likelihood of the observations w.r.t. the parameters, is not defined.

Several authors have proposed solutions to this problem. Indeed, in the change-point estimation framework, the CRB has already been studied using approximations (see, e.g., [7]–[12]). Depending on the particular parametrization of the data likelihood, two main challenges have been addressed concerning the CRB computation on the change-point time index: i) the discrete nature of the aforementioned parameter and ii) the regularity conditions of the likelihood of the observation. The former implies that the parameter does not have a defined derivative because of its discrete nature [10], and the latter implies that the likelihood of the observations has to be smooth (details are given in [6] and [13]), which is not the case for signal parameters with sudden changes. To overcome the discrete nature of the change-point time index, a continuous parametrization has been proposed (see, e.g., [12] and [14]). To satisfy the regularity conditions of the data likelihood, the step-like function, which represents a change in parameter, is generally approximated by another function with smooth properties (e.g., the so-called sigmoidal function introduced in [9] and [12] or a Heaviside function filtered by a Gaussian filter, as in [7]). This new function depends on parameters that have to be adjusted, and it tends to the step-like function when the appropriate values of these parameters are used. The main problem that appears when using this technique is that the CRB tends to zero when the approximate function tends to the step-like function [8], [12].

Moreover, it is noteworthy that these previous works concerning change-point estimation were always done in the framework of a single change point. To the best of our knowledge, performance bounds have never been derived in the multiple change-point context. The latter is important in off-line estimation of change points where batch-data are available, for example, in biomedical applications, such as

Manuscript received January 22, 2010; accepted July 21, 2010. Date of publication August 09, 2010; date of current version October 13, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ta-Hsin Li. This work was partially presented at the Second International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), St. Thomas, U.S. Virgin Islands, December 12–14, 2007. This work was supported by the Department of Defense under the Air Force Office of Scientific Research MURI Grant FA9550-05-1-0443 and ONR Grant N000140810849.

P. S. La Rosa is with the General Medical Sciences Division, Department of Medicine, Washington University in St. Louis, St. Louis, MO 63110 USA (e-mail: plarosa@wustl.edu).

A. Nehorai is with the Department of Electrical and Systems Engineering at Washington University in St. Louis, St. Louis, MO 63130 USA (e-mail: nehorai@ese.wustl.edu).

A. Renaux is with the Laboratoire des Signaux et Systèmes, Laboratory of Signals and Systems, University Paris-Sud 11, 91192 Gif-sur-Yvette cedex, France (e-mail: renaux@lss.supelec.fr).

C. Muravchik is with the CIC-PBA and Departamento de Electrotecnia, Universidad Nacional de La Plata, La Plata, Argentina (e-mail: carlosm@ing.unlp.edu.ar).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2010.2064771

DNA sequence segmentation [15], rat EEG segmentation (see [3, ch. 2]), and uterine MMG contraction detection [16], and in signal segmentation in general such as speech segmentation [17], astronomical data analysis [18].

In this paper, we analyze the Barankin bound (BB) [19] for multiple change-point estimation in the context of an independent vector sequence. The Barankin bound is the greatest lower bound for any unbiased estimator. Moreover, in contrast to the CRB, its computation is not limited by the discrete nature of the parameter and the regularity assumptions on the likelihood of the observations [13], [20]. However, the BB requires the use of parameters called test points. These test points choice is left to the user, and, in order to obtain the best (i.e., the tightest) bound, a nonlinear maximization over these test points has to be performed. This explains why this bound is so much less used and known than the CRB; nevertheless, the BB is often a practical bound for realistic scenarios (see, e.g., [21]).

To the best of our knowledge, minimal bounds other than the CRB have been proposed in the context of change-point estimation only in the foundational communication of Ferrari and Tourneret [22]. A simplified and practical version of the BB (i.e., one test point per parameter), the so-called Hammersley–Chapman–Robbins (HCR) bound [20], [23], is studied in that paper. As in the previous works on the CRB, only one change point is considered.

In this paper we extend the results presented in [22] to the case of multiple change points. We consider the multi-parameter HCR bound and we show that the so-called Barankin information matrix (BIM), which has to be inverted, has an interesting structure (*viz.*, a block diagonal matrix structure). We show that the estimation of one change point is corrupted by its neighboring change points and we give the details of the computation for the two change-point case. This case facilitates the derivation of a closed-form expression for the inverse of the BIM. Note that it is possible to find tighter bounds by using more test-points per parameter, however, such approach does not allow for obtaining closed-form expressions of the BIM and its inverse as derived here. We also discuss on the existence of the supremum of the finite set formed by all possible BB solutions and, following ideas from [24] and from convex optimization, we compute a suitable minimal-upper bound to this candidate set with respect to the Loewner cone, the set of semipositive definite matrices. In particular, we show that its computation is given by the matrix associated with the Loewner–John ellipsoid of the candidate set, which is the minimum-volume hyper-ellipsoid covering the set of hyper-ellipsoids associated to each matrix in the candidate set. We apply the bounds to the case of changes in the parameters of Gaussian and Poisson observations. We finally present numerical examples for comparing our bound to the performance achieved by the maximum likelihood estimator (MLE).

The notational convention adopted in this paper is as follows: italic indicates a scalar quantity, as in A; lowercase boldface indicates a vector quantity, as in $\mathbf{a}$; uppercase boldface indicates a matrix quantity, as in $\mathbf{A}$. The matrix transpose is indicated by a superscript ^{T} as in $\mathbf{A}^T$. The mth-row and nth-column element of the matrix $\mathbf{A}$ is denoted by $[\mathbf{A}]_{mn}$. The identity matrix of size $N \times N$ is denoted $\mathbf{I}_N$. We define by $\mathbf{1}_{M\times N}$ the matrix such that $[\mathbf{1}]_{mn} = 1$, $\forall m = 1 \dots M$ and $\forall n = 1 \dots N$, and $\mathbf{D}(\mathbf{a})$ is a diagonal matrix formed by the elements of the row vector $\mathbf{a}$. The trace operator is defined as $\mathrm{Tr}\{\cdot\}$. The determinant of a matrix is denoted by $|\cdot|$ and cardinality when applying to a set. $\mathbb{S}^n$ denotes the vector space of symmetric $n \times n$ matrices and the subsets of nonnegative definite matrices and positive definite matrices are denoted by $\mathbb{S}^n_+$ and $\mathbb{S}^n_{++}$, respectively. The notation $\mathbf{A} \succeq \mathbf{B}$ means that for $\mathbf{A}, \mathbf{B} \in \mathbb{S}^n$, $\mathbf{A}-\mathbf{B} \in \mathbb{S}^n_+$, also known as *Loewner partial ordering* of symmetric matrices [25], [26]. The absolute value is denoted by $\mathrm{abs}(\cdot)$. The indicator function of a set S is denoted by $I_S(\cdot)$. The expectation operator is denoted by $\mathbb{E}[\cdot]$. The observation and parameter spaces are denoted, respectively, by Ω and Θ.

The remainder of this paper is organized as follows. In Section II, we present the signal model, the assumptions, and we introduce the general structure for Barankin bound for the signal model parameters. The computation and analysis of the Barankin bound for the change-point localization parameters are provided in Section III. In Section IV, we analyze the cases of changes in the parameters of Gaussian and Poisson distributions. To illustrate our results, simulations are presented in Section V. Finally, in Section VI we conclude this work.

II. Problem Formulation

A. Observation Model

We consider the general case of N independent vector observations $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N] \in \mathbb{R}^{M\times N}$, which can be obtained, for example, by a multiple sensor system and are modeled as follows:

$$\begin{cases} \mathbf{x}_i \sim p_1(\mathbf{x}_i;\boldsymbol{\eta}_1) \text{ for } i = 1, \dots, t_1 \\ \mathbf{x}_i \sim p_2(\mathbf{x}_i;\boldsymbol{\eta}_2) \text{ for } i = t_1+1, \dots, t_2 \\ \vdots \\ \mathbf{x}_i \sim p_{q+1}(\mathbf{x}_i;\boldsymbol{\eta}_{q+1}) \text{ for } i = t_q+1, \dots, N \end{cases} \tag{1}$$

where M is the size of the sample vector (e.g., the number of sensors), q is the number of change-points, and p_j is a probability density function (or mass function for discrete random variables) with parameters $\boldsymbol{\eta}_j \in \mathbb{R}^L$. In other words, $\mathbf{x}_i \sim p_j(\mathbf{x}_i;\boldsymbol{\eta}_j)$ for $i = t_{j-1}+1, \dots, t_j$, with $j = 1, \dots, q+1$, where we define $t_0 = 0$ and $t_{q+1} = N$. Note that if $M = 1$, the problem is reduced to the estimation of changes in a time series. We assume that all probability density functions p_j belong to a common distribution. The unknown parameters of interest are the change-point locations $\{t_1, t_2, \dots, t_q\}$ with $\{t_k \in \mathbb{N} - \{0\}, \quad k = 1, \dots, q\}, 1 < t_1 < t_2 < \cdots < t_q < N$, and $q < N - 2$. The observations between two consecutive change points are assumed to be stationary. Consequently, the $q \times 1$ vector of unknown true parameters for this model is $\mathbf{t} = [t_1, t_2, \dots, t_q]^T$.

The observation model (1) is useful in signal processing; several examples were mentioned in the Introduction. Note that, since we focus on the change-point estimation, we assume that the parameters $\boldsymbol{\eta}_j$ are known. The resulting bound will still be useful if these parameters are unknown, but overly optimistic. Moreover, the complexity of the bound derivation increases for unknown $\boldsymbol{\eta}_j$ and therefore we do not consider this case in this work.

B. Barankin Bound

The P-order BB of a vector $\boldsymbol{\theta}_0 \in \mathbb{R}^q$, denoted by $\mathbf{BB}_P(\boldsymbol{\theta}_0, \mathbf{H}_{q\times P})$, is given as follows (see [27]–[30] for more details):

$$\begin{aligned}\mathbf{Cov}(\hat{\boldsymbol{\theta}}) &\succeq \mathbf{BB}_P(\boldsymbol{\theta}_0, \mathbf{H}_{q\times P}) \\ &= \mathbf{H}_{q\times P}(\boldsymbol{\Phi} - \mathbf{1}_{P\times P})^{-1}\mathbf{H}_{q\times P}^T \end{aligned} \tag{2}$$

where $\mathbf{Cov}(\hat{\boldsymbol{\theta}})$ is the covariance matrix of an unbiased estimator $\hat{\boldsymbol{\theta}}$ of the parameter vector $\boldsymbol{\theta}_0$. The matrix $\mathbf{H} = [\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0, \ldots, \boldsymbol{\theta}_P - \boldsymbol{\theta}_0]$ is a function of the set $\{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_P\}$, the so-called "test points," which are left to the user's choice. We define $\mathbf{h}_i = \boldsymbol{\theta}_i - \boldsymbol{\theta}_0$ such that the matrix $\mathbf{H} \in \mathbb{R}^{q\times P}$ becomes $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_P]$. Moreover, note that $\boldsymbol{\theta}_0 + \mathbf{h}_j \in \Theta$. In the following, for simplicity, we use the term "test point" for the vectors $\mathbf{h}_i$. Finally, $\boldsymbol{\Phi}$ is a $\mathbb{R}^{P\times P}$ matrix whose elements $[\boldsymbol{\Phi}]_{kl}$ are given by

$$[\boldsymbol{\Phi}]_{kl} = \mathbb{E}[L(\mathbf{X}, \boldsymbol{\theta}_0, \mathbf{h}_k) L(\mathbf{X}, \boldsymbol{\theta}_0, \mathbf{h}_l)] \tag{3}$$

where $L(\mathbf{X}, \boldsymbol{\theta}_0, \mathbf{h}_j)$ is defined by

$$L(\mathbf{X}, \boldsymbol{\theta}_0, \mathbf{h}_j) = \frac{p(\mathbf{X}; \boldsymbol{\theta}_0 + \mathbf{h}_j)}{p(\mathbf{X}; \boldsymbol{\theta}_0)} \tag{4}$$

where $p(\mathbf{X}; \boldsymbol{\varphi})$ is the likelihood of the observations with parameter vector $\boldsymbol{\varphi}$. Note that the matrix $\boldsymbol{\Phi} - \mathbf{1}_{P\times P}$ is sometimes referred to as the Barankin information matrix (BIM) [31].

As already stated, test points choice is left to the user, since any set of test points in $\mathbf{BB}_P(\boldsymbol{\theta}_0)$ satisfies the inequality (2). Thus, the tightest BB, denoted by $\mathbf{BB}(\boldsymbol{\theta}_0)$, is given as follows:

$$\mathbf{BB}(\boldsymbol{\theta}_0) = \lim_{P \longrightarrow |\Theta|} \sup_{\{\mathbf{h}_1, \ldots, \mathbf{h}_P\}} \mathbf{BB}_P(\boldsymbol{\theta}_0, \mathbf{H}_{q\times P}) \succeq \mathbf{CRB}(\boldsymbol{\theta}_0) \tag{5}$$

where $|\Theta|$ is the cardinality of the set Θ formed by all possible parameter values, and $\mathbf{CRB}(\boldsymbol{\theta}_0)$ is the CRB of $\boldsymbol{\theta}_0$, which, assuming that it exists, is smaller than the $\mathbf{BB}(\boldsymbol{\theta}_0)$ in the Loewner ordering sense. The computation of $\mathbf{BB}(\boldsymbol{\theta}_0)$ is costly, since the limit on P usually implies that a large, possibly infinite, number of test points needs to be considered, a nonlinear maximization over the test points has to be performed, and the inverse of the BIM has to be computed.

Concerning the BB for the parameter vector $\boldsymbol{\theta}_0 = \mathbf{t}$, $|\Theta|$ depends on the number of samples N and change points q as follows:

$$|\Theta| = \sum_{t_1=1}^{N-q} \sum_{t_2=t_1+1}^{N-q+1} \cdots \sum_{t_{q-1}=t_{q-2}+1}^{N-1} (N - t_{q-1} - 1) = \binom{N-1}{q}. \tag{6}$$

Note that $|\Theta| \to \infty$ as $(N - q) \to \infty$, and for N finite then $|\Theta|$ is finite. In practice, the number of test points and the particular structure of the matrix $\mathbf{H}$ is usually chosen based on the analytical and computational complexity associated with it, which lead to approximated versions of the BB. In the latter case it would be useful to have some knowledge of how different Barankin bound approximations compare among each other w.r.t. Loewner partial ordering. In the following proposition, we provide with a general guideline for this purpose.

Lemma 1: Let $\mathbf{A} \in \mathbb{S}_{++}^q, \mathbf{B} \in \mathbb{S}_{+}^q$ with $\operatorname{rank}(\mathbf{B}) = m < q$, and let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m > 0$ and $\lambda_{m+1} = \cdots = \lambda_q = 0$ be the roots of the characteristic equation $|\mathbf{B} - \lambda\mathbf{A}| = 0$. If $\lambda_1 \leq 1$, then $\mathbf{A} \succ \mathbf{B}$, otherwise $\mathbf{A}$ and $\mathbf{B}$ are not mutually comparable.

Proof: See Appendix A.

If $\operatorname{rank}(\mathbf{H}_{q\times P}) = q$, then $\mathbf{BB}_P(\boldsymbol{\theta}_0, \mathbf{H}_{q\times P}) \in \mathbb{S}_{++}^q$, since $(\boldsymbol{\Phi} - 1)^{-1} \in \mathbb{S}_{++}^q$ by construction, and if $\operatorname{rank}(\mathbf{H}_{q\times P'}) < q$ then $\mathbf{BB}_P(\boldsymbol{\theta}_0, \mathbf{H}_{q\times P'}) \in \mathbb{S}_{+}^q$. The lemma can now be used with $\mathbf{A} = \mathbf{BB}_P(\boldsymbol{\theta}_0, \mathbf{H}_{q\times P})$ and $\mathbf{B} = \mathbf{BB}_{P'}(\boldsymbol{\theta}_0, \tilde{\mathbf{H}}_{q\times P'})$ provided $\operatorname{rank}(\mathbf{H}_{q\times P}) = q > \operatorname{rank}(\tilde{\mathbf{H}}_{q\times P'})$. Note that $\operatorname{rank}(\tilde{\mathbf{H}}_{q\times P'}) < q$ implies that the number of test-points $P' < q$, therefore, a matrix bound $\mathbf{BB}_{P'}(\boldsymbol{\theta}_0, \tilde{\mathbf{H}}_{q\times P'})$ cannot be larger, w.r.t. Loewner partial ordering, than any matrix bound given by a test-point matrix $\mathbf{H}_{q\times P}$ consisting of $P = q$ independent test-point vectors. Consequently, in the following we will use an approximate version of the BB that allows us to derive efficiently computed closed-form expressions for the BIM and its inverse in the context of our multiple change-point estimation problem. In particular, we will compute the multi-parameter HCR bound [27] with the classical assumption of one test point per parameter ($P = q$), i.e., $\mathbf{h}_j = [0, \ldots, \alpha_j, \ldots, 0]^T$. Then, $\mathbf{H}$ is a diagonal matrix given by

$$\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_q] = \mathbf{D}(\boldsymbol{\alpha}) \tag{7}$$

where the vector $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_q]^T$ corresponds to the set of test points associated to the parameters $\mathbf{t} = [t_1, t_2, \ldots, t_q]^T$. Note that $\alpha_j \neq 0$ is defined such that $t_j + \alpha_j$ ranges over all possible values of t_j, for $j = 1, \ldots, q$. Thus, $\alpha_j \in \{\mathbb{Z} \cap [t_{j-1} - t_j + 1, t_{j+1} - t_j - 1] - \{0\}\}$. Let $S \subset \mathbb{Z}^q$ be a set formed by all possible values of $\boldsymbol{\alpha}$. The set S is finite, given that t_{q+1} is finite.

The matrix $\boldsymbol{\Phi} - \mathbf{1}_{q\times q}$ corresponds to the BIM for change-point locations $\mathbf{t}$, denoted here by $\mathbf{BIM}_{\mathbf{t}}$. The approximated BB, $\mathbf{BB}_{\mathbf{t},q}$, is then obtained from

$$\mathbf{BB}_{\mathbf{t},q} = \sup_{[\mathbf{h}_1, \ldots, \mathbf{h}_q]} \mathbf{BB}_q(\mathbf{t}, \mathbf{H}_{q\times q}) = \sup_{\boldsymbol{\alpha} \in S} \mathbf{D}(\boldsymbol{\alpha}) \mathbf{BIM}_{\mathbf{t}}^{-1} \mathbf{D}(\boldsymbol{\alpha})^T. \tag{8}$$

By construction, the finite set $C := \{\mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha})), \boldsymbol{\alpha} \in S\}$ is a subset of the partially ordered set $(\mathbb{S}^q, \preceq)$ with partial order "$\preceq$" given by the Loewner ordering [25], [26]. This partial order is not a lattice ordering, i.e., each finite subset of $\mathbb{S}^q$ may not be closed under the least-upper (infimum) and greatest-lower bounds (supremum) [26]. In other words, the notion of a unique supremum or an infimum of C might not exist with respect to the Loewner ordering. The supremum does not exist if there is no upper bound to the set, or if the set of upper bounds does not have a least element. If the supremum exists, it does not need to be defined in the set, but if it belongs to it, then it is the greatest element[1] in the set. Note that a set with respect to the partially

[1] $\mathbf{B}_i \in C$ is the greatest element of C *w.r.t.* $(\mathbb{S}^q, \preceq)$ if $\mathbf{B}_i \succeq \mathbf{Y}$ for all $\mathbf{Y} \in C$. If the greatest element exists it is an upper-bound of C contained in it. The least element of C is defined similarly considering $\mathbf{B}_i \preceq \mathbf{Y}$.

ordered set $(\mathbb{S}^q, \succeq)$ may have several maximal[2] and minimal elements without having a greatest and least element in the set, respectively. If the set has a greatest or least element, then it is the unique maximal or minimal element, and therefore it is the supremum or infimum of the set. Here, we will approach the computation of the supremum by computing a suitable minimal element of the set of upper bounds of C, namely, a minimal-upper bound $\mathbf{B}_q \in \mathbb{S}^q_{++}$ such that $\mathbf{B}_q \succeq C$ and which is minimal in the sense that there is not smaller matrix $\mathbf{B}'_q \preceq \mathbf{B}_q$ such that $\mathbf{B}'_q \succeq C$. From (2), $\mathbf{Cov}(\hat{\boldsymbol{\theta}})$ belongs to set of upper bounds of C, therefore if the set of upper bounds has a unique minimal element, i.e., a least element, then $\mathbf{Cov}(\hat{\boldsymbol{\theta}}) \succeq \mathbf{B}_q$. However, if the set of upper bounds has several minimal elements, then in general we can expect that $\mathbf{Cov}(\hat{\boldsymbol{\theta}}) \succeq \mathbf{B}_q$, or that $\mathbf{Cov}(\hat{\boldsymbol{\theta}})$ and $\mathbf{B}_q$ are not mutually comparable.

Having a closed form for $\mathbf{BIM}_\mathbf{t}^{-1}$ makes the task of computing $\mathbf{B}_q$ much less computationally demanding than having to invert $\mathbf{BIM}_\mathbf{t}$ for every $\boldsymbol{\alpha} \in S$. In the following section, we will first derive the elements of $\mathbf{BIM}_\mathbf{t}$ and obtain closed-form expressions for $\mathbf{BIM}_\mathbf{t}^{-1}$. Then, we will introduce the approach for computing the minimal-upper bound $\mathbf{B}_q$.

III. Barankin Bound Type for Multiple Change-Point Estimation

To compute the BB for the change point localization parameters, we first need to compute $\mathbf{BIM}_\mathbf{t}$, which depends on the matrix $\boldsymbol{\Phi}$. From (3) and (4), the elements of $[\boldsymbol{\Phi}]_{kl}$, for $k, l = 1, \ldots, q$ are given by

$$[\boldsymbol{\Phi}]_{kl} = \int_\Omega \frac{p(\mathbf{X}; \mathbf{t} + \mathbf{h}_k) p(\mathbf{X}; \mathbf{t} + \mathbf{h}_l)}{p(\mathbf{X}; \mathbf{t})} d\mathbf{X} \tag{9}$$

where $p(\mathbf{X}; \mathbf{t})$ is given by

$$p(\mathbf{X}; \mathbf{t}) = \Pi_{i=1}^{t_1} p_1(\mathbf{x}_i; \boldsymbol{\eta}_1) \cdots \Pi_{i=t_{k-1}+1}^{t_k} p_k(\mathbf{x}_i; \boldsymbol{\eta}_k) \cdots \Pi_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i; \boldsymbol{\eta}_{q+1}) \tag{10}$$

and $p(\mathbf{X}; \mathbf{t} + \mathbf{h}_k)$ is given by

$$p(\mathbf{X}; \mathbf{t} + \mathbf{h}_k) = \Pi_{i=1}^{t_1} p_1(\mathbf{x}_i; \boldsymbol{\eta}_1) \cdots \Pi_{i=t_{k-1}+1}^{t_k+\alpha_k} p_k(\mathbf{x}_i; \boldsymbol{\eta}_k) \cdots \Pi_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i; \boldsymbol{\eta}_{q+1}) \tag{11}$$

and where $p(\mathbf{X}; \mathbf{t} + \mathbf{h}_l)$ is same as (11) ($k = l$).

In order to study and to simplify $\boldsymbol{\Phi}$, we will analyze its diagonal and non-diagonal elements separately.

A. Diagonal Elements of $\boldsymbol{\Phi}$

Replacing $k = l$ in (9) and using (11), we obtain the following expression:

$$[\boldsymbol{\Phi}]_{kk} = \int_\Omega \frac{\Pi_{i=1}^{t_1} p_1^2(\mathbf{x}_i; \boldsymbol{\eta}_1) \cdots \Pi_{i=t_{k-1}+1}^{t_k+\alpha_k} p_k^2(\mathbf{x}_i; \boldsymbol{\eta}_k) \cdots}{\Pi_{i=1}^{t_1} p_1(\mathbf{x}_i; \boldsymbol{\eta}_1) \cdots \Pi_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i; \boldsymbol{\eta}_{q+1})} \cdots \Pi_{i=t_q+1}^{N} p_{q+1}^2(\mathbf{x}_i; \boldsymbol{\eta}_{q+1}) d\mathbf{X}. \tag{12}$$

[2] $\mathbf{B}_i \in C$ is a maximal element of C w.r.t. $(\mathbb{S}^q, \preceq)$ if there is not $\mathbf{Y} \in C$ such that $\mathbf{Y} \succeq \mathbf{B}_i$ and is a minimal element if there is not $\mathbf{Y} \in C$ such that $\mathbf{B}_i \succeq \mathbf{Y}$.

This equation can be further simplified by considering the cases $\alpha_k > 0$ and $\alpha_k < 0$, obtaining the following expression (see Appendix B for details on its derivation):

$$[\boldsymbol{\Phi}]_{kk} = \begin{cases} \left(\int_\Omega \frac{p_k^2(\mathbf{x}; \boldsymbol{\eta}_k)}{p_{k+1}(\mathbf{x}; \boldsymbol{\eta}_{k+1})} d\mathbf{x} \right)^{\alpha_k}, & \text{If } \alpha_k > 0 \\ \left(\int_\Omega \frac{p_{k+1}^2(\mathbf{x}; \boldsymbol{\eta}_{k+1})}{p_k(\mathbf{x}; \boldsymbol{\eta}_k)} d\mathbf{x} \right)^{-\alpha_k}, & \text{If } \alpha_k < 0. \end{cases} \tag{13}$$

B. Non-Diagonal Elements of $\boldsymbol{\Phi}$

The computation of the off-diagonal elements of $\boldsymbol{\Phi}$ can be simplified by using the fact that the matrix $\boldsymbol{\Phi}$ is symmetric; therefore, we can focus on either the upper or lower triangular part of $\boldsymbol{\Phi}$. In our derivations below we consider the upper triangular part, i.e., $k < l$, then by using (9) and (11), we obtain the following expression for the elements of $\boldsymbol{\Phi}$:

$$\begin{aligned} [\boldsymbol{\Phi}]_{kl} = \int_\Omega & \frac{\Pi_{i=1}^{t_1} p_1(\mathbf{x}_i; \boldsymbol{\eta}_1) \cdots \Pi_{i=t_{k-1}+1}^{t_k+\alpha_k} p_k(\mathbf{x}_i; \boldsymbol{\eta}_k)}{\Pi_{i=1}^{t_1} p_1(\mathbf{x}_i; \boldsymbol{\eta}_1) \cdots \cdots \Pi_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i; \boldsymbol{\eta}_{q+1})} \\ & \times \Pi_{i=t_k+\alpha_k+1}^{t_{k+1}} p_{k+1}(\mathbf{x}_i; \boldsymbol{\eta}_{k+1}) \cdots \\ & \Pi_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i; \boldsymbol{\eta}_{q+1}) \\ & \times \Pi_{i=1}^{t_1} p_1(\mathbf{x}_i; \boldsymbol{\eta}_1) \cdots \Pi_{i=t_{l-1}+1}^{t_l+\alpha_l} p_l(\mathbf{x}_i; \boldsymbol{\eta}_l) \cdots \\ & \Pi_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i; \boldsymbol{\eta}_{q+1}) d\mathbf{X}. \end{aligned} \tag{14}$$

Following the same idea as for the diagonal elements, $[\boldsymbol{\Phi}]_{kl}$ can be simplified by analyzing the four possible combinations of test-point ranges, namely,

$$\begin{cases} \text{Case 1: } \alpha_k > 0 \text{ and } \alpha_l > 0 \\ \text{Case 2: } \alpha_k < 0 \text{ and } \alpha_l < 0 \\ \text{Case 3: } \alpha_k < 0 \text{ and } \alpha_l > 0 \\ \text{Case 4: } \alpha_k > 0 \text{ and } \alpha_l < 0. \end{cases} \tag{15}$$

For the last case, i.e., $\alpha_k > 0$ and $\alpha_l < 0$, two subcases have to be analyzed: i) $t_k + \alpha_k < t_l + \alpha_l$ and ii) $t_k + \alpha_k > t_l + \alpha_l$. These two cases correspond to non-overlapping and overlapping test points, respectively. Note that since $k < l, t_k < t_l$ and since $\alpha_j \in \{\mathbb{Z} \cap [t_{j-1} - t_j + 1, t_{j+1} - t_j - 1] - \{0\}\}$, the case $t_k + \alpha_k > t_l + \alpha_l$ which corresponds to an overlapping between two test points, can appear only when $l = k + 1$, or, in other words, when we are analyzing two neighboring change points. Then, for Cases 1-3 and subcase i), (14) becomes (see Appendix C)

$$[\boldsymbol{\Phi}]_{kl} = 1, \quad \text{for } l > k \tag{16}$$

and for subcase (ii), keeping in mind that $\alpha_k > 0$ and $\alpha_{k+1} < 0$, (14) becomes

$$[\boldsymbol{\Phi}]_{kl} = \begin{cases} \left(\int_\Omega \frac{p_k(\mathbf{x}; \boldsymbol{\eta}_k) p_{k+2}(\mathbf{x}; \boldsymbol{\eta}_{k+2})}{p_{k+1}(\mathbf{x}; \boldsymbol{\eta}_{k+1})} d\mathbf{x} \right)^{\beta_k}, & \text{for } l = k+1 \\ 1, & \text{for } l > k+1 \end{cases} \tag{17}$$

where $\beta_k = (t_k + \alpha_k) - (t_{k+1} + \alpha_{k+1})$.

Remark: This last result is fundamental because it proves the natural intuition that the estimation of q change points is not equivalent to q times the estimation of one change point. In other

words, it means that the estimation of one change point is perturbed by its two neighbors. We now summarize the previous results.

C. Barankin Information Matrix $\mathbf{\Phi} - \mathbf{1}_{q\times q}$

Using (13), (16), and (17), it is clear that $\mathbf{BIM_t}$ has at least a tridiagonal structure:

$$\mathbf{BIM_t} = \begin{bmatrix} A_1 & B_1 & 0 & \cdots & 0 \\ B_1 & A_2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & A_{q-1} & B_{q-1} \\ 0 & \cdots & 0 & B_{q-1} & A_q \end{bmatrix} \tag{18}$$

where

$$\begin{aligned} A_k &= [\mathbf{\Phi}]_{kk} - 1, \quad \text{for } k = 1, \ldots, q \\ &= \begin{cases} \left(\int_\Omega \frac{p_k^2(\mathbf{x};\boldsymbol{\eta}_k)}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})} d\mathbf{x} \right)^{\alpha_k} - 1 & \text{If} \alpha_k > 0 \\ \left(\int_\Omega \frac{p_{k+1}^2(\mathbf{x};\boldsymbol{\eta}_{k+1})}{p_k(\mathbf{x};\boldsymbol{\eta}_k)} d\mathbf{x} \right)^{-\alpha_k} - 1 & \text{If } \alpha_k < 0 \end{cases} \end{aligned} \tag{19}$$

and

$$\begin{aligned} B_k &= [\mathbf{\Phi}]_{kk+1} - 1, \quad \text{for } k = 1, \ldots, q-1 \\ &= \begin{cases} 0, & \text{If } \beta_k < 0 \\ \left(\int_\Omega \frac{p_k(\mathbf{x};\boldsymbol{\eta}_k) p_{k+2}(\mathbf{x};\boldsymbol{\eta}_{k+2})}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})} d\mathbf{x} \right)^{\beta_k} - 1, & \text{If } \beta_k > 0. \end{cases} \end{aligned} \tag{20}$$

In the case of one change-point estimation, $\mathbf{BIM_t}$ is reduced to a scalar A_1, and by replacing $\alpha_1 = \alpha$ we re-obtain the result proposed by Ferrari and Tourneret (see (5) and (6) in [22]):

$$A_1 = \begin{cases} \left(\int_\Omega \frac{p_1^2(\mathbf{x};\boldsymbol{\eta}_1)}{p_2(\mathbf{x};\boldsymbol{\eta}_2)} d\mathbf{x} \right)^{\alpha} - 1, & \text{If } \alpha > 0 \\ \left(\int_\Omega \frac{p_2^2(\mathbf{x};\boldsymbol{\eta}_2)}{p_1(\mathbf{x};\boldsymbol{\eta}_1)} d\mathbf{x} \right)^{-\alpha} - 1, & \text{If } \alpha < 0. \end{cases} \tag{21}$$

Note also that the diagonal elements of $\mathbf{BIM_t}$ can be computed numerically in one step (i.e., $\forall \alpha_k \gtrless 0$) as follows:

$$A_k = \left(\int_\Omega \left(\frac{p_k(\mathbf{x};\boldsymbol{\eta}_k)}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})} \right)^{\epsilon_k} p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1}) d\mathbf{x} \right)^{\mathrm{abs}(\alpha_k)} - 1 \tag{22}$$

where $\epsilon_k = (1/2)(3(\alpha_k)/(\mathrm{abs}(\alpha_k) + 1)$.

The next step of our analysis is to compute $(\mathbf{BIM_t})^{-1}$. For a given set of test points, it is clear that $t_k + \alpha_k > t_{k+1} + \alpha_{k+1} \Longrightarrow t_{k+1} + \alpha_{k+1} < t_{k+2} + \alpha_{k+2}$, since $\alpha_j \in \{\mathbb{Z} \cap [t_{j-1} - t_j + 1, t_{j+1} - t_j - 1] - \{0\}\}$. In other words, $\forall k$, if $B_k \neq 0$, then $B_{k+1} = B_{k-1} = 0$; therefore, $\mathbf{BIM_t}$ is block diagonal and the maximum size of one block is 2×2. Since the problem is reduced to finding, at worst, the inverse of several 2×2 matrices with the same structure, we will have a straightforward inversion. In this section, we detail the case of two change points, we give the generalization to two neighboring points, and we use this to derive a closed-form expression for the inverse of $\mathbf{BIM_t}$ and thus $\mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}))$.

1) The Case of Two Change Points: In this case we have $q = 2, \mathbf{t} = [t_1, t_2]^T$, and $\mathbf{BB}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}))$ becomes

$$\mathbf{BB}_2(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha})) = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \begin{bmatrix} A_1 & B_1 \\ B_1 & A_2 \end{bmatrix}^{-1} \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \tag{23}$$

with

$$A_1 = \begin{cases} \left(\int_\Omega \frac{p_1^2(\mathbf{x};\boldsymbol{\eta}_1)}{p_2(\mathbf{x};\boldsymbol{\eta}_2)} d\mathbf{x} \right)^{\alpha_1} - 1, & \text{If } \alpha_1 > 0 \\ \left(\int_\Omega \frac{p_2^2(\mathbf{x};\boldsymbol{\eta}_2)}{p_1(\mathbf{x};\boldsymbol{\eta}_1)} d\mathbf{x} \right)^{-\alpha_1} - 1, & \text{If } \alpha_1 < 0 \end{cases}$$

$$A_2 = \begin{cases} \left(\int_\Omega \frac{p_2^2(\mathbf{x};\boldsymbol{\eta}_2)}{p_3(\mathbf{x};\boldsymbol{\eta}_3)} d\mathbf{x} \right)^{\alpha_2} - 1, & \text{If } \alpha_2 > 0 \\ \left(\int_\Omega \frac{p_3^2(\mathbf{x};\boldsymbol{\eta}_3)}{p_2(\mathbf{x};\boldsymbol{\eta}_2)} d\mathbf{x} \right)^{-\alpha_2} - 1, & \text{If } \alpha_2 < 0 \end{cases} \tag{24}$$

$$B_1 = \begin{cases} 0, & \text{If } \beta_1 < 0 \\ \left(\int_\Omega \frac{p_1(\mathbf{x};\boldsymbol{\eta}_1) p_3(\mathbf{x};\boldsymbol{\eta}_3)}{p_2(\mathbf{x};\boldsymbol{\eta}_2)} d\mathbf{x} \right)^{\beta_1} - 1, & \text{If } \beta_1 > 0 \end{cases} \tag{25}$$

where $\beta_1 = (t_1 + \alpha_1) - (t_2 + \alpha_2)$.

Consequently, depending on the given set of test points, the following five combinations, corresponding respectively to Cases 1, 2, 3, and 4 in (15), are possible for $\mathbf{BB}_2(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}))$:

$$\begin{aligned} &\left\{ \begin{bmatrix} \frac{\alpha_1^2}{\Delta_{112}^{\alpha_1} - 1} & 0 \\ 0 & \frac{\alpha_2^2}{\Delta_{223}^{\alpha_2} - 1} \end{bmatrix}, \begin{bmatrix} \frac{\alpha_1^2}{\Delta_{221}^{\mathrm{abs}(\alpha_1)} - 1} & 0 \\ 0 & \frac{\alpha_2^2}{\Delta_{332}^{\mathrm{abs}(\alpha_2)} - 1} \end{bmatrix}, \right. \\ &\quad \begin{bmatrix} \frac{\alpha_1^2}{\Delta_{221}^{\mathrm{abs}(\alpha_1)} - 1} & 0 \\ 0 & \frac{\alpha_2^2}{\Delta_{223}^{\alpha_2} - 1} \end{bmatrix}, \begin{bmatrix} \frac{\alpha_1^2}{\Delta_{112}^{\alpha_1} - 1} & 0 \\ 0 & \frac{\alpha_2^2}{\Delta_{332}^{\mathrm{abs}(\alpha_2)} - 1} \end{bmatrix}, \\ &\quad \left. \kappa^{-1} \begin{bmatrix} \alpha_1^2 \left(\Delta_{332}^{\mathrm{abs}(\alpha_2)} - 1 \right) & \alpha_1 \alpha_2 \left(1 - \Delta_{132}^{\beta_1} \right) \\ \alpha_1 \alpha_2 \left(1 - \Delta_{132}^{\beta_1} \right) & \alpha_2^2 \left(\Delta_{112}^{\alpha_1} - 1 \right) \end{bmatrix} \right\} \end{aligned} \tag{26}$$

where we define $\Delta_{ijk} = \int_\Omega (p_i(\mathbf{x};\boldsymbol{\eta}_i) p_j(\mathbf{x};\boldsymbol{\eta}_j) / p_k(\mathbf{x};\boldsymbol{\eta}_k)) d\mathbf{x}$, and $\kappa = (\Delta_{112}^{\alpha_1} - 1)(\Delta_{332}^{\mathrm{abs}(\alpha_2)} - 1) - (\Delta_{132}^{\beta_1} - 1)^2$.

2) Generalization to Q Change Points: Note that for more change points the process is the same, except that the inversion has to be computed because of the increase of possibilities. However, the matrix to be inverted is block diagonal, with block of size 1×1 or 2×2, as stated in the previous section. In particular, depending on the values of $\boldsymbol{\alpha}$, the elements of $[\mathbf{BIM_t^{-1}}]_{kl}$ for $1 < k < q$ and $l = \{k, k+1\}$, with $\mathbf{BIM_t}$, A_k, and B_k given by (18), (19), and (20), respectively, and $\alpha_0 = \alpha_q = 0$ and $B_q = 0$, have the following possible values:

If $t_{k+1} + \alpha_{k+1} < t_k + \alpha_k$, then $\alpha_k > 0, B_k \neq 0$ and $B_{k-1} = B_{k+1} = 0$, thus

$$\left[\mathbf{BIM_t^{-1}}\right]_{kl} = \begin{cases} \frac{A_{k+1}}{A_k A_{k+1} - B_k^2}, & \text{for } l = k, \\ -\frac{B_k}{A_k A_{k+1} - B_k^2}, & \text{for } l = k+1. \end{cases} \tag{27}$$

If $t_k + \alpha_k < t_{k-1} + \alpha_{k-1}$, then $\alpha_k < 0, B_{k-1} \neq 0$ and $B_{k-2} = B_k = 0$, thus

$$\left[\mathbf{BIM_t^{-1}}\right]_{kl} = \begin{cases} \frac{A_{k-1}}{A_k A_{k-1} - B_{k-1}^2}, & \text{for } l = k \\ 0, & \text{for } l = k+1. \end{cases} \tag{28}$$

If $t_{k-1} + \alpha_{k-1} < t_k + \alpha_k < t_{k+1} + \alpha_{k+1}$, then $B_{k-1} = B_k = 0$, thus

$$\left[\mathbf{BIM_t^{-1}}\right]_{kl} = \begin{cases} \frac{1}{A_k}, & \text{for } l = k \\ 0, & \text{for } l = k+1. \end{cases} \tag{29}$$

Therefore, the elements of $[\mathbf{BIM}_{\mathbf{t}}^{-1}]_{kl}$ for $k,l = 1,\ldots,q$, which is a symmetric matrix, are given by

$$\left[\mathbf{BIM}_{\mathbf{t}}^{-1}\right]_{kl} = \begin{cases} \frac{A_{k-1}I_{[-\infty,-1]}(\alpha_k)+A_{k+1}I_{[1,\infty]}(\alpha_k)}{A_k(A_{k-1}I_{[-\infty,-1]}(\alpha_k)+A_{k+1}I_{[1,\infty]}(\alpha_k))-\left(B_{k-1}^2+B_k^2\right)}, & \text{for } l=k \\ \frac{-B_k}{A_kA_{k+1}-B_k^2}, & \text{for } l=k+1, \\ 0, & \text{for } l>k+1. \end{cases} \tag{30}$$

Since the matrix $\mathbf{BB}_q(\mathbf{t},\mathbf{D}(\boldsymbol{\alpha}))=\mathbf{D}(\boldsymbol{\alpha})\mathbf{BIM}_{\mathbf{t}}^{-1}\mathbf{D}(\boldsymbol{\alpha})^T$, then $[\mathbf{BB}_q(\mathbf{t},\mathbf{D}(\boldsymbol{\alpha}))]_{kl}$ for $k,l=1,\ldots,q$ is given as follows:

$$[\mathbf{BB}_q(\mathbf{t},\mathbf{D}(\boldsymbol{\alpha}))]_{kl} = \begin{cases} \frac{\alpha_k^2(A_{k-1}I_{[-\infty,-1]}(\alpha_k)+A_{k+1}I_{[1,\infty]}(\alpha_k))}{A_k(A_{k-1}I_{[-\infty,-1]}(\alpha_k)+A_{k+1}I_{[1,\infty]}(\alpha_k))-\left(B_{k-1}^2+B_k^2\right)}, & \text{for } l=k \\ -\frac{\alpha_k\alpha_{k+1}B_k}{A_kA_{k+1}-B_k^2}, & \text{for } l=k+1 \\ 0, & \text{for } l>k+1 \end{cases} \tag{31}$$

where $[\mathbf{BB}_q(\mathbf{t},\mathbf{D}(\boldsymbol{\alpha}))]_{kl} = [\mathbf{BB}_q(\mathbf{t},\mathbf{D}(\boldsymbol{\alpha}))]_{lk}$. If for a given set of test points there is no overlap with the neighboring change-points t_{k-1} and t_{k+1}, then $B_{k-1}=B_k=0$ in (31) and we obtain the particular result $[\mathbf{BB}_q(\mathbf{t},\mathbf{D}(\boldsymbol{\alpha}))]_{kk} = \alpha_k^2/A_k$ and $[\mathbf{BB}_q(\mathbf{t},\mathbf{D}(\boldsymbol{\alpha}))]_{kk+1} = [\mathbf{BB}_q(\mathbf{t},\mathbf{D}(\boldsymbol{\alpha}))]_{k+1k} = 0$. This is equivalent to the bound obtained using the same set of test points and assuming one change-point located in the time interval between t_{k-1} and t_{k+1} with total numbers of time-samples $N = t_{k+1} - t_{k-1}$.

D. Computation of the Supremum

To obtain the tightest bound from the finite set $C := \{\mathbf{BB}_q(\mathbf{t},\mathbf{D}(\boldsymbol{\alpha})), \boldsymbol{\alpha} \in S\} \subset \mathbb{S}_{++}^q$, we need to compute the supremum of C with respect to the partially ordered set $(\mathbb{S}^q, \succeq)$. The partial order is given by the Loewner ordering, which is defined via the cone of positive semidefinite matrices [25], [26]. In general, this problem is indeed very complex since it requires to look for $\boldsymbol{\alpha}^* \in S$ such that $\mathbf{BB}_q(\mathbf{t},\mathbf{D}(\boldsymbol{\alpha}^*)) \succeq \mathbf{BB}_q(\mathbf{t},\mathbf{D}(\boldsymbol{\alpha}))$ for all $\boldsymbol{\alpha} \in S$. To the best of our knowledge, no formal approach for solving this problem has been proposed in the technical literature of minimal bounds. For example, in [28] and [32], the choice of the test point $\boldsymbol{\alpha}$ is guided by some physical considerations of the model being studied. Also, from an optimal design context [25], an approximation for solving this problem is to compute the matrix in C with the largest trace, $\mathbf{BB}_{\mathrm{tr}}$. However, the fact that $\mathrm{Tr}\{\mathbf{BB}_{\mathrm{tr}}\} > \mathrm{Tr}\{\mathbf{B}_i\}$ for $\mathbf{B}_i \in C$, does not imply that $\mathbf{BB}_{\mathrm{tr}} \succ \mathbf{B}_i$, only the converse statement is valid. In fact, only if C has a greatest element, i.e., the supremum of the set, then it is given by the matrix in C with the largest trace Let $\mathbf{B}_j = \sup C$, with $\mathbf{B}_j \in C$, then by definition $\mathbf{B}_j \succeq \mathbf{B}_i$, for all $\mathbf{B}_i \in C$ with $i \neq j$. Let $\mathbf{G} = \mathbf{B}_j - \mathbf{B}_i$, thus $\mathbf{G} \in \mathbb{S}_+^q$ and $\mathrm{Tr}\{\mathbf{G}\} > 0$. Hence, $\mathrm{Tr}\{\mathbf{B}_j\} > \mathrm{Tr}\{\mathbf{B}_i\}$, for all $\mathbf{B}_i \in C$ with $i \neq j$, but as we discussed at the end of Section II, a unique supremum or an infimum with respect to the Loewner partial ordering in the finite set C might not exist.

Here we address the computation of the supremum by finding a minimal-upper bound $\mathbf{B}_q \in \mathbb{S}_{++}^q$ to the set C such that $\mathbf{B}_q \succeq C$ and which is minimal in the sense that there is no smaller matrix $\mathbf{B}_q' \preceq \mathbf{B}_q$ such that $\mathbf{B}_q' \succeq C$. In [24], the authors implicitly introduced an algorithm for computing a minimal-upper bound to a finite set of positive definite matrices and redefined this element as the supremum of the set. Before discussing more details about it, we need to introduce the so-called penumbra $P(\mathbf{M})$ of a matrix $\mathbf{M} \in \mathbb{S}^q$ as the set $P(\mathbf{M}) := \{\mathbf{N} \in \mathbb{S}^q : \mathbf{N} \preceq \mathbf{M}\}$ [24], [25] and the following proposition.

Proposition 2: Define $\mathbf{M}$ and $\mathbf{N} \in \mathbb{S}^q$, then $\mathbf{M} \succeq \mathbf{N}$ iff $P(\mathbf{N}) \subseteq P(\mathbf{M})$.

Proof: If $P(\mathbf{N}) \subseteq P(\mathbf{M})$, then $\mathbf{N} \in P(\mathbf{M})$ and then, by the definition of penumbra, $\mathbf{M} \succeq \mathbf{N}$. To prove the other implication, we define a matrix $\mathbf{G} \in \mathbb{S}^q$ such that $\mathbf{N} \succeq \mathbf{G}$. Then if $\mathbf{M} \succeq \mathbf{N}$ we have, by the transitivity property of the Loewner order, $\mathbf{M} \succeq \mathbf{G}$, namely, $\mathbf{M} \succeq \mathbf{N} \succeq \mathbf{G}$. Therefore, all the matrix elements in $P(\mathbf{N})$ are also in $P(\mathbf{M})$, thus, $P(\mathbf{N}) \subseteq P(\mathbf{M})$. ■

The penumbra $P(\mathbf{M})$ is seen as an inverted cone of vertex $\mathbf{M}$ characterizing all matrices that are smaller than $\mathbf{M}$ [24], [25]. The authors in [24] and [25] redefined the supremum of a set of matrices as the matrix associated to the vertex of the minimal penumbra covering the penumbras of all the matrices in the set. The minimal-penumbra vertex is a minimal-upper bound to the set with respect to the partially ordered set $(\mathbb{S}^q, \succeq)$. In [24], the minimal-penumbra vertex is computed by associating with each matrix $\mathbf{M} \in \mathbb{S}^q$ a ball in the subspace $\mathbb{S}_{\mathbf{A}} = \{\mathbf{A} : \mathrm{Tr}\{\mathbf{A}\} = 0\}$, and the authors show that it is determined by the smallest ball enclosing the set of balls associated to each matrix in the set. The latter algorithm is implemented in an approximate manner, by solving instead the problem of finding the smallest enclosing ball of a set of points which correspond to samples from the boundaries of each ball. The success of this method to obtain a minimal-upper bound matrix depends on the samples chosen. For example, in the case of two balls, it is easy to show that the smallest enclosing ball is tangent to each ball border at the two farthest points from the set of points defined by the intersection of a line passing through each ball center and each ball boundary. Therefore, if the sampling procedure does not include this pair of points, then the resulting ball does not completely enclose both balls and, thus, the resulting matrix is not a minimal-upper bound. Moreover, when the dimension is larger than two, a simple analytical computation shows that this algorithm fails to obtain a minimal-upper bound matrix for the set formed by two diagonal matrices not comparable to each other according to Loewner order.

Here, instead, we propose a method for computing a suitable $\mathbf{B}_q$ for any dimension. First, we show that computing $\mathbf{B}_q$ is equivalent to finding the minimum-volume hyper-ellipsoid covering the set of hyper-ellipsoids associated to each matrix in the set C. And second, we show that this problem can be written as a convex objective function with convex constraints which can be solved efficiently using semidefinite programming. An hyper-ellipsoid $\varepsilon \subset \mathbb{R}^q$ with non-empty interior and centered at the origin can be represented by the set $\varepsilon(\mathbf{F}) = \{\mathbf{x}^T\mathbf{F}^{-1}\mathbf{x} \leq \mathbf{1}\}$, where $\mathbf{F} \in \mathbb{S}_{++}^q$. Suppose $\varepsilon(\tilde{\mathbf{F}})$ is another hyper-ellipsoid similarly represented, where $\tilde{\mathbf{F}} \in \mathbb{S}_{++}^q$. Then, the following statement holds.

Lemma 3: $\mathbf{F} \succeq \tilde{\mathbf{F}}$ iff $\varepsilon(\mathbf{F}) \supseteq \varepsilon(\tilde{\mathbf{F}})$.

Proof: By the S-procedure [33], we have that $\varepsilon(\tilde{\mathbf{F}}) \subseteq \varepsilon(\mathbf{F})$ if and only if there is a $\lambda > 0$ such that

$$\begin{bmatrix} \mathbf{F}^{-1} & 0 \\ 0 & -1 \end{bmatrix} \preceq \lambda \begin{bmatrix} \tilde{\mathbf{F}}^{-1} & 0 \\ 0 & -1 \end{bmatrix},$$

with equality when $\lambda = 1$, implying the necessary condition $\tilde{\mathbf{F}} \preceq \mathbf{F}$. ■

Given a finite set of hyper-ellipsoids $C_\varepsilon := \{\varepsilon(\mathbf{F}_i) | \mathbf{F}_i \in \mathbb{S}_{++}^q, i = 1, \ldots, R\}$, we can always find a unique minimum volume hyper-ellipsoid, $\varepsilon(\mathbf{F}_{jl})$, containing the set C_ε, i.e., containing all $\varepsilon(\mathbf{F}_i)$ [33]. Since C_ε is convex, $\varepsilon(\mathbf{F}_{jl})$ is known as the Loewner–John ellipsoid of C_ε [33] and, as we show in the following statement, $\mathbf{F}_{jl}$ is a minimal-upper bound of the set $C_F := \{\mathbf{F}_i, i = 1, \ldots, R\}$ formed by all the matrices associated to the hyper-ellipsoids in C_ε.

Theorem 4: The matrix $\mathbf{F}_{jl}$, associated to the Loewner-John ellipsoid of the set C_ε, is a minimal-upper bound of the set C_F *w.r.t* to the Loewner partial ordering.

Proof: We will demonstrate this by contradiction. From Lemma 3, we have that $\mathbf{F}_{jl} \succeq \mathbf{F}_i$, $i = 1, \ldots, R$. Assume that there exists a matrix $\mathbf{F}_o \notin C_F$ such that $\mathbf{F}_{jl} \succeq \mathbf{F}_o \succeq \mathbf{F}_i$, therefore $\varepsilon(\mathbf{F}_o) \supseteq \varepsilon(\mathbf{F}_i)$, for $i = 1, \ldots, R$, and thus $\varepsilon(\mathbf{F}_o) \supseteq \cup_{i=1}^{P} \varepsilon(\mathbf{F}_i)$. Given that the volume of $\varepsilon(\mathbf{F}_{jl})$ is less than the volume of $\varepsilon(\mathbf{F}_o)$, since it is the minimum volume hyper-ellipsoid enclosing all $\mathbf{F}_i$, then $|\mathbf{F}_{jl}| \leq |\mathbf{F}_o|$, but by construction $\mathbf{F}_{jl} \succeq \mathbf{F}_o$, thus $|\mathbf{F}_{jl}| \geq |\mathbf{F}_o|$ which is a contradiction. Thus, $\mathbf{F}_o = \mathbf{F}_{jl}$ and $\mathbf{F}_{jl}$ is a minimal-upper bound of the set C_F. ■

Therefore, computing a minimal-upper bound matrix $\mathbf{B}_q$ of the set $C := \{\mathbf{B}\mathbf{B}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha})), \boldsymbol{\alpha} \in S\} \subset \mathbb{S}_{++}^q$ is equivalent to finding the Loewner-John ellipsoid of the set of hyper-ellipsoids associated to C. This is a particular case of a more general problem of computing the minimum volume hyper-ellipsoid $\varepsilon(\mathbf{B}) = \{\mathbf{x}^T\mathbf{B}^{-1}\mathbf{x} + 2(\mathbf{B}^{-1/2}\mathbf{b})^T\mathbf{x} + \mathbf{b}^T\mathbf{b} \leq \mathbf{1}\}$ which covers the union of a set of non centered hyper-ellipsoids parameterized by the quadratic inequalities $\varepsilon_i(\mathbf{B}_i) = \{\mathbf{x}^T\mathbf{B}_i^{-1}\mathbf{x} + 2\mathbf{b}_i^T\mathbf{x} + c_i \leq 0\}$ for $i = 1, \ldots, m$. This problem can be posed as [33] (32), shown at the bottom of the page.

The objective function and the set of constraints are convex, so it can be solved efficiently using semidefinite programming. In particular, we solve this problem using CVX, a package for specifying and solving convex programs [34], [35], for $\mathbf{B}_i = \mathbf{B}\mathbf{B}_q(\mathbf{t}, \mathbf{D}(\boldsymbol{\alpha}_i))$ for $\boldsymbol{\alpha}_i \in S$, $\mathbf{b}_i = \mathbf{b} = \mathbf{0}$, and $c_i = 1$. Therefore, the minimal-upper bound $\mathbf{B}_q$ of the set C is given by $\mathbf{B}_q = \mathbf{B}_*$, where $\mathbf{B}_*$ is the optimal solution of (32). Using the following statement, we can even reduce the number of constraints in the above problem by considering only the set $C_m \subseteq C$ formed by all the maximal elements of C.

Theorem 5: Define C_{Fm} as the subset of C_F formed by all the maximal elements of C_F. Then, the Loewner-John ellipsoid $\varepsilon(\mathbf{F}_{jl})$ of C_ε is also the Loewner-John ellipsoid of the set $C_{\varepsilon m}$ formed by the hyper-ellipsoids associated to the matrices in C_{Fm}.

Proof: Since C_{Fm} is formed by all the maximal elements of C_F, then for $\mathbf{F}_i \in C_{Fm}$ and any $\mathbf{F}_j \in C_{Fc} = C_F - C_{Fm}$, we have that $\mathbf{F}_i \succeq \mathbf{F}_j$. From Lemma 3, $\varepsilon(\mathbf{F}_i) \supseteq \{\varepsilon(\mathbf{F}_j), \text{for all } \mathbf{F}_j \in C_{Fc}\}$, which is true for all $\mathbf{F}_i \in C_{Fm}$, i.e., for all $\varepsilon(\mathbf{F}_i) \in C_{\varepsilon m}$, thus $C_{\varepsilon m} \supseteq \{\varepsilon(\mathbf{F}_j), \text{for all } \mathbf{F}_j \in C_{Fc}\}$ and $C_\varepsilon = C_{\varepsilon m} \cup \{\varepsilon(\mathbf{F}_j), \text{for all } \mathbf{F}_j \in C_{Fc}\} = C_{\varepsilon m}$. Therefore, $\varepsilon(\mathbf{F}_{jl})$ is the Loewner–John ellipsoid for the set C_ε and $C_{\varepsilon m}$. ■

Hence, using the above result we decrease the number of constraints in (32) by performing a pre-step which identifies the set C_m. Note that if C has a greatest element, it is the unique maximal element of C. Therefore, it is the supremum of the set, and its associated hyper-ellipsoid is the Loewner–John ellipsoid of the set of hyper-ellipsoids associated to C. Therefore, there is no need to solve problem (32). Our algorithm searches and removes from the set of constraints the matrices whose hyper-ellipsoid is fully enclosed by other hyper-ellipsoids. In particular, we evaluate in an iterative manner the membership in C_m of all elements in C. We define a membership indicator vector $\mathbf{i}_{C_m}$ where $[\mathbf{i}_{C_m}]_i = I_{C_m}(\mathbf{F}_i)$ and the algorithm begins by assuming that all elements belong to C_m, namely, $\mathbf{i}_{C_m} = \mathbf{1}_{R\times 1}$, where $R = |C|$. Then, all the values of the elements of $\mathbf{i}_{C_m}$ are evaluated using the following iterative procedure.

- *Step 0*: Initialize $\mathbf{i}_{C_m} = \mathbf{1}_{R\times 1}$ and set indexes $k = 1$, $l = 1$.
- *Step 1*: Evaluate membership of $\mathbf{F}_k$ to C_m (if $k > R$, terminate the algorithm):
  $$\text{If } I_{C_m}(\mathbf{F}_k) = \begin{cases} 0, \text{set } k = k+1 \text{ and restart Step 1} \\ 1, \text{ set } l = l+1 \text{ and go to Step 2.} \end{cases}$$
- *Step 2*: Evaluate membership of $\mathbf{F}_l$ to C_m (if $l > R$, set $k = k+1, l = 1$, and go to *Step 1*):
  $$\text{If } I_{C_m}(\mathbf{F}_l) = \begin{cases} 0, \text{ set } l = l+1 \text{ and restart Step 2} \\ 1, \text{go to Step 3.} \end{cases}$$
- *Step 3*: Compare $\mathbf{F}_k$ versus $\mathbf{F}_l$ w.r.t. the Loewner ordering [see the equation shown at the bottom of the page].

$$\max_{\{\mathbf{B},\mathbf{b}\}} \{\log(\det(\mathbf{B}^{1/2}))\} \quad (32)$$
$$\text{subject to } \begin{cases} \tau_1 \geq 0, \tau_2 \geq 0, \ldots, \tau_m \geq 0, \\ \begin{bmatrix} \mathbf{B}^{-1} - \tau_i\mathbf{B}_i^{-1} & \mathbf{B}^{-1/2}\mathbf{b} - \tau_i\mathbf{b}_i \\ (\mathbf{B}^{-1/2}\mathbf{b} - \tau_i\mathbf{b}_i)^T & \mathbf{b}^T\mathbf{b} - 1 - \tau_i c_i \end{bmatrix} \preceq 0, \ i = 1, \ldots, m. \end{cases}$$

$$\begin{cases} \text{If } \mathbf{F}_k \succeq \mathbf{F}_l, \text{set } I_{C_m}(\mathbf{F}_l) = 0, l = l+1, \text{ and go to Step 2} \\ \text{If } \mathbf{F}_l \succeq \mathbf{F}_k, \text{set } I_{C_m}(\mathbf{F}_k) = 0, k = k+1, l = 1, \text{and go to Step 1} \\ \text{If not comparable, set } I_{C_m}(\mathbf{F}_l) = 1, l = l+1, \text{and go to Step 2.} \end{cases}$$

Finally, once the algorithm terminates, the set C_m will be given by all elements such that $I_{C_m}(\mathbf{F}_i) = 1$. To compare $\mathbf{F}_k$ versus $\mathbf{F}_l$, w.r.t. to the Loewner ordering, we apply the determinant test [36] to the matrix, $\mathbf{G} = \mathbf{F}_k - \mathbf{F}_l$. This test evaluates the principal minors of $\mathbf{G}$ and concludes on the matrix definiteness as follows: i) $\mathbf{G}$ is positive definite, i.e., , $\mathbf{F}_k \succ \mathbf{F}_l$, if and only if all its leading principal minors are strictly positive, and it is negative definite, i.e., $\mathbf{F}_l \succ \mathbf{F}_k$, if its kth order leading principal minor is < 0 for k odd and > 0 for k even; ii) $\mathbf{G}$ is positive semidefinite, i.e., $\mathbf{F}_k \succeq \mathbf{F}_l$, if and only if all the principal minors are non-negative, and it is negative semidefinite, i.e., $\mathbf{F}_l \succeq \mathbf{F}_k$, if all the kth-order principal minors are ≤ 0 for k odd and ≥ 0 for k even; iii) $\mathbf{G}$ is indefinite, i.e., $\mathbf{F}_k$ and $\mathbf{F}_l$ are not comparable, if none of the previous conditions are satisfied. Since all the matrices in the set C are block diagonal and the maximum size of one block is 2×2, then every matrix $\mathbf{G}$ is a symmetric tridiagonal matrix, of which the leading principal minors $\{f_{\mathbf{G}}(r), r = 1, \ldots, q\}$ can be computed iteratively as follows [37]:

$$f_{\mathbf{G}}(r) = \begin{cases} 1, & \text{for } r = 0 \\ [\mathbf{G}]_{11}, & \text{for } r = 1 \\ [\mathbf{G}]_{rr} f_{\mathbf{G}}(r-1) - ([\mathbf{G}]_{rr-1})^2 f_{\mathbf{G}}(r-2), & \text{for } 2 < r < q. \end{cases}$$

Note that the determinant of the tridiagonal matrix $\mathbf{G}$ is given by $|\mathbf{G}| = f_{\mathbf{G}}(q)$, and since all the principal minors of $\mathbf{G}$ are also tridiagonal matrices, then their values are computed efficiently using the above expression.

Following the ideas of [24], the issue of having a unique supremum of a the set positive definite matrices can be overcome by redefining the supremum as the matrix associated to the Loewner–John ellipsoid of the set of hyperellipsoids associated to the maximal elements of the set C formed by the P-order BB matrices. This matrix $\mathbf{B}_q$ is unique in the sense that there is no other ellipsoid with minimal volume covering the hyper-ellipsoids associated to the set of maximal elements of C. It also has the properties of continuity, namely, it is positive definite. In the following section we will derive the elements of the Barankin information matrix for the problem of changes in the parameters of Gaussian and Poisson distributions.

IV. Change in Parameters of Gaussian and Poisson Distributions

In this section, we apply the proposed bound for two distributions generally encountered in signal processing. We analyze these two cases in a very general way, which means that the results presented here can be applied to a wide variety of estimation problems. Indeed, the parameters involved in the Gaussian distribution (mean and covariance) and in the Poisson distribution are assumed to be a function of the parameters $\boldsymbol{\eta}_j$, which generally represent physical parameters of interest in signal processing. An example of change of parameters in a Gaussian distribution in the radar context is direction-of-arrival (DOA) estimation. The varying cross section fluctuations are modeled with a Swerling 0 model [38], where the DOAs are hidden in the mean of the observations, leading, for example, to the so-called conditional MLE [39]. On the other hand, when the emitted signals are modeled with a Swerling 1–2, the DOAs are hidden in the covariance of the observations, leading, for example, to the so-called unconditional MLE [40]. In the context of particle detection, the Poisson distribution is generally used to model the particle counting process; i.e., the observations and the parameter involved in the Poisson distribution become a function of the DOA [41].

A. Gaussian Case

Let us assume that the vector of observations $\mathbf{x}_i \in \mathbb{R}^M$, for $i = 1, \ldots, N$, is modeled as $\mathbf{x}_i = \mathbf{f}(\boldsymbol{\nu}_j) + \mathbf{n}_i$, where $\mathbf{f}(\cdot)$ is a vector of known functions; $\boldsymbol{\nu}_j \in \mathbb{R}^F$ is a known parameter vector; $\mathbf{n}_i$ is a zero-mean Gaussian random vector with covariance matrix $\mathbf{M}(\boldsymbol{\varphi}_j)$, with $\mathbf{M}(\cdot)$ a symmetric positive definite matrix of known functions; and $\boldsymbol{\varphi}_j \in \mathbb{R}^G$ is a known parameter vector. Then $\boldsymbol{\eta}_j = [\boldsymbol{\nu}_j^T, \boldsymbol{\varphi}_j^T]^T \in \mathbb{R}^L$, with $L = F + G$, and $\mathbf{x}_i$ are distributed as $\mathcal{N}(\mathbf{f}(\boldsymbol{\nu}_j), \mathbf{M}(\boldsymbol{\varphi}_j))$. Here, we are interested in deriving the elements of the Barankin information matrix for changes in the pdf parameters of $\mathbf{x}_i$, i.e., the mean and covariance matrix. First, we analyze the general case of piecewise changes of the mean and covariance. Second, we deduce two particular cases: i) piecewise changes of mean and constant covariance matrix, i.e., $\mathbf{M}(\boldsymbol{\varphi}_j) = \mathbf{M}(\boldsymbol{\varphi}) = \boldsymbol{\Sigma}$, and ii) piecewise changes of covariance and constant mean vector, i.e., $\mathbf{f}(\boldsymbol{\nu}_j) = \mathbf{f}(\nu) = \mu$. Note that we restrict our analysis to the set of parameter vectors $\{\boldsymbol{\nu}_j\}$ and $\{\boldsymbol{\varphi}_j\}$ such that the functions in $\mathbf{f}(\boldsymbol{\nu}_j)$ and $\mathbf{M}(\boldsymbol{\varphi}_j)$ are injective. In other words, a change in the values of $\boldsymbol{\nu}_j$ changes the values of $\mathbf{f}(\boldsymbol{\nu}_j)$, the mean of the distribution of $\mathbf{x}_i$. Similarly, a change in the values of $\boldsymbol{\varphi}_j$ implies a change in values of the covariance matrix $\mathbf{M}(\boldsymbol{\varphi}_j)$. Below, we compute the elements of the Barankin information matrix $\mathbf{BIM_t}$. Then, for each case, respectively, we derive closed-form expressions for the elements $\boldsymbol{\Phi} - \mathbf{1}_{q \times q}$ (see Appendix D for details on their derivation) which are different from zero; namely, we evaluate $[\boldsymbol{\Phi}]_{kk}$ for $\alpha_k > 0, \alpha_k < 0$, and $[\boldsymbol{\Phi}]_{kk+1}$for $t_k + \alpha_k > t_{k+1} + \alpha$.

1) Piecewise Changes of Mean and Covariance Matrix: For $\alpha_k > 0$, using (13), we have that $[\boldsymbol{\Phi}]_{kk}$ is given by (33), shown at the bottom of the page, where $\mathbf{M}_k = (2(\mathbf{M}(\boldsymbol{\varphi}_k))^{-1} - (\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1})$ and $\mathbf{g}_k = 2(\mathbf{M}(\boldsymbol{\varphi}_k))^{-1}\mathbf{f}(\boldsymbol{\nu}_k) - (\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+1})$. For $\alpha_k < 0$, using (13), we have

$$[\boldsymbol{\Phi}]_{kk} = \begin{cases} \left(\frac{|\mathbf{M}(\boldsymbol{\varphi}_{k+1})|^{1/2} |\mathbf{M}_k^{-1}|^{1/2}}{|\mathbf{M}(\boldsymbol{\varphi}_k)|} \right)^{\alpha_k} \exp\left\{ \frac{\alpha_k}{2} \mathbf{g}_k^T \mathbf{M}_k^{-1} \mathbf{g}_k - \alpha_k \mathbf{f}^T(\boldsymbol{\nu}_k)(\mathbf{M}(\boldsymbol{\varphi}_k))^{-1}\mathbf{f}(\boldsymbol{\nu}_k) \right\} \\ \times \exp\left\{ \frac{\alpha_k}{2} \mathbf{f}^T(\boldsymbol{\nu}_{k+1})(\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+1}) \right\}, \text{ for } \mathbf{M}_k \in \mathbb{S}_{++}^M \\ \infty, \text{ otherwise} \end{cases} \tag{33}$$

that $[\mathbf{\Phi}]_{kk}$ is given by (34), shown at the bottom of the page, where $\mathbf{M}_{k+1} = 2(\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1} - (\mathbf{M}(\boldsymbol{\varphi}_k))^{-1}$ and $\mathbf{g}_{k+1} = 2(\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+1}) - (\mathbf{M}(\boldsymbol{\varphi}_k))^{-1}\mathbf{f}(\boldsymbol{\nu}_k)$. For $t_k + \alpha_k > t_{k+1} + \alpha_{k+1}$, using (17), we have that $[\mathbf{\Phi}]_{kk+1}$ is given as (35), shown at the bottom of the page, where $\bar{\mathbf{M}}_k = (\mathbf{M}(\boldsymbol{\varphi}_k))^{-1} + (\mathbf{M}(\boldsymbol{\varphi}_{k+2}))^{-1} - (\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1}$, and $\bar{\mathbf{g}}_k = (\mathbf{M}(\boldsymbol{\varphi}_k))^{-1}\mathbf{f}(\boldsymbol{\nu}_k) + (\mathbf{M}(\boldsymbol{\varphi}_{k+2}))^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+2}) - (\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+1})$.

2) Piecewise Changes of Mean and Constant Covariance Matrix: In this case $\mathbf{M}(\boldsymbol{\varphi}_j) = \mathbf{M}(\boldsymbol{\varphi}) = \mathbf{\Sigma}, \boldsymbol{\eta}_j = [\boldsymbol{\nu}_j^T, \boldsymbol{\varphi}^T]^T$, and $[\mathbf{\Phi}]_{kl}$ is given as follows.

For $\alpha_k > 0$, using (33) and replacing $\mathbf{M}(\boldsymbol{\varphi}_k)$ and $\mathbf{M}(\boldsymbol{\varphi}_{k+1})$ by $\mathbf{\Sigma}$, we have straightforwardly for $[\mathbf{\Phi}]_{kk}$:

$$[\mathbf{\Phi}]_{kk} = \exp\{\alpha_k(\mathbf{f}(\boldsymbol{\nu}_k) - \mathbf{f}(\boldsymbol{\nu}_{k+1}))^T\mathbf{\Sigma}^{-1}(\mathbf{f}(\boldsymbol{\nu}_k) - \mathbf{f}(\boldsymbol{\nu}_{k+1}))\}. \tag{36}$$

For $\alpha_k < 0$, using (34), $[\mathbf{\Phi}]_{kk}$ is given as follows:

$$[\mathbf{\Phi}]_{kk} = \exp\{-\alpha_k(\mathbf{f}(\boldsymbol{\nu}_{k+1}) - \mathbf{f}(\boldsymbol{\nu}_k))^T \times \mathbf{\Sigma}^{-1}(\mathbf{f}(\boldsymbol{\nu}_{k+1}) - \mathbf{f}(\boldsymbol{\nu}_k))\}. \tag{37}$$

For $t_k + \alpha_k > t_{k+1} + \alpha_{k+1}$, using (35), then$[\mathbf{\Phi}]_{kk+1}$is given as follows:

$$[\mathbf{\Phi}]_{kk+1} = \left(\exp\left\{\frac{\beta_k}{2}((\mathbf{f}(\boldsymbol{\nu}_{k+1}) - \mathbf{f}(\boldsymbol{\nu}_k))\mathbf{\Sigma}^{-1}(\mathbf{f}(\boldsymbol{\nu}_{k+1}) - \mathbf{f}(\nu_k))^T + (\mathbf{f}(\boldsymbol{\nu}_{k+2}) - \mathbf{f}(\nu_{k+1}))\mathbf{\Sigma}^{-1}(\mathbf{f}(\boldsymbol{\nu}_{k+2}) - \mathbf{f}(\nu_{k+1}))^T - (\mathbf{f}(\boldsymbol{\nu}_k) - \mathbf{f}(\nu_{k+2}))\mathbf{\Sigma}^{-1}(\mathbf{f}(\boldsymbol{\nu}_k) - \mathbf{f}(\nu_{k+2}))^T)\right\}\right). \tag{38}$$

3) Piecewise Changes of Covariance Matrix and Constant Mean Vector: In this case $\mathbf{f}(\boldsymbol{\nu}_j) = \mathbf{f}(\nu) = \mu, \boldsymbol{\eta}_j = [\boldsymbol{\nu}^T, \boldsymbol{\varphi}_j^T]^T$, and $[\mathbf{\Phi}]_{kl}$ is given as follows.

For $\alpha_k > 0$ using (33) and replacing $\mathbf{f}(\boldsymbol{\nu}_k)$ and $\mathbf{f}(\boldsymbol{\nu}_{k+1})$ by $\boldsymbol{\mu}$, we have straightforwardly for $[\mathbf{\Phi}]_{kk}$:

$$[\mathbf{\Phi}]_{kk} = \begin{cases} \left(\frac{|\mathbf{M}(\boldsymbol{\varphi}_{k+1})|^{1/2}}{|\mathbf{M}(\boldsymbol{\varphi}_k)||\mathbf{M}_k|^{1/2}}\right)^{\alpha_k}, \text{ for } \mathbf{M}_k \in \mathbb{S}_{++}^M \\ \infty, \text{otherwise} \end{cases} \tag{39}$$

where $\mathbf{M}_k = 2(\mathbf{M}(\boldsymbol{\varphi}_k))^{-1} - (\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1}$.

For $\alpha_k < 0$, using (34), $[\mathbf{\Phi}]_{kk}$ is given as follows:

$$[\mathbf{\Phi}]_{kk} = \begin{cases} \left(\frac{|\mathbf{M}(\boldsymbol{\varphi}_k)|^{1/2}}{|\mathbf{M}(\boldsymbol{\varphi}_{k+1})||\mathbf{M}_{k+1}|^{1/2}}\right)^{-\alpha_k}, \text{ for } \mathbf{M}_{k+1} \in \mathbb{S}_{++}^M \\ \infty, \text{otherwise} \end{cases} \tag{40}$$

where $\mathbf{M}_{k+1} = 2(\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1} - (\mathbf{M}(\boldsymbol{\varphi}_k))^{-1}$.

For $t_k + \alpha_k > t_{k+1} + \alpha_{k+1}$, using (35), then $[\mathbf{\Phi}]_{kk+1}$is given as follows:

$$[\mathbf{\Phi}]_{kk+1} = \begin{cases} \left(\frac{|\mathbf{M}(\boldsymbol{\varphi}_{k+1})|^{1/2}}{|\mathbf{M}(\boldsymbol{\varphi}_k)|^{1/2}|\mathbf{M}(\boldsymbol{\varphi}_{k+2})|^{1/2}|\bar{\mathbf{M}}_k|^{1/2}}\right)^{\beta_k}, & \text{for } \bar{\mathbf{M}}_k \in \mathbb{S}_{++}^M \\ \infty, \text{otherwise} \end{cases} \tag{41}$$

where $\bar{\mathbf{M}}_k = (\mathbf{M}(\boldsymbol{\varphi}_k))^{-1} + (\mathbf{M}(\boldsymbol{\varphi}_{k+2}))^{-1} - (\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1}$.

The elements of Barankin bound for each case are obtained by using (31), recalling that $A_k = [\mathbf{\Phi}]_{kk} - 1$ and $B_k = [\mathbf{\Phi}]_{kk+1} - 1$, from (19) and (20), respectively.

B. Poisson Case

Assume that the measurements $x_i \in \mathbb{N} + \{0\}$, for $i = 1, \ldots, N$, are distributed as a Poisson distribution with parameter $f(\boldsymbol{\eta}_j)$, where $f(\cdot)$ is a known function and $\boldsymbol{\eta}_j \in \mathbb{R}^L$ is a known parameter vector. Similarly to the Gaussian case, we restrict our analysis to the set of parameter vectors $\{\boldsymbol{\eta}_j\}$ such that the function $f(\boldsymbol{\eta}_j)$ is injective. Therefore, we derive closed-form expressions for the elements of the matrix $\mathbf{\Phi} - \mathbf{1}_{q\times q}$ for piecewise changes of the parameter $\boldsymbol{\eta}_j$. Below, we evaluate $[\mathbf{\Phi}]_{kk}$ for $\alpha_k > 0$ and $\alpha_k < 0$, and $[\mathbf{\Phi}]_{kk+1}$ for $t_k + \alpha_k > t_{k+1} + \alpha_{k+1}$. Note that since $x_i \in \mathbb{N}$ we replace the integral operator by the summation operator.

For $\alpha_k > 0$, $[\mathbf{\Phi}]_{kk}$ becomes

$$[\mathbf{\Phi}]_{kk} = \exp\left\{\frac{\alpha_k(f(\boldsymbol{\eta}_{k+1}) - f(\boldsymbol{\eta}_k))^2}{f(\boldsymbol{\eta}_{k+1})}\right\}. \tag{42}$$

For $\alpha_k < 0$, $[\mathbf{\Phi}]_{kk}$ becomes

$$[\mathbf{\Phi}]_{kk} = \exp\left\{\frac{-\alpha_k(f(\boldsymbol{\eta}_k) - f(\boldsymbol{\eta}_{k+1}))^2}{f(\boldsymbol{\eta}_k)}\right\}. \tag{43}$$

For $t_k + \alpha_k > t_{k+1} + \alpha_{k+1}$, $[\mathbf{\Phi}]_{kk+1}$ is given as follows:

$$[\mathbf{\Phi}]_{kk+1} = \exp\left\{\beta_k\left(\frac{(f(\boldsymbol{\eta}_{k+1}) - f(\boldsymbol{\eta}_k))^2}{2f(\boldsymbol{\eta}_{k+1})} + \frac{(f(\boldsymbol{\eta}_{k+2}) - f(\boldsymbol{\eta}_{k+1}))^2}{2f(\boldsymbol{\eta}_{k+1})} - \frac{(f(\boldsymbol{\eta}_k) - f(\boldsymbol{\eta}_{k+2}))^2}{2f(\boldsymbol{\eta}_{k+1})}\right)\right\}. \tag{44}$$

$$[\mathbf{\Phi}]_{kk} = \begin{cases} \left(\frac{|\mathbf{M}(\boldsymbol{\varphi}_k)|^{1/2}|\mathbf{M}_{k+1}^{-1}|^{1/2}}{|\mathbf{M}(\boldsymbol{\varphi}_{k+1})|}\right)^{-\alpha_k} \exp\left\{\frac{-\alpha_k}{2}\mathbf{g}_{k+1}^T\mathbf{M}_{k+1}^{-1}\mathbf{g}_{k+1} + \alpha_k\mathbf{f}^T(\boldsymbol{\nu}_{k+1})(\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+1})\right\} \\ \times \exp\left\{\frac{-\alpha_k}{2}\mathbf{f}^T(\boldsymbol{\nu}_k)(\mathbf{M}(\boldsymbol{\varphi}_k))^{-1}\mathbf{f}(\boldsymbol{\nu}_k)\right\}, \text{ for } \mathbf{M}_{k+1} \in \mathbb{S}_{++}^M \\ \infty, \text{ otherwise} \end{cases} \tag{34}$$

$$[\mathbf{\Phi}]_{kk+1} = \begin{cases} \left(\frac{|\mathbf{M}(\boldsymbol{\varphi}_{k+1})|^{1/2}|\bar{\mathbf{M}}_k^{-1}|^{1/2}}{|\mathbf{M}(\boldsymbol{\varphi}_k)|^{1/2}|\mathbf{M}(\boldsymbol{\varphi}_{k+2})|^{1/2}}\right)^{\beta_k} \exp\left\{\frac{\beta_k}{2}\bar{\mathbf{g}}_k^T\bar{\mathbf{M}}_k^{-1}\bar{\mathbf{g}}_k^T - \frac{\beta_k}{2}\mathbf{f}^T(\boldsymbol{\nu}_k)(\mathbf{M}(\boldsymbol{\varphi}_k))^{-1}\mathbf{f}(\boldsymbol{\nu}_k)\right\} \\ \times \exp\left\{-\frac{\beta_k}{2}\mathbf{f}^T(\boldsymbol{\nu}_{k+2})(\mathbf{M}(\boldsymbol{\varphi}_{k+2}))^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+2}) + \frac{\beta_k}{2}\mathbf{f}^T(\boldsymbol{\nu}_{k+1})(\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+1})\right\}, \text{ for } \bar{\mathbf{M}}_k^{-1} \in \mathbb{S}_{++}^M \\ \infty, \text{ otherwise} \end{cases} \tag{35}$$

Similarly, as in the Gaussian case, the elements of the Barankin bound for each case are obtained by using (31) with $A_k = [\mathbf{\Phi}]_{kk} - 1$ and $B_k = [\mathbf{\Phi}]_{kk+1} - 1$.

V. Numerical Examples

In this section, as an illustration, we compare the MSE between the true values of the change-point locations and their maximum likelihood estimations with our bounds. In particular, we first introduce the MLE of change-point locations assuming the total number of changes is known. Then we analyze the cases of multiple changes in i) the mean of a Gaussian distribution with fixed variance, ii) the variance of a Gaussian assuming a fixed mean, and iii) the mean rate of a Poisson distribution.

A. Maximum Likelihood Estimation

The MLE of $\mathbf{t}$ is the solution to the following problem:

$$\hat{\mathbf{t}}_{\mathrm{ML}} = \arg\max_{\mathbf{t}} \sum_{i=1}^{q+1} \ln p_i(\mathbf{x}_{t_{i-1}+1}, \ldots, \mathbf{x}_{t_i}; \boldsymbol{\eta}_i) \qquad (45)$$

where $t_0 = 0$ and $t_{q+1} = N$ by definition. There is no known closed-form expression for $\hat{\mathbf{t}}_{\mathrm{ML}}$ so it has to be estimated via numerical computations. To solve this multidimensional optimization problem efficiently, we apply dynamic programming (DP), explained in detail in [42], in our context of change-point estimation. The main advantage of the DP approach is that it does not need to evaluate all the possible combinations of values for $\mathbf{t}$ in (45). In all our examples below, we illustrate the average MSE performance of the MLE for 1000 Monte Carlo experiments. We studied the performance as a function of the signal-to-noise ratio (SNR), which is defined accordingly in each example, and as a function of the distance between change points. Here we chose $q = 3$ and the number of samples $N = 80$. In each example below, we set $t_2 = 40$ and $t_3 = 60$, and we analyze two scenarios for change point t_1: In the first one, we set $t_1 = 20$ such that each segment has the same number of samples, and in the second scenario, $t_1 \in [2, 38]$. Note that the unbiasedness properties of the MLE have been studied in [43] for a single change-point and for multiple change-points in [44]. The asymptotic results derived in [43] and [44] are applicable only for the case of a Gaussian distribution with changes in the mean. However, in the case of having a finite interval the MLE is expected to be biased independently of the distribution. On the other hand, it seems reasonable to assume that for large SNR values the MLE is unbiased for a subset of the parameter space, i.e., subintervals, and specially for change-points located equidistant from their neighboring change-points or the interval limits. For example, in all the examples below, the bias of the MLE for $\mathbf{t} = [20\ 40\ 60]^T$ is approximately zero for all the SNR ranges considered in each scenario.

B. Changes in the Mean of a Gaussian Distribution

We consider the scenario of a time series with three change points in the mean values of a Gaussian distribution with common variance. We recall the closed-form expressions obtained for computing $[\mathbf{\Phi}]_{kk}$, namely, (36) and (37), and define the SNR for the kth change point as

$$\mathrm{SNR}_k = (\mathbf{f}(\boldsymbol{\nu}_{k+1}) - \mathbf{f}(\boldsymbol{\nu}_k))^T \Sigma^{-1} (\mathbf{f}(\boldsymbol{\nu}_{k+1}) - \mathbf{f}(\boldsymbol{\nu}_k)) \qquad (46)$$

where $\mathbf{f}(\boldsymbol{\nu}_k) \in \mathbb{R}^M$ is the mean vector of the kth segment and $\mathbf{\Sigma} \in \mathbb{R}^{M \times M}$ is the common covariance matrix. In our example, $M = 1$ and, without loss of generality, we choose $f(\boldsymbol{\nu}_k) = \nu_k$ and $\Sigma = \sigma^2 = 1$, thus $\boldsymbol{\eta}_k = [\boldsymbol{\nu}_k^T, 1]^T$. Here, we set $\nu_1 = 1$, and ν_2, ν_3, and ν_4 are set such that $\mathrm{SNR}_1 = \mathrm{SNR}_2 = \mathrm{SNR}_3 = \mathrm{SNR}$. In particular, $\nu_k = \nu_{k-1} + (-1)^k \sqrt{\sigma^2 \mathrm{SNR}}$ for $k = 2, 3, 4$. Fig. 1(a) illustrates the mean values as a function of sample time for different SNR values. In Fig. 1(c), we illustrate the MSE performance of the MLE for the change-point vector, and the BB as a function of the SNR. In particular, $\mathrm{MSE}_{\mathrm{known}}$ is the MSE performance of the MLE for the change-point vector, assuming knowledge of the means and variance. $\mathrm{MSE}_{\mathrm{unknown}}$ is the MSE performance of the MLE for a more realistic case when no knowledge of the distribution parameters is available. The $\mathbf{BB}_{\mathrm{sup}}$ is given by the minimal-upper bound matrix $\mathbf{B}_q$ of the set C computed using the algorithm presented in Section III-D, and $\mathbf{BB}_{\mathrm{tr}}$ is the matrix in C that has the maximum trace. We illustrate the trace of $\mathbf{BB}_{\mathrm{sup}}$ and $\mathbf{BB}_{\mathrm{tr}}$ since we are comparing the MSE performance for the change-point vector estimates. Note that, in view of the discussion presented in Section III-D, we compute$\mathbf{BB}_{\mathrm{tr}}$ only in this example to show that $\mathbf{BB}_{\mathrm{tr}}$ does not necessarily coincide with supremum of the set unless $\mathbf{BB}_{\mathrm{sup}} \in C$. In this particular scenario, we found that $\mathbf{BB}_{\mathrm{sup}}$ belongs to the set C for SNR values equal to and larger than 2 dB. Therefore, we have the optimal test points $\{\alpha_1^*, \alpha_2^*, \alpha_3^*\}$ associated to the matrix $\mathbf{BB}_{\mathrm{sup}}$ defining the Loewner-John Ellipsoid, which are presented in Fig. 1(b). For SNR values above 2 dB, no change point is overlapped, therefore, each bound depends only on its corresponding diagonal element $[\mathbf{\Phi}]_{ii}$, which is equivalent to the resulting analysis of considering one change point located at $t = 20$, assuming $N = 40$. Moreover, it is important to mention that in this example, $[\mathbf{\Phi}]_{ii}$ is symmetric with respect to α_i, and since all segments have the same length, then both α_i and $-\alpha_i$ are optimal solutions for the bound on t_i. In Fig. 1(b), we illustrate only one optimal solution. When the $\mathrm{SNR} > 2$ dB, we found the set C had several maximal elements that were not mutually comparable, thus, $\mathbf{BB}_{\mathrm{sup}} \notin C$ and does not show up in Fig. 1(b). Finally, it can be seen that the test point approached the true change point values as the SNR increases; i.e., α_2 tends to -1 as SNR increases.

In Fig. 1(d), we illustrate the $\mathrm{MSE}_{\mathrm{known}}$ and $\mathbf{BB}_{\mathrm{sup}}$ for change-point $t_i, i = 1, 2, 3$ as a function of SNR. It is noteworthy to mention that we did not illustrate the performance for higher SNR range in this example, since we found that for SNR values larger than 10 dB the bound tends quickly to zero. On the other hand, computing MSE values in these examples for larger SNR requires a large number of Monte Carlo simulations since the higher the SNR, the smaller the probability of an error. For example, a single realization with an error of only 1 unit in one of the change-points, among 1000 realizations in the Monte Carlo simulation, amounts to an MSE of -30 dB. Similar observations hold for the example of changes in the mean rate of a Poisson distribution.

We also analyze the MSE performance as a function of the distance between change points for a fixed SNR value. In Fig. 1(e), for $\mathrm{SNR} = -6$ dB, we illustrate the diagonal elements of $\mathbf{BB}_{\mathrm{sup}}$ and the MSE of the MLE for the change-point

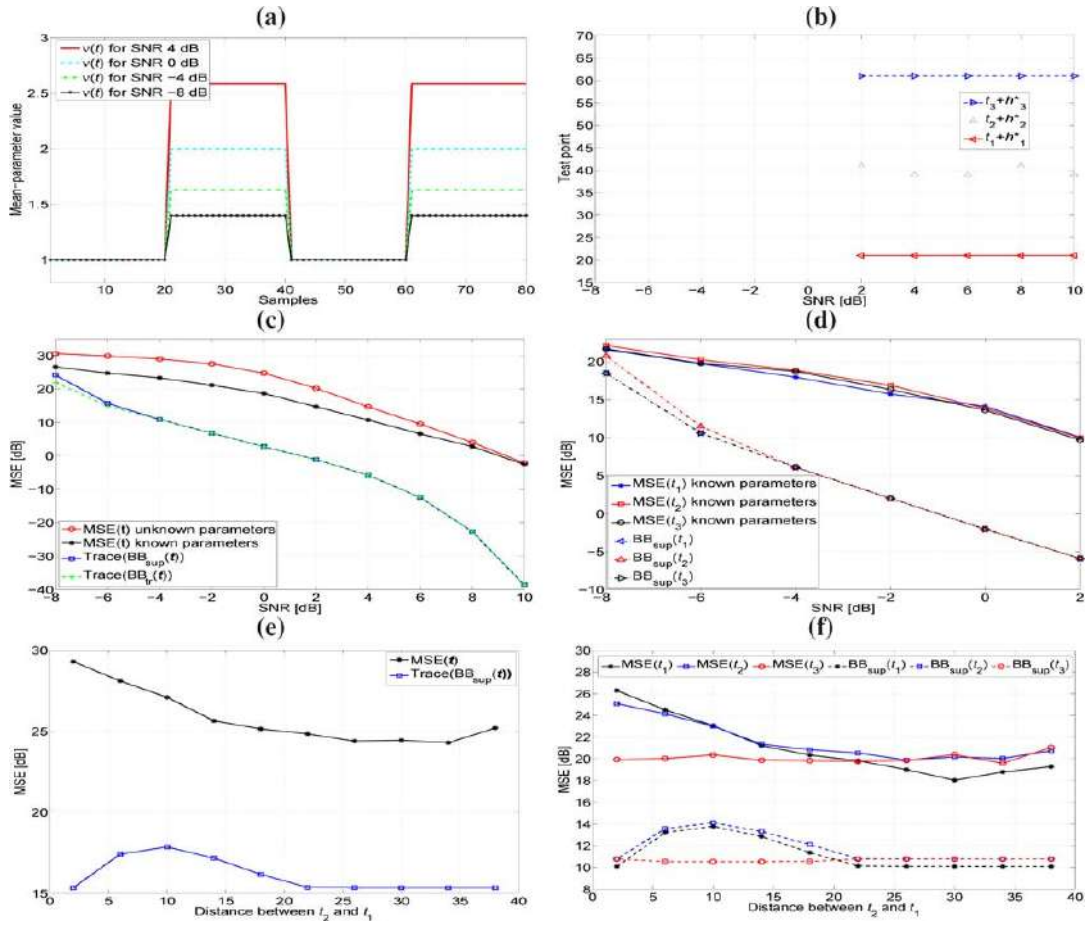


Fig. 1. Performance analysis for estimating change-points of the mean in a Gaussian distribution: (a) Mean values as a function of sample time for different SNR values; (b) Test points associated with the BB given by the minimal-upper bound of C, $\mathbf{BB}_{\sup}$, as a function of SNR; (c) MSE of the change-point vector using the MLE of $\mathbf{t}$ and its Barankin bound given by $\mathbf{BB}_{\sup}$, and by the matrix with maximum trace in C, $\mathbf{BB}_{\mathrm{tr}}$; (d) MSE of each change-point as a function of SNR using the MLE of t_1, t_2, and t_3 and their corresponding Barankin bound $\mathbf{BB}_{\sup}(t_i)$, $i = 1, \ldots, 3$; (e) MSE of change-point vector using the MLE of $\mathbf{t}$ and its Barankin bound, $\mathbf{BB}_{\sup}(\mathbf{t})$, as a function of the distance between t_2 and t_1 for $\mathrm{SNR} = -6$ [dB]; (f) MSE of each change-point and their respective $\mathbf{BB}_{\sup}$ as a function of the distance between t_2 and t_1 for $\mathrm{SNR} = -6$ [dB].

vector $\mathbf{t}$, assuming knowledge of the distribution parameters, as a function of the distance between change points t_1 and t_2. In Fig. 1(f), we illustrate the BB and the MSE of the MLE for each change-point. We observe that the MSE of the MLE for t_1 and t_2 increases as the distance between change points t_1 and t_2 decreases. Similarly, their respective BB predict the same behavior for distances between t_1 and t_2 equal to and larger than 10 time-units; however, for distances smaller than 10 time-units their respective bounds decrease to the same value, as they did for distances larger than 22 time-units. This bound behavior is expected to take place as our Barankin-type lower bound approximation considers only one change-point per parameter. Therefore, in our problem the test-point values are lower and upper bounded by the adjacent change-point parameters, which does not allow for evaluating errors, in estimating each change-point, beyond these limits. Thus, as the change-points get closer, the test-point domains become limited, and the bound cannot take into account either estimated errors given by estimates of t_1 which are larger than the true value of t_2, or estimated errors given by estimates of t_2 which are lower than the true value of t_1.

C. Changes in the Variance of a Gaussian Distribution

We consider the same scenarios as above, but with a time series with three change points in the variance of a Gaussian distribution and a common mean. We recall the closed-form expressions obtained for computing $[\mathbf{\Phi}]_{kk}$, namely, (39) and (40), and define, the SNR for the kth change point as $\mathrm{SNR}_k = (|\mathbf{M}(\boldsymbol{\varphi}_{k+1})|)/(|\mathbf{M}(\boldsymbol{\varphi}_k)|)$, where $\mathbf{M}(\boldsymbol{\varphi}_k) \in \mathbb{R}^{M\times M}$ is the covariance matrix of the kth segment. In our example, $M = 1$, and, without loss of generality, we choose $\mathbf{M}(\boldsymbol{\varphi}_k) = \varphi_k$, and the mean equal to zero since the BIM does not depend on the mean, thus $\boldsymbol{\eta}_k = [0, \varphi_k]^T$. Here, we set $\varphi_1 = 1$, and variances φ_2, φ_3, and φ_4 are set such that $\mathrm{SNR}_1 = \mathrm{SNR}_2 = \mathrm{SNR}_3 = \mathrm{SNR}$. In

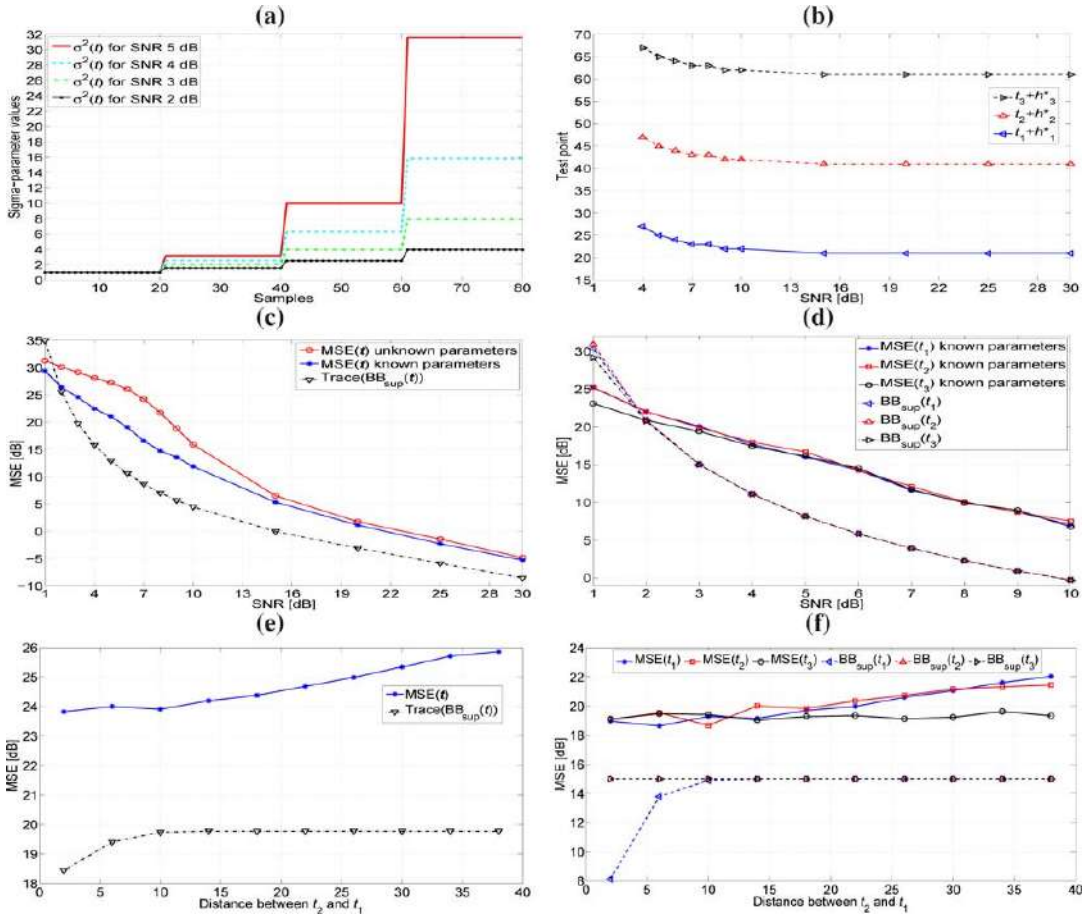


Fig. 2. Performance analysis for estimating change-points of the variance in a Gaussian distribution: (a) Sigma-parameter values as a function of sample time for different SNR values; (b) Test points associated with the BB given by the minimal-upper bound of C, $\mathbf{BB}_{\sup}$, as a function of SNR; (c) MSE of the change-point vector using the MLE of $\mathbf{t}$ and its Barankin bound given by $\mathbf{BB}_{\sup}$; (d) MSE of each change-point as a function of SNR using the MLE of t_1, t_2, and t_3 and their corresponding Barankin bound $\mathbf{BB}_{\sup}(t_i)$, $i = 1, \ldots, 3$; (e) MSE of change-point vector using the MLE of $\mathbf{t}$ and its Barankin bound, $\mathbf{BB}_{\sup}(\mathbf{t})$, as a function of the distance between t_2 and t_1 for $\mathrm{SNR} = \mathbf{4}$ [dB]; (f) MSE of each change-point and their respective $\mathbf{BB}_{\sup}$ as a function of the distance between t_2 and t_1 for $\mathrm{SNR} = 4$ [dB].

practice, $\varphi_k = \varphi_{k-1}\mathrm{SNR}$. In Fig. 2(a), we illustrate sigma-parameter values as a function of sample time for different SNR values. In Fig. 2(c), we illustrate the MSE performance of the MLE for the change-point vector as a function of SNR and its respective Barankin bound, $\mathbf{BB}_{\sup}$. In particular, we illustrate the $\mathrm{MSE}_{\mathrm{unknown}}$ and $\mathrm{MSE}_{\mathrm{known}}$ of $\mathbf{t}$ for SNR ranging from 1 to 30 dB. In Fig. 2(d), we focus on SNR ranging between 1 to 10 dB, and we illustrate the MSE for change-point estimate of t_1, t_2, and t_3, using the MLE and their respective bounds given by the diagonal elements of $\mathbf{BB}_{\sup}$. In this scenario $\mathbf{BB}_{\sup}$ belongs to set C for SNR values larger than 4 dB, and the MSE of the MLE slowly approaches the BB as the SNR increases. In this example, the BB is the same for all change-points for SNR values above 2 dB, and for all the SNR ranges illustrated, the maximum differences between the BB and both the $\mathrm{MSE}_{\mathrm{known}}$ and $\mathrm{MSE}_{\mathrm{unknown}}$ are approximately 7 and 17 dB, respectively. For SNR values lower than 2 dB, the $\mathbf{BB}_{\sup}$ is greater than the MSE of the MLE because the Barankin bound derivation does not consider the set of admissible values of the estimator. In our example, the MLE computation restricts the search to the range between 1 and N, and thus the MLE variance has an upper limit, which the BB computation does not consider. Moreover, the BB assumes that the estimator is unbiased at the test-points; thus for low SNR the comparison against the MLE's MSE is inappropriate because the optimal test-points tend to go to the extreme of the intervals associated to each change-point causing some bias. Fig. 2(b) illustrates the optimal test points $[\alpha_1^*, \alpha_2^*, \alpha_3^*]^T$ associated to the matrix $\mathbf{BB}_{\sup}$. It can be seen that for all the SNR range there are no overlaps between test points and, as in the previous example, all test points approach to 1 or -1, namely, they are close to the true change-point values as SNR increases. Therefore, for large SNR values $[\mathbf{BB}_{\sup}]_{kk} = (\sqrt{2\mathrm{SNR}_k - 1})/(\mathrm{SNR}_k)$, which tends to 0 as $\mathrm{SNR}_k \to \infty$.

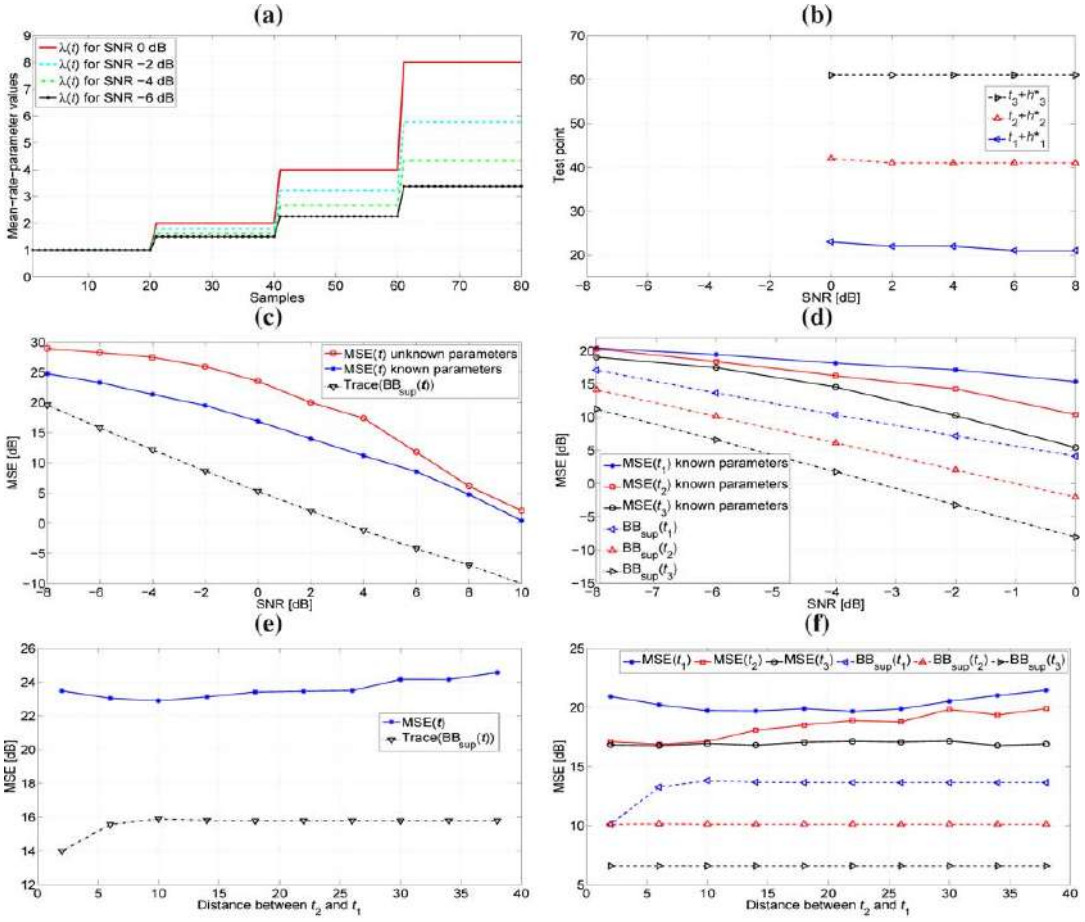


Fig. 3. Performance analysis for estimating change-points in the mean rate of a Poisson distribution: (a) Mean-rate-values as a function of sample time for different SNR values; (b) Test points associated with the BB given by the minimal-upper bound of C, $\mathbf{BB}_{\sup}$, as a function of SNR; (c) MSE of the change-point vector using the MLE of $\mathbf{t}$ and its Barankin bound given by $\mathbf{BB}_{\sup}$; (d) MSE of each change-point as a function of SNR using the MLE of t_1, t_2, and t_3 and their corresponding Barankin bound $\mathbf{BB}_{\sup}(t_i)$, $i = 1, \ldots, 3$; (e) MSE of change-point vector using the MLE of $\mathbf{t}$ and its Barankin bound, $\mathbf{BB}_{\sup}(\mathbf{t})$, as a function of the distance between t_2 and t_1 for $\mathrm{SNR} = -6$ [dB]; (f) MSE of each change-point and their respective $\mathbf{BB}_{\sup}$ as a function of the distance between t_2 and t_1 for $\mathrm{SNR} = -6$ [dB].

In Fig. 2(e) and (f), for $\mathrm{SNR} = 4$ dB, we illustrate the BB and the MSE of the MLE for t_1, t_2, and t_3, assuming knowledge of the distribution parameters, as a function of the distance between change points t_1 and t_2. Above 10 units, the BB for all the change-points remains the same for distances between change-points t_1 and t_2. The BB for t_1 increases as the distance between change-points t_1 and t_2 increases from zero to 10 units. As in the previous example, the bound in this range is overly optimistic since the test-point domains become limited.

D. Changes in the Mean Rate of a Poisson Distribution

Now we consider a time series with three change points in the mean rate of a Poisson distribution. As in the previous examples, we recall the closed-form expressions for $[\mathbf{\Phi}]_{kk}$, i.e., (42) and (43). Then we define the SNR for the kth change point detector as $\mathrm{SNR}_k = ((f(\boldsymbol{\eta}_k) - f(\boldsymbol{\eta}_{k+1}))^2)/(f(\boldsymbol{\eta}_k)^2)$, where $f(\boldsymbol{\eta}_k)$ is the mean rate of the kth segment. Here, without loss of generality, we set $f(\boldsymbol{\eta}_k) = \eta_k$. The mean rate is set to $\eta_1 = 1$, and the mean rates η_2, η_3, and, η_4 are set such that $\mathrm{SNR}_1 = \mathrm{SNR}_2 = \mathrm{SNR}_3 = \mathrm{SNR}$. In practice, $\eta_k = \eta_{k-1}(1 + \sqrt{\mathrm{SNR}})$. In Fig. 3(a), we illustrate the mean-rate-values as a function of sample time for different SNR values. Fig. 3(c) and (d) illustrates the $\mathrm{MSE}_{\mathrm{unknown}}$ and $\mathrm{MSE}_{\mathrm{known}}$ performance for the change-point vector and change points t_1, t_2, and t_3, respectively. In this case, the MSE values, as well as the bounds for t_1, t_2, and t_3, are not the same for the same SNR values. In fact, it can be seen that the MSE values for t_3 are lower than the MSE values for t_2, and these last are lower than the MSE values for t_1. This difference in performance is due to the fact that in our example the difference between the means of contiguous segments are not the same, which is a direct consequence of the definition used for SNR. In practice, for any SNR, the differences between the means for segments $[t_3+1, N]$ and $[t_2+1, t_3]$

is larger than the difference between the means for segments $[t_2+1, t_3]$ and $[t_1+1, t_2]$. In Fig. 3(b), we illustrate the test points associated to the matrix $\mathbf{BB}_{\text{sup}}$. As in the previous examples, the test points tend to the true change-point values as the SNR increases. Finally, in Fig. 3(e) and (f), we illustrate the MSE performance, assuming known mean rates, as a function of the distance between change points for $\text{SNR} = -6$ dB. The bounds for change-point t_2 and t_3 are constant in all the illustrated range, though, the MSE of the MLE for t_2 slightly varies as t_1 approaches t_2. As we discussed in the previous examples, the bound for t_1 is overly optimistic for small distances between t_2 and t_1, due to the constrained test-point domain.

VI. Conclusion

We investigated a simplified version of the Barankin bound on multiple change-point estimation. The approximate Barankin information matrix was spelled, revealing an interesting tridiagonal structure, meaning that the estimation of one change point is naturally perturbed by its two neighbors. Moreover, the Barankin information matrix can be reduced to a block diagonal structure leading to closed-form for the elements of its inverse. The main limitation posed by this HCR approximation is a reduced search space for the BIM that leads to a loose Barankin bound. We also discussed the existence and computation of the supremum with respect to the Loewner partial ordering, on the finite set of candidate BB solutions. To overcome this problem, we computed a suitable minimal-upper bound to this set given by the matrix associated with the Loewner–John ellipsoid of the set of hyper-ellipsoids associated to each maximal element of the set of candidate bound matrices. Two important distributions in signal and image processing were investigated, the Gaussian case and the Poisson case, for which we obtained closed-form expressions for all the elements of the Barankin information matrix. Finally, we illustrated our analysis by presenting various simulation results. In a future work, we will analyze Barankin-type lower bounds, considering all distribution parameters in addition to the multiple change-point localizations.

Appendix

A. Proof of Lemma 1

Proof: We need to proof that for all $\mathbf{y} \in \mathbb{R}^q$ with $\mathbf{y} \neq \mathbf{0}, \mathbf{y}^T(\mathbf{A}-\mathbf{B})\mathbf{y} > 0$ if $\lambda_1 \leq 1$. Since $\mathbf{A}$ is pd and $\mathbf{B}$ is psd, there exist a non-singular matrix $\mathbf{F}$ such that

$$\mathbf{F}^T\mathbf{B}\mathbf{F} = \mathbf{diag}(\lambda_1, \ldots, \lambda_m, \lambda_{m+1}, \ldots, \lambda_q) = \Lambda \text{ and}$$
$$\mathbf{F}^T\mathbf{A}\mathbf{F} = \mathbf{I}.$$

Thus, $\mathbf{B} = (\mathbf{F}^T)^{-1}\Lambda(\mathbf{F})^{-1}$ and $\mathbf{A} = (\mathbf{F}^T)^{-1}\mathbf{I}(\mathbf{F})^{-1}$ and $\mathbf{y}^T(\mathbf{A}-\mathbf{B})\mathbf{y} = \mathbf{y}^T(\mathbf{F}^T)^{-1}(\mathbf{I}-\Lambda)(\mathbf{F})^{-1}\mathbf{y}$. Let $\mathbf{z} = (\mathbf{F})^{-1}\mathbf{y}$, because $\mathbf{F}$ is not singular $(\mathbf{F})^{-1}\mathbf{y} = \mathbf{0}$ for $\mathbf{y} = \mathbf{0}$; therefore, our problem is equivalent to analyze the positiveness of $r = \mathbf{z}^T(\mathbf{I}-\Lambda)\mathbf{z}$, for $\mathbf{z} \neq \mathbf{0}$. Since $\lambda_{m+1} = \cdots = \lambda_q = 0, r = \sum_{i=1}^{m}(1-\lambda_i)z_i^2 + \sum_{i=m+1}^{q} z_i^2$. Hence, if $\lambda_1 \leq 1$, then $(1-\lambda_i) \geq 0$, for $i = 1, \ldots, m$ thus $r > 0$ and $\mathbf{A} \succ \mathbf{B}$. On the other hand, if $\lambda_1 > 1$, we can always find a $\mathbf{z}$ vector such that $r \leq 0$ or $r > 0$, thus $\mathbf{A}$ and $\mathbf{B}$ are not mutually comparable. ■

B. Computing Diagonal Elements of $\mathbf{\Phi}$

For $\alpha_k > 0$, (12) becomes

$$\begin{aligned}[\Phi]_{kk} = \int_{\Omega} & \Pi_{i=1}^{t_1} p_1(x_i;\eta_1)\cdots\Pi_{i=t_{k-1}+1}^{t_k} p_k(x_i;\eta_k) \\ & \times \frac{\Pi_{i=t_k+1}^{t_k+\alpha_k} p_k^2(x_i;\eta_k)}{\Pi_{i=t_k+1}^{t_k+\alpha_k} p_{k+1}(x_i;\eta_{k+1})} \\ & \times \Pi_{i=t_k+\alpha_k+1}^{t_{k+1}} p_{k+1}(x_i;\eta_{k+1})\cdots \\ & \Pi_{i=t_q+1}^{N} p_{q+1}(x_i;\eta_{q+1}) dX.\end{aligned}$$

After some straightforward simplifications, we have that

$$\begin{aligned}[\Phi]_{kk} &= \int_{\Omega} \Pi_{i=t_k+1}^{t_k+\alpha_k} \frac{p_k^2(\mathbf{x}_i;\boldsymbol{\eta}_k)}{p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1})} dx_{t_k+1}\cdots dx_{t_k+\alpha_k} \\ &= \left(\int_{\Omega} \frac{p_k^2(\mathbf{x};\boldsymbol{\eta}_k)}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})} d\mathbf{x}\right)^{\alpha_k}.\end{aligned}$$

Similar analysis can be applied to solve for $\alpha_k < 0$.

C. Computing Non-Diagonal Elements of $\mathbf{\Phi}$

For $\alpha_k > 0$ and $\alpha_l > 0$, (14) becomes

$$\begin{aligned}[\mathbf{\Phi}]_{kl} = \int_{\Omega} & \frac{\Pi_{i=t_k+1}^{t_k+\alpha_k} p_k(\mathbf{x}_i;\boldsymbol{\eta}_k)}{\Pi_{i=t_k+1}^{t_k+\alpha_k} p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1})} \\ & \times \Pi_{i=1}^{t_1} p_1(\mathbf{x}_i;\boldsymbol{\eta}_1)\cdots\Pi_{i=t_k+1}^{t_k+\alpha_k} p_{k+1} \\ & \times (\mathbf{x}_i;\boldsymbol{\eta}_{k+1})\Pi_{i=t_k+\alpha_k+1}^{t_{k+1}} p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1})\cdots \\ & \Pi_{i=t_{l-1}+1}^{t_l+\alpha_l} p_l(\mathbf{x}_i;\boldsymbol{\eta}_l)\cdots\Pi_{i=t_q+1}^{N} p_{q+1} \\ & \times (\mathbf{x}_i;\boldsymbol{\eta}_{q+1}) d\mathbf{X} = \mathbf{1}.\end{aligned}$$

The cases $(\alpha_k < 0, \alpha_l < 0), (\alpha_k < 0, \alpha_l > 0)$, and $t_k + \alpha_k < t_l + \alpha_l$ are solved using same approach as above. For the overlapping case, i.e., $t_k + \alpha_k > t_l + \alpha_l$, is more difficult. Replacing $l = k+1$ and keeping in mind that $\alpha_k > 0$ and $\alpha_{k+1} < 0$, (14) becomes

$$\begin{aligned}[\mathbf{\Phi}]_{kk+1} &= \int_{\Omega} \Pi_{i=1}^{t_1} p_1(\mathbf{x}_i;\boldsymbol{\eta}_1)\Pi_{i=t_{k+1}+\alpha_{k+1}+1}^{t_k+\alpha_k} \\ &\quad \times \frac{p_k(\mathbf{x}_i;\boldsymbol{\eta}_k)p_{k+2}(\mathbf{x}_i;\boldsymbol{\eta}_{k+2})}{p_{k+1}(\mathbf{x}_i;\boldsymbol{\eta}_{k+1})}\cdots \\ &\quad \Pi_{i=t_q+1}^{N} p_{q+1}(\mathbf{x}_i;\boldsymbol{\eta}_{q+1}) d\mathbf{X} \\ &= \left(\int_{\Omega} \frac{p_k(\mathbf{x};\boldsymbol{\eta}_k)p_{k+2}(\mathbf{x};\boldsymbol{\eta}_{k+2})}{p_{k+1}(\mathbf{x};\boldsymbol{\eta}_{k+1})} d\mathbf{x}\right)^{\beta_k}\end{aligned}$$

where $\beta_k = (t_k + \alpha_k) - (t_{k+1} + \alpha_{k+1})$.

D. Computing the Elements of $\mathbf{\Phi}$ for Changes in Mean and Covariance Matrix of Gaussian Distribution

In this case $\boldsymbol{\eta}_j = [\boldsymbol{\nu}_j^T, \boldsymbol{\varphi}_j^T]^T$, and the data likelihood is given as follows:

$$p(\mathbf{X};\mathbf{t}) = \frac{1}{(2\pi)^{NM/2}\Pi_{j=1}^{q+1}|\mathbf{M}(\boldsymbol{\varphi}_j)|^{(t_j - t_{j-1})/2}} \times \exp\left\{-\frac{1}{2}\mathrm{Tr}\left\{\sum_{j=1}^{q+1}\mathbf{M}(\boldsymbol{\varphi}_j)^{-1} \times \left(\sum_{i=t_{j-1}+1}^{t_j}(\mathbf{x}_i - \mathbf{f}(\boldsymbol{\nu}_j))(\mathbf{x}_i - \mathbf{f}(\boldsymbol{\nu}_j))^T\right)\right\}\right\}.$$

For $\alpha_k > 0$, using (13), we have that $[\mathbf{\Phi}]_{kk}$ is given as follows:

$$[\mathbf{\Phi}]_{kk} = \left(\frac{|\mathbf{M}(\boldsymbol{\varphi}_{k+1})|^{1/2}}{(2\pi)^{M/2}|\mathbf{M}(\boldsymbol{\varphi}_k)|}\right)^{\alpha_k} \times \left(\int_{\mathbb{R}^M}\exp\left\{-\frac{1}{2}\left(\mathbf{x}_i^T\mathbf{M}_k\mathbf{x}_i - 2\mathbf{g}_k^T\mathbf{x}_i\right)\right\}d\mathbf{x}_i\right)^{\alpha_k} \times \exp\left\{-\frac{\alpha_k}{2}\mathbf{f}^T(\boldsymbol{\nu}_k)2(\mathbf{M}(\boldsymbol{\varphi}_k))^{-1}\mathbf{f}(\boldsymbol{\nu}_k) + \frac{\alpha_k}{2}\mathbf{f}^T(\boldsymbol{\nu}_{k+1})(\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+1})\right\}$$

where $M_k = (2(\mathbf{M}(\boldsymbol{\varphi}_k))^{-1} - (\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1})$ and $g_k = 2(\mathbf{M}(\boldsymbol{\varphi}_k))^{-1}f(\nu_k) - (\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1}f(\nu_{k+1})$. The integral above has a finite value for $\mathbf{M}_k$ positive definite (pd). Hence, and after some straightforward algebraic derivations, we obtain the expression in (33). The case $\alpha_k < 0$ is obtained proceeding similarly as above. For $t_k + \alpha_k > t_{k+1} + \alpha_{k+1}$, using (17), we have that $[\mathbf{\Phi}]_{kk+1}$ is given as follows:

$$[\mathbf{\Phi}]_{kk+1} = \left(\frac{|\mathbf{M}(\boldsymbol{\varphi}_{k+1})|^{1/2}}{(2\pi)^{M/2}|\mathbf{M}(\boldsymbol{\varphi}_k)|^{1/2}|\mathbf{M}(\boldsymbol{\varphi}_{k+2})|^{1/2}}\right)^{\beta_k} \times \left(\int_{\mathbb{R}^M}\exp\left\{-\frac{1}{2}\left(\mathbf{x}_i^T\bar{\mathbf{M}}_k\mathbf{x}_i - 2\bar{\mathbf{g}}_k^T\mathbf{x}_i\right)\right\}d\mathbf{x}\right)^{\beta_k} \times \exp\left\{-\frac{\beta_k}{2}\mathbf{f}^T(\boldsymbol{\nu}_k)(\mathbf{M}(\boldsymbol{\varphi}_k))^{-1}\mathbf{f}(\boldsymbol{\nu}_k) - \frac{\beta_k}{2}\mathbf{f}^T(\nu_{k+2})(\mathbf{M}(\boldsymbol{\varphi}_{k+2}))^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+2}) + \frac{\beta_k}{2}\mathbf{f}^T(\boldsymbol{\nu}_{k+1})(\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+1})\right\}$$

where $\bar{\mathbf{M}}_k = (\mathbf{M}(\boldsymbol{\varphi}_k))^{-1} + (\mathbf{M}(\boldsymbol{\varphi}_{k+2}))^{-1} - (\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1}$, and $\{\bar{\mathbf{g}}_k = (\mathbf{M}(\boldsymbol{\varphi}_k))^{-1}\mathbf{f}(\boldsymbol{\nu}_k) + (\mathbf{M}(\boldsymbol{\varphi}_{k+2}))^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+2}) - (\mathbf{M}(\boldsymbol{\varphi}_{k+1}))^{-1}\mathbf{f}(\boldsymbol{\nu}_{k+1})$. Hence, and after some straightforward algebraic derivations, we obtain the expression in (35).

REFERENCES

[1] A. Benveniste and M. Basseville, *Detection of Abrupt Changes in Signals and Dynamical Systems*. Berlin, Germany: Springer-Verlag, 1986.

[2] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes, Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, Apr. 1993.

[3] F. Gustafsson, *Adaptive Filtering and Change Detection*, 1st ed. New York: Wiley, Oct. 2000.

[4] B. Brodskya and B. Darkhovskyb, *Non-Parametric Statistical Diagnosis. Problems and Methods*. Boston, MA: Kluwer, 2000.

[5] B. Brodskya and B. Darkhovskyb, "Asymptotically optimal methods of change-point detection for composite hypotheses," *J. Stat. Planning Inference*, vol. 133, pp. 123–138, 2005.

[6] H. Cramér, *Mathematical Methods of Statistics*. New York: Princeton Univ. Press, Sep. 1946, vol. 9.

[7] R. Kakarala and A. O. Hero, "On achievable accuracy in edge localization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 7, pp. 777–781, Jul. 1992.

[8] A. Bartov and H. Messer, "Analysis of inherent limitations in localizing step-like singularities in a continuous signal," in *Proc. IEEE SP Int. Symp. Time-Frequency Time-Scale Analysis*, Paris, France, Jun. 1996, pp. 21–24.

[9] A. M. Reza and M. Doroodchi, "Cramér–Rao lower bound on locations of sudden changes in a steplike signal," *IEEE Trans. Signal Process.*, vol. 44, no. 10, pp. 2551–2556, Oct. 1996.

[10] J.-Y. Tourneret, M. Chabert, and M. Ghogho, "Detection and estimation of multiplicative jumps," in *Proc. 8th IEEE Signal Processing Workshop Statistical Signal Array Processing*, Jun. 1996, pp. 20–23.

[11] A. Bartov and H. Messer, "Lower bound on the achievable DSP performance for localizing step-like continuous signals in noise," *IEEE Trans. Signal Process.*, vol. 46, no. 8, pp. 2195–2201, Aug. 1998.

[12] A. Swami and B. Sadler, "Analysis of multiscale products for step detection and estimation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 1043–1051, Apr. 1999.

[13] S. Zacks, *The Theory of Statistical Inference*. New York: Wiley, 1971.

[14] A. Swami and B. Sadler, "Cramér–Rao bounds for step-change localization in additive and multiplicative noise," in *Proc. IEEE Signal Processing Workshop Statistical Signal Array Processing*, Sep. 1999, pp. 403–406.

[15] J. V. Braun and H. G. Muller, "Statistical methods for DNA sequence segmentation," *Stat. Sci.*, vol. 13, no. 2, pp. 142–162, 1998.

[16] P. S. La Rosa, A. Nehorai, H. Eswaran, C. Lowery, and H. Preissl, "Detection of uterine MMG contractions using a multiple change point estimator and the K-means cluster algorithm," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 2, pp. 453–467, Feb. 2008.

[17] C. H. Wu and C. H. Hsieh, "Multiple change-point audio segmentation and classification using an MDL-based Gaussian model," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 2, pp. 647–657, Mar. 2006.

[18] N. Dobigeon, J.-Y. Tourneret, and J. D. Scargle, "Change-point detection in astronomical data by using a hierarchical model and a Bayesian sampling approach," in *Proc. IEEE/SP 13th Workshop Statistical Signal Processing*, Jul. 2005, pp. 369–374.

[19] E. W. Barankin, "Locally best unbiased estimates," *Ann. Math. Stat.*, vol. 20, no. 4, pp. 477–501, Dec. 1949.

[20] D. G. Chapman and H. Robbins, "Minimum variance estimation without regularity assumptions," *Ann. Math. Stat.*, vol. 22, no. 4, pp. 581–586, Dec. 1951.

[21] L. Knockaert, "The Barankin bound and threshold behavior in frequency estimation," *IEEE Trans. Signal Process.*, vol. 45, no. 9, pp. 2398–2401, Sep. 1997.

[22] A. Ferrari and J.-Y. Tourneret, "Barankin lower bound for change-points in independent sequences," in *Proc. IEEE Workshop Statistical Signal Processing (SSP)*, St. Louis, MO, USA, Sep. 2003, pp. 557–560.

[23] J. M. Hammersley, "On estimating restricted parameters," *J. Roy. Stat. Soc. Series B (Methodol.)*, vol. 12, no. 2, pp. 192–240, 1950.

[24] B. Burgeth, A. Bruhna, N. Papenberg, M. Welka, and J. Weickert, "Mathematical morphology for matrix fields induced by the Loewner ordering in higher dimensions," *Signal Process.*, vol. 87, no. 2, pp. 277–290, Feb. 2007.

[25] F. Pukelsheim, *Optimal Design of Experiments*, 1st ed. New York: Wiley, 1993.

[26] J. Borwein and A. Lewis, *Convex Analysis and Nonlinear Optimization*, 1st ed. New York: Springer-Verlag, 2000.

[27] J. D. Gorman and A. O. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Trans. Inf. Theory*, vol. 26, no. 6, Nov. 1990.

[28] I. Reuven and H. Messer, "A Barankin-type lower bound on the estimation error of a hybrid parameter vector," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 1084–1093, May 1997.

[29] R. J. McAulay and L. P. Seidman, "A useful form of the Barankin lower bound and its application to PPM threshold analysis," *IEEE Trans. Inf. Theory*, vol. 15, pp. 273–279, Mar. 1969.

[30] R. J. McAulay and E. M. Hofstetter, "Barankin bounds on parameter estimation," *IEEE Trans. Inf. Theory*, vol. 17, pp. 669–676, Nov. 1971.

[31] J. S. Abel, "A bound on mean square estimate error," *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 1675–1680, Sep. 1993.

[32] J. Tabrikian and J. Krolik, "Barankin bounds for source localization in an uncertain ocean environment," *IEEE Trans. Signal Process.*, vol. 47, no. 11, pp. 2917–2927, Nov. 1999.

[33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[34] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming. Dec. 2008 [Online]. Available: http://stanford.edu/boyd/cvx, Web page and software

[35] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," *Recent Adv. Learn. Control (a Tribute to M. Vidyasagar)*, vol. 371, pp. 95–110, Dec. 2008.

[36] G. Strang, *Introduction to Linear Algebra*. Wellesley, MA: Wellesley-Cambridge Press, 2005.

[37] T. Muir, *A Treatise on the Theory of Determinants*. New York: Dover, 1960.

[38] P. Swerling, "Probability of detection for fluctuating targets," *IEEE Trans. Inf. Theory*, vol. 6, no. 2, pp. 269–308, Apr. 1960.

[39] J. F. Böhme, "Estimation of source parameters by maximum likelihood and non linear regression," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, 1984, pp. 731–734.

[40] J. F. Böhme, "Estimation of spectral parameters of correlated signals in wavefields," *Signal Process.*, vol. 10, pp. 329–337, 1986.

[41] Z. Liu, A. Nehorai, and E. Paldi, "Statistical analysis of a generalized compound eye detector array," in *Proc. 4th IEEE Sensor Conf.*, Oct. –Nov. 2005, pp. 568–571.

[42] S. M. Kay, *Fundamentals of Statistical Signal Processing Vol. 2: Detection Theory*, 1st ed. Englewood Cliffs, NJ: Prentice-Hall PTR, Jan. 1998.

[43] D. V. Hinkley, "Inference about the change-point in a sequence of random variables," *Biometrika*, vol. 57, no. 1, pp. 1–17, 1970.

[44] Y.-C. Yao and S. T. Au, "Least-squares estimation of a step function," *Sankhya: The Indian J. Stat., Series A*, vol. 51, no. 3, pp. 370–381, Oct. 1989.

Patricio S. La Rosa (S'07–M'09) received the B.Sc. degree in engineering from the Pontifical Catholic University of Chile (PUC) in 1999, the M.Sc. degree (with maximum distinction) in electrical engineering from the University of Chile in 2003, and the M.Sc. and the Ph.D. degrees in electrical engineering from Washington University in St. Louis (WUSTL) in 2010, under the guidance of Prof. A. Nehorai.

He was Research Assistant in the Signal and Image Research Laboratory at the University of Illinois at Chicago, from 2003 to 2005, and in the Center for Sensor Signal and Information Processing at WUSTL, from 2006 to 2010. Since August 2010, he has been a Postdoctoral Research Associate at the General Medical Sciences Division, Department of Internal Medicine, Washington University School of Medicine. His research interests are in statistical array signal processing and its applications to medical imaging and biosignal processing, and in biophysical modeling, including nonlinear waveform phenomena in excitable media, bioelectromagnetism, and optical mapping.

Dr. La Rosa received the John Paul II Foundation scholarship between the years 1995 and 2000 for undergraduate studies in engineering sciences at PUC.

Alexandre Renaux (S'06–M'07) was born in Vosges, France in 1978. He was born close to Christmas 1978. He received the French Agregation degree and also the M.Sc. and Ph.D. degrees in electrical engineering from the Ecole Normale Suprieure de Cachan, France, in 2002, 2003 and 2006, respectively.

From 2006 to 2007, he was a Postdoctoral Research Associate in the Department of Electrical and Systems Engineering at Washington University in St. Louis, MO. Since 2007, he has been an Assistant Professor in the Physics Department and member of the Laboratory of Signals and Systems (L2S) at the Universit Paris-Sud 11, Orsay, France. His research interests are in the not very well known, but interesting, field of lower bounds on the mean-square error applied to statistical signal processing.

Carlos H. Muravchik (S'81–M'83–SM'99) was born in Argentina on June 11, 1951. He graduated as an Electronics Engineer from the National University of La Plata, Argentina, in 1973, and he received the M.Sc. degree in electrical engineering, the M.Sc. degree in statistics, and the Ph.D. degree in electrical engineering, all from Stanford University, Stanford, CA, in 1980, 1983, and 1983, respectively.

He is a Professor at the Department of the Electrical Engineering of the National University of La Plata and a member of its Industrial Electronics, Control and Instrumentation Laboratory (LEICI). He is also a member of the Comision de Investigaciones Cientificas de la Pcia. de Buenos Aires. He was a Visiting Professor to Yale University in 1983 and 1994; to the University of Illinois at Chicago in 1996, 1997, 1999, and 2003; and to Washington University in St. Louis in 2006 and 2010. His research interests are in the area of statistical and array signal processing with biomedical, communications and control applications, and in nonlinear control systems.

Dr. Muravchik has been a member of the Advisory Board of the journal *Latin American Applied Research* since 1999 and was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2003 to 2006.

Arye Nehorai (S'80–M'83–SM'90–F'94) received the B.Sc. and M.Sc. degrees from the Technion—Israel Institute of Technology, Haifa, and the Ph.D. from Stanford University, Stanford, CA.

He is currently the Eugene and Martha Lohman Professor and Chair of the Department of Electrical and Systems Engineering at Washington University in St. Louis (WUSTL). He serves as the Director of the Center for Sensor Signal and Information Processing at WUSTL. Previously, he was a faculty member at Yale University and the University of Illinois at Chicago.

Dr. Nehorai has served as Editor-in-Chief of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2000 to 2002. From 2003 to 2005, he was Vice-President (Publications) of the IEEE Signal Processing Society (SPS), Chair of the Publications Board, and member of the Executive Committee of this Society. He was the Founding Editor of the special columns on Leadership Reflections in the *IEEE Signal Processing Magazine* from 2003 to 2006. He received the 2006 IEEE SPS Technical Achievement Award and the 2010 IEEE SPS Meritorious Service Award. He was elected Distinguished Lecturer of the IEEE SPS for the term 2004 to 2005. He was corecipient of the IEEE SPS 1989 Senior Award for Best Paper coauthor of the 2003 Young Author Best Paper Award and corecipient of the 2004 Magazine Paper Award. In 2001, he was named University Scholar of the University of Illinois. He is the Principal Investigator of the Multidisciplinary University Research Initiative (MURI) project entitled Adaptive Waveform Diversity for Full Spectral Dominance. He has been a Fellow of the Royal Statistical Society since 1996.

Annexe F

Statistical resolution limit for the multidimensional harmonic retrieval model : Hypothesis test and Cramér-Rao bound approaches

A paraître dans EURASIP Journal on Advances in Signal Processing, special issue on Advances in Angle-of-Arrival and Multidimensional Signal Processing for Localization and Communications

El Korso *et al. EURASIP Journal on Advances in Signal Processing* 2011, **2011**:12
http://asp.eurasipjournals.com/content/2011/1/12

EURASIP Journal on Advances in Signal Processing
a SpringerOpen Journal

RESEARCH Open Access

Statistical resolution limit for the multidimensional harmonic retrieval model: hypothesis test and Cramér-Rao Bound approaches

Mohammed Nabil El Korso*, Rémy Boyer, Alexandre Renaux and Sylvie Marcos

Abstract

The statistical resolution limit (SRL), which is defined as the minimal separation between parameters to allow a correct resolvability, is an important statistical tool to quantify the ultimate performance for parametric estimation problems. In this article, we generalize the concept of the SRL to the multidimensional SRL (MSRL) applied to the multidimensional harmonic retrieval model. In this article, we derive the SRL for the so-called multidimensional harmonic retrieval model using a generalization of the previously introduced SRL concepts that we call multidimensional SRL (MSRL). We first derive the MSRL using an hypothesis test approach. This statistical test is shown to be asymptotically an uniformly most powerful test which is the *strongest* optimality statement that one could expect to obtain. Second, we link the proposed asymptotic MSRL based on the hypothesis test approach to a new extension of the SRL based on the Cramér-Rao Bound approach. Thus, a closed-form expression of the asymptotic MSRL is given and analyzed in the framework of the multidimensional harmonic retrieval model. Particularly, it is proved that the optimal MSRL is obtained for equi-powered sources and/or an equi-distributed number of sensors on each multi-way array.

Keywords: Statistical resolution limit, Multidimensional harmonic retrieval, Performance analysis, Hypothesis test, Cramér-Rao bound, Parameter estimation, Multidimensional signal processing

Introduction

The multidimensional harmonic retrieval problem is an important topic which arises in several applications [1]. The main reason is that the multidimensional harmonic retrieval model is able to handle a large class of applications. For instance, the joint angle and carrier estimation in surveillance radar system [2,3], the underwater acoustic multisource azimuth and elevation direction finding [4], the 3-D harmonic retrieval problem for wireless channel sounding [5,6] or the detection and localization of multiple targets in a MIMO radar system [7,8].

One can find many estimation schemes adapted to the multidimensional harmonic retrieval estimation problem, see, e.g., [1,2,4-7,9,10]. However, to the best of our knowledge, no work has been done on the resolvability of such a multidimensional model.

The resolvability of closely spaced signals, in terms of parameter of interest, for a given scenario (e.g., for a given signal-to-noise ratio (SNR), for a given number of snapshots and/or for a given number of sensors) is a former and challenging problem which was recently updated by Smith [11], Shahram and Milanfar [12], Liu and Nehorai [13], and Amar and Weiss [14]. More precisely, the concept of statistical resolution limit (SRL), i. e., the minimum distance between two closely spaced signals[a] embedded in an additive noise that allows a correct resolvability/parameter estimation, is rising in several applications (especially in problems such as radar, sonar, and spectral analysis [15].)

The concept of the SRL was defined/used in several manners [11-14,16-24], which could turn in it to a confusing concept. There exist essentially three approaches

* Correspondence: elkorso@lss.supelec.fr
Laboratoire des Signaux et Systèmes (L2S), Université Paris-Sud XI (UPS), CNRS, SUPELEC, 3 Rue Joliot Curie, Gif-Sur-Yvette 91192, France

to define/obtain the SRL. (i) The first is based on the concept of mean null spectrum: assuming, e.g., that two signals are parameterized by the frequencies f_1 and f_2, the Cox criterion [16] states that these sources are resolved, w.r.t. a given high-resolution estimation algorithm, *if the mean null spectrum at each frequency f_1 and f_2 is lower than the mean of the null spectrum at the midpoint* $\frac{f_1+f_2}{2}$. Another commonly used criterion, also based on the concept of the mean null spectrum, is the Sharman and Durrani criterion [17], which states that two sources are resolved *if the second derivative of the mean of the null spectrum at the midpoint* $\frac{f_1+f_2}{2}$ *is negative*. It is clear that the SRL based on the mean null spectrum is relevant to a specific high-resolution algorithm (for some applications of these criteria one can see [16-19] and references therein.) (ii) The second approach is based on detection theory: the main idea is to use a hypothesis test to decide if one or two closely spaced signals are present in the set of the observations. Then, the challenge herein is to link the minimum separation, between two sources (e.g., in terms of frequencies) that is detectable at a given SNR, to the probability of false alarm, P_{fa} and/or to the probability of detection P_{d}. In this spirit, Sharman and Milanfar [12] have considered the problem of distinguishing whether the observed signal contains one or two frequencies at a given SNR using the generalized likelihood ratio test (GLRT). The authors have derived the SRL expressions w.r.t. P_{fa} and P_{d} in the case of real received signals, and unequal and unknown amplitudes and phases. In [13], Liu and Nehorai have defined a statistical angular resolution limit using the asymptotic equivalence (in terms of number of observations) of the GLRT. The challenge was to determine the minimum angular separation, in the case of complex received signals, which allows to resolve two sources knowing the direction of arrivals (DOAs) of one of them for a given P_{fa} and a given P_{d}. Recently, Amar and Weiss [14] have proposed to determine the SRL of complex sinusoids with nearby frequencies using the Bayesian approach for a given correct decision probability. (iii) The third approach is based on a estimation accuracy criteria independent of the estimation algorithm. Since the Cramér-Rao Bound (CRB) expresses a lower bound on the covariance matrix of any unbiased estimator, then it expresses also the ultimate estimation accuracy [25,26]. Consequently, it could be used to describe/obtain the SRL. In this context, one distinguishes two main criteria for the SRL based on the CRB: (1) the first one was introduced by Lee [20] and states that: *two signals are said to be resolvable w.r.t. the frequencies if the maximum standard deviation is less than twice the difference between f_1 and f_2*. Assuming that the CRB is a tight bound (under mild/weak conditions), the standard deviation, $\sigma_{\hat{f}_1}$ and $\sigma_{\hat{f}_2}$, of an unbiased estimator $\hat{\boldsymbol{f}} = [\hat{f}_1\ \hat{f}_2]^{\text{T}}$ is given by $\sqrt{\text{CRB}(f_1)}$ and $\sqrt{\text{CRB}(f_2)}$, respectively. Consequently, the SRL is defined, in the Lee criterion sense, as $2\max\left\{\sqrt{\text{CRB}(f_1)}, \sqrt{\text{CRB}(f_2)}\right\}$. One can find some results and applications in [20,21] where this criterion is used to derive a matrix-based expression (i.e., without analytic inversion of the Fisher information matrix) of the SRL for the frequency estimates in the case of the conditional and unconditional signal source models. On the other hand, Dilaveroglu [22] has derived a closed-form expression of the frequency resolution for the real and complex conditional signal source models. However, one can note that the coupling between the parameters, $\text{CRB}(f_1, f_2)$ (i.e., the CRB for the cross parameters f_1 and f_2), is ignored by this latter criterion. (2) To extend this, Smith [11] has proposed the following criterion: *two signals are resolvable w.r.t. the frequencies if the difference between the frequencies, δ_f, is greater than the standard deviation of the DOA difference estimation*. Since, the standard deviation can be approximated by the CRB, then, the SRL, in the Smith criterion sense, is defined as the limit of δ_f for which $\delta_f < \sqrt{\text{CRB}(\delta_f)}$ is achieved. This means that, the SRL is obtained by solving the following implicit equation

$$\delta_f^2 = \text{CRB}(\delta_f) = \text{CRB}(f_1) + \text{CRB}(f_2) - 2\text{CRB}(f_1, f_2).$$

In [11,23], Smith has derived the SRL for two closely spaced sources in terms of DOA, each one modeled by one complex pole. In [24], Delmas and Abeida have derived the SRL based on the Smith criterion for DOA of discrete sources under QPSK, BPSK, and MSK model assumptions. More recently, Kusuma and Goyal [27] have derived the SRL based on the Smith criterion in sampling estimation problems involving a powersum series.

It is important to note that all the criteria listed before take into account only one parameter of interest per signal. Consequently, all the criteria listed before cannot be applied to the aforementioned the multidimensional harmonic model. To the best of our knowledge, no results are available on the SRL for multiple parameters of interest per signal. The goal of this article is to fill this lack by proposing and deriving the so-called MSRL for the multidimensional harmonic retrieval model.

More precisely, in this article, the MSRL for multiple parameters of interest per signal using a hypothesis test is derived. This choice is motivated by the following arguments: (i) the hypothesis test approach is not specific to a certain high-resolution algorithm (unlike the mean null spectrum approach), (ii) in this article, we

link the asymptotic MSRL based on the hypothesis test approach to a new extension of the MSRL based on the CRB approach. Furthermore, we show that the MSRL based on the CRB approach is equivalent to the MSRL based on the hypothesis test approach for a fixed couple (P_{fa}, P_{d}), and (iii) the hypothesis test is shown to be asymptotically an uniformly most powerful test which is the *strongest* statement of optimality that one could expect to obtain [28].

The article is organized as follows. We first begin by introducing the multidimensional harmonic model, in section "Model setup". Then, based on this model, we obtain the MSRL based on the hypothesis test and on the CRB approach. The link between theses two MSRLs is also described in section "Determination of the MSRL for two sources" followed by the derivation of the MSRL closed-form expression, where, as a by product the exact closed-form expressions of the CRB for the multidimensional retrieval model is derived (note that to the best of our knowledge, no exact closed-form expressions of the CRB for such model is available in the literature). Furthermore, theoretical and numerical analyses are given in the same section. Finally, conclusions are given.

Glossary of notation

The following notations are used through the article. Column vectors, matrices, and multi-way arrays are represented by lower-case bold letters ($\boldsymbol{a}$, ...), upper-case bold letters ($\boldsymbol{A}$, ...) and bold calligraphic letters ($\boldsymbol{\mathcal{A}}$, ...), whereas

- $\mathbb{R}$ and $\mathbb{C}$ denote the body of real and complex values, respectively,
- $\mathbb{R}^{D_1\times D_2\times\cdots\times D_I}$ and $\mathbb{C}^{D_1\times D_2\times\cdots\times D_I}$ denote the real and complex multi-way arrays (also called tensors) body of dimension $D_1 \times D_2 \times ... \times D_I$, respectively,
- j = the complex number $\sqrt{-1}$.
- $\mathbf{I}_{\mathbf{Q}}$ = the identity matrix of dimension Q,
- $\mathbf{0}_{\mathbf{Q}_1\times\mathbf{Q}_2}$ = the $Q_1 \times Q_2$ matrix filled by zeros,
- $[\boldsymbol{a}]_i$ = the ith element of the vector $\boldsymbol{a}$,
- $[\boldsymbol{A}]_{i_1,i_2}$ = the i_1th row and the i_2th column element of the matrix $\boldsymbol{A}$,
- $[\boldsymbol{\mathcal{A}}]_{i_1,i_2,...,i_N}$ = the $(i_1, i_2, ..., i_N)$th entry of the multi-way array $\boldsymbol{\mathcal{A}}$,
- $[\boldsymbol{A}]_{i,p:q}$ = the row vector containing the $(q - p + 1)$ elements $[\boldsymbol{A}]_{i,k}$, where $k = p, ..., q$,
- $[\boldsymbol{A}]_{p:q,k}$ = the column vector containing the $(q - p + 1)$ elements $[\boldsymbol{A}]_{i,k}$, where $i = p, ..., q$,
- the derivative of vector $\boldsymbol{a}$ *w.r.t.* to vector $\boldsymbol{b}$ is defined as follows: $\left[\frac{\partial \boldsymbol{a}}{\partial \boldsymbol{b}}\right]_{i,j} = \frac{\partial [\boldsymbol{a}]_i}{\partial [\boldsymbol{b}]_j}$,
- $\boldsymbol{A}^{\mathrm{T}}$ = the transpose of the matrix $\boldsymbol{A}$,
- $\boldsymbol{A}^*$ = the complex conjugate of the matrix $\boldsymbol{A}$,
- $\boldsymbol{A}^{\mathrm{H}} = (\boldsymbol{A}^*)^{\mathrm{T}}$,
- tr $\{\boldsymbol{A}\}$ = the trace of the matrix $\boldsymbol{A}$,
- det $\{\boldsymbol{A}\}$ = the determinant of the matrix $\boldsymbol{A}$,
- $\Re\{a\}$ = the real part of the complex number a,
- $\mathbb{E}\{a\}$ = the expectation of the random variable a,
- $||\boldsymbol{a}||^2 = \frac{1}{L}\sum_{t=1}^{L} [\boldsymbol{a}]_t^2$ denotes the normalized norm of the vector $\boldsymbol{a}$ (in which L is the size of $\boldsymbol{a}$),
- sgn (a) = 1 if $a \geq 0$ and -1 otherwise.
- diag$(\boldsymbol{a})$ is the diagonal operator which forms a diagonal matrix containing the vector $\boldsymbol{a}$ on its diagonal,
- vec(.) is the vec-operator stacking the columns of a matrix on top of each other,
- $\odot$ stands for the Hadamard product,
- $\otimes$ stands for the Kronecker product,
- $\circ$ denotes the multi-way array outer-product (recall that for a given multi-way arrays $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{A_1\times A_2\times\cdots\times A_I}$ and $\boldsymbol{\mathcal{B}} \in \mathbb{C}^{B_1\times B_2\times\cdots\times B_J}$, the result of the outer-product of $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$ denoted by $\boldsymbol{\mathcal{C}}^{A_1\times\cdots\times A_I\times B_1\times\cdots\times B_J}$ is given by $[\boldsymbol{\mathcal{C}}]_{a_1,...,a_I,b_1,...,b_J} = [\boldsymbol{\mathcal{A}}\circ\boldsymbol{\mathcal{B}}]_{a_1,...,a_I,b_1,...,b_J} = [\boldsymbol{\mathcal{A}}]_{a_1,...,a_I}[\boldsymbol{\mathcal{B}}]_{b_1,...,b_J}$).

Model setup

In this section, we introduce the multidimensional harmonic retrieval model in the multi-way array form (also known as *tensor form* [29]). Then, we use the PARAFAC (PARallel FACtor) decomposition to obtain a vector form of the observation model. This vector form will be used to derive the closed-form expression of the MSRL.

Let us consider a multidimensional harmonic model consisting of the superposition of two harmonics each one of dimension P contaminated by an additive noise. Thus, the observation model is given as follows [8,9,26,30-32]:

$$[\boldsymbol{\mathcal{Y}}(t)]_{n_1,...,n_P} = [\boldsymbol{\mathcal{X}}(t)]_{n_1,...,n_P} + [\boldsymbol{\mathcal{N}}(t)]_{n_1,...,n_P}, \quad t = 1, \ldots, L, \quad \text{and} \quad n_p = 0, \ldots, N_p - 1, \tag{1}$$

where $\boldsymbol{\mathcal{Y}}(t)$, $\boldsymbol{\mathcal{X}}(t)$, and $\boldsymbol{\mathcal{N}}(t)$ denote the noisy observation, the noiseless observation, and the noise multi-way array at the tth snapshot, respectively. The number of snapshots and the number of sensors on each array are denoted by L and $(N_1, ...,N_P)$, respectively. The noiseless observation multi-way array can be written as follows[b] [26,30-32]:

$$[\boldsymbol{\mathcal{X}}(t)]_{n_1,...,n_P} = \sum_{m=1}^{2} s_m(t) \prod_{p=1}^{P} e^{j\omega_m^{(p)} n_p}, \tag{2}$$

where $\omega_m^{(p)}$and $s_m(t)$ denote the mth frequency viewed along the pth dimension or array and the mth complex signal source, respectively. Furthermore, the signal source is given by $s_m(t) = \alpha_m(t)e^{j\phi_m(t)}$ where $\alpha_m(t)$ and

$\varphi_m(t)$ denote the real positive amplitude and the phase for the mth signal source at the tth snapshot, respectively.

Since,

$$\prod_{p=1}^{P} e^{j\omega_m^{(p)} n_p} = \left[\boldsymbol{a}(\omega_m^{(1)}) \circ \boldsymbol{a}(\omega_m^{(2)}) \circ \cdots \circ \boldsymbol{a}(\omega_m^{(P)})\right]_{n_1,n_2,\ldots,n_P},$$

where $\boldsymbol{a}(.)$ is a Vandermonde vector defined as

$$\boldsymbol{a}(\omega_m^{(p)}) = \left[1 \quad e^{j\omega_m^{(p)}} \quad \cdots \quad e^{j}(N_p - 1)\omega_m^{(p)}\right]^{\mathrm{T}},$$

then, the multi-way array $\boldsymbol{\mathcal{X}}(t)$ follows a PARAFAC decomposition [7,33]. Consequently, the noiseless observation multi-way array can be rewritten as follows:

$$\boldsymbol{\mathcal{X}}(t) = \sum_{m=1}^{2} s_m(t)\left(\boldsymbol{a}(\omega_m^{(1)}) \circ \boldsymbol{a}(\omega_m^{(2)}) \circ \cdots \circ \boldsymbol{a}(\omega_m^{(P)})\right). \quad (3)$$

First, let us vectorize the noiseless observation as follows:

$$\mathrm{vec}(\boldsymbol{\mathcal{X}}(t)) = [[\boldsymbol{\mathcal{X}}(t)]_{0,0,\ldots,0} \cdots [\boldsymbol{\mathcal{X}}(t)]_{N_1-1,0,\cdots,0} [\boldsymbol{\mathcal{X}}(t)]_{0,1,\ldots,0} \cdots [\boldsymbol{\mathcal{X}}(t)]_{N_1-1,N_2-1,\ldots,N_P-1}]^{\mathrm{T}}. \quad (4)$$

Thus, the full noise-free observation vector is given by

$$\mathbf{x} = \left[\mathrm{vec}^{\mathrm{T}}(\boldsymbol{\mathcal{X}}(1)) \quad \mathrm{vec}^{\mathrm{T}}(\boldsymbol{\mathcal{X}}(2)) \cdots \mathrm{vec}^{\mathrm{T}}(\boldsymbol{\mathcal{X}}(L))\right]^{\mathrm{T}}.$$

Second, and in the same way, we define $\mathbf{y}$, the noisy observation vector, and $\mathbf{n}$, the noise vector, by the concatenation of the proper multi-way array's entries, i.e.,

$$\mathbf{y} = \left[\mathrm{vec}^{\mathrm{T}}(\boldsymbol{\mathcal{Y}}(1)) \quad \mathrm{vec}^{\mathrm{T}}(\boldsymbol{\mathcal{Y}}(2)) \cdots \mathrm{vec}^{\mathrm{T}}(\boldsymbol{\mathcal{Y}}(L))\right]^{\mathrm{T}} = \mathbf{x} + \mathbf{n}. \quad (5)$$

Consequently, in the following, we will consider the observation model in (5). Furthermore, the unknown parameter vector is given by

$$\boldsymbol{\xi} = \left[\boldsymbol{\omega}^{\mathrm{T}} \boldsymbol{\rho}^{\mathrm{T}}\right]^{\mathrm{T}}, \quad (6)$$

where $\boldsymbol{\omega}$ denotes the unknown parameter vector of interest, i.e., containing all the unknown frequencies

$$\boldsymbol{\omega} = \left[\left(\boldsymbol{\omega}^{(1)}\right)^{\mathrm{T}} \cdots \left(\boldsymbol{\omega}^{(P)}\right)^{\mathrm{T}}\right]^{\mathrm{T}},$$

in which

$$\boldsymbol{\omega}^{(p)} = \left[\omega_1^{(p)} \ \omega_2^{(p)}\right]^{\mathrm{T}}. \quad (7)$$

whereas $\boldsymbol{\rho}$ contains the unknown nuisance/unwanted parameters vector, i.e., characterizing the noise covariance matrix and/or amplitude and phase of each source (e.g., in the case of a covariance noise matrix equal to $\sigma^2 \mathbf{I}_{LN_1 \ldots N_P}$ and unknown deterministic amplitudes and phases, the unknown nuisance/unwanted parameters vector $\boldsymbol{\rho}$ is given by $\boldsymbol{\rho} = [\alpha_1(1) \ldots \alpha_2(L)\varphi_1(1) \ldots \varphi_2(L)\sigma^2]^{\mathrm{T}}$.

In the following, we conduct a hypothesis test formulation on the observation model (5) to derive our MSRL expression in the case of two sources.

Determination of the MSRL for two sources

Hypothesis test formulation

Resolving two closely spaced sources, with respect to their parameters of interest, can be formulated as a binary hypothesis test [12-14] (for the special case of P = 1). To determine the MSRL (i.e., $P \geq 1$), let us consider the hypothesis $\mathcal{H}_0$ which represents the case where the two emitted signal sources are combined into one signal, i.e., the two sources have the same parameters (this hypothesis is described by $\forall p \in [1 \ldots P], \omega_1^{(p)} = \omega_2^{(p)}$), whereas the hypothesis $\mathcal{H}_1$ embodies the situation where the two signals are resolvable (the latter hypothesis is described by $\exists p \in [1 \ldots P]$, such that $\omega_1^{(p)} \neq \omega_2^{(p)}$). Consequently, one can formulate the hypothesis test, as a simple one-sided binary hypothesis test as follows:

$$\begin{cases} \mathcal{H}_0 : \delta = 0, \\ \mathcal{H}_1 : \delta > 0, \end{cases} \quad (8)$$

where the parameter δ is the so-called MSRL which indicates us in which hypothesis our observation model belongs. Thus, the question addressed below is how can we define the MSRL δ such that all the P parameters of interest are taken into account? A natural idea is that δ reflects a distance between the P parameters of interest. Let the MSRL denotes the l_1 norm[c] between two sets containing the parameters of interest of each source (which is the naturally used norm, since in the mono-parameter frequency case that we extend here, the SRL is defined as $\delta = f_1 - f_2$ [13,14,34]). Meaning that, if we denote these sets as C_1 and C_2 where $C_m = \left\{\omega_m^{(1)}, \omega_m^{(2)}, \ldots, \omega_m^{(P)}\right\}$, m = 1,2, thus, δ can be defined as

$$\delta \triangleq \sum_{p=1}^{P} \left|\omega_2^{(p)} - \omega_1^{(p)}\right|. \quad (9)$$

First, note that the proposed MSRL describes well the hypothesis test (8) (i.e., δ = 0 means that the two emitted signal sources are combined into one signal and $\delta \neq 0$ the two signals are resolvable). Second, since the MSRL δ is unknown, it is impossible to design an optimal detector in the Neyman-Pearson sense. Alternatively, the GLRT [28,35] is a well-known approach appropriate to solve such a problem. To conduct the GLRT on (8), one has to express the probability density function (pdf) of (5) w.r.t. δ. Assuming (without loss of generality) that $\omega_1^{(1)} > \omega_2^{(1)}$, one can notice that $\boldsymbol{\xi}$ is known if and only if δ and $\boldsymbol{\vartheta} \triangleq \left[\omega_2^{(1)} \left(\boldsymbol{\omega}^{(2)}\right)^{\mathrm{T}} \ldots \left(\boldsymbol{\omega}^{(P)}\right)^{\mathrm{T}}\right]^{\mathrm{T}}$

are fixed (i.e., there is a one to one mapping between δ, $\boldsymbol{\vartheta}$, and $\boldsymbol{\xi}$). Consequently, the pdf of (5) can be described as $p(\mathbf{y}|\delta,\boldsymbol{\vartheta})$. Now, we are ready to conduct the GLRT for this problem:

$$L_G(\mathbf{y}) = \frac{\max_{\delta,\boldsymbol{\vartheta}_1} p(\mathbf{y}|\delta,\boldsymbol{\vartheta}_1,\mathcal{H}_1)}{\max_{\boldsymbol{\vartheta}_0} p(\mathbf{y}|\boldsymbol{\vartheta}_0,\mathcal{H}_0)} = \frac{p(\mathbf{y}|\hat{\delta},\hat{\boldsymbol{\vartheta}}_1,\mathcal{H}_1)}{p(\mathbf{y}|\hat{\boldsymbol{\vartheta}}_0,\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \varsigma', \tag{10}$$

where $\hat{\delta}$, $\hat{\boldsymbol{\vartheta}}_1$, and $\hat{\boldsymbol{\vartheta}}_0$ denote the maximum likelihood estimates (MLE) of δ under $\mathcal{H}_1$, the MLE of $\boldsymbol{\vartheta}$ under $\mathcal{H}_1$ and the MLE of $\boldsymbol{\vartheta}$ under $\mathcal{H}_0$, respectively, and where ς' denotes the test threshold. From (10), one obtains

$$T_G(\mathbf{y}) = \mathrm{Ln}\, L_G(\mathbf{y}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\lessgtr}} \varsigma = \mathrm{Ln}\varsigma', \tag{11}$$

in which Ln denotes the natural logarithm.

Asymptotic equivalence of the MSRL

Finding the analytical expression of $T_G(\mathbf{y})$ in (11) is not tractable. This is mainly due to the fact that the derivation of $\hat{\delta}$ is impossible since from (2) one obtains a multimodal likelihood function [36]. Consequently, in the following, and as in[d] [13], we consider the asymptotic case (in terms of the number of snapshots). In [35, eq (6C.1)], it has been proven that, for a large number of snapshots, the statistic $T_G(\mathbf{y})$ follows a chi-square pdf under $\mathcal{H}_0$ and $\mathcal{H}_1$ given by

$$T_G(\mathbf{y}) \sim \begin{cases} \chi^2_1 & \text{under } \mathcal{H}_0, \\ \chi'^2_1(\kappa'(P_{\mathrm{fa}},P_{\mathrm{d}})) & \text{under } \mathcal{H}_1, \end{cases} \tag{12}$$

where χ^2_1 and $\chi'^2_1(\kappa'(P_{\mathrm{fa}},P_{\mathrm{d}}))$ denote the central chi-square and the noncentral chi-square pdf with one degree of freedom, respectively. P_{fa} and P_{d} are, respectively, the probability of false alarm and the probability of detection of the test (8). In the following, CRB(δ) denotes the CRB for the parameter δ where the unknown vector parameter is given by $[\delta\ \boldsymbol{\vartheta}^{\mathrm{T}}]^{\mathrm{T}}$. Consequently, assuming that CRB(δ) exists (under $\mathcal{H}_0$ and $\mathcal{H}_1$), is well defined (see section "MSRL closed-form expression" for the necessary[e] and sufficient conditions) and is a tight bound (i.e., achievable under quite general/weak conditions [36,37]), thus the noncentral parameter $\kappa'(P_{\mathrm{fa}}, P_{\mathrm{d}})$ is given by [[35], p. 239]

$$\kappa'(P_{\mathrm{fa}},P_{\mathrm{d}}) = \delta^2(\mathrm{CRB}(\delta))^{-1}. \tag{13}$$

On the other hand, one can notice that the noncentral parameter $\kappa'(P_{\mathrm{fa}}, P_{\mathrm{d}})$ can be determined numerically by the choice of P_{fa} and P_{d} [13,28] as the solution of

$$\mathcal{Q}^{-1}_{\chi^2_1}(P_{\mathrm{fa}}) = \mathcal{Q}^{-1}_{\chi'^2_1(\kappa'(P_{\mathrm{fa}},P_{\mathrm{d}}))}(P_{\mathrm{d}}), \tag{14}$$

in which $\mathcal{Q}^{-1}_{\chi^2_1}(\varpi)$ and $\mathcal{Q}^{-1}_{\chi'^2_1(\kappa'(P_{\mathrm{fa}},P_{\mathrm{d}}))}(\varpi)$ are the inverse of the right tail of the χ^2_1 and $\chi'^2_1(\kappa'(P_{\mathrm{fa}},P_{\mathrm{d}}))$ pdf starting at the value ϖ. Finally, from (13) and (14) one obtains[f]

$$\delta = \kappa(P_{\mathrm{fa}},\ P_{\mathrm{d}})\sqrt{\mathrm{CRB}(\delta)}, \tag{15}$$

where $\sqrt{\kappa(P_{\mathrm{fa}},P_{\mathrm{d}})} = \kappa'(P_{\mathrm{fa}},P_{\mathrm{d}})$ is the so-called translation factor [13] which is determined for a given probability of false alarm and probability of detection (see Figure 1 for the behavior of the translation factor versus P_{fa} and P_{d}).

Result 1: The asymptotic MSRL for model (5) in the case of P parameters of interest per signal ($P \geq 1$) is given by δ which is the solution of the following equation:

$$\delta^2 - \kappa^2(P_{\mathrm{fa}},P_{\mathrm{d}})(A_{\mathrm{direct}} + A_{\mathrm{cross}}) = 0, \tag{16}$$

where A_{direct} denotes the contribution of the parameters of interest belonging to the same dimension as follows

$$A_{\mathrm{direct}} = \sum_{p=1}^{P} \mathrm{CRB}(\omega_1^{(p)}) + \mathrm{CRB}(\omega_2^{(p)}) - 2\mathrm{CRB}(\omega_1^{(p)},\omega_2^{(p)}),$$

and where A_{cross} is the contribution of the cross terms between distinct dimension given by

$$A_{\mathrm{cross}} = \sum_{p=1}^{P}\sum_{\substack{p'=1 \\ p'\neq p}}^{P} g_p g_{p'}(\mathrm{CRB}(\omega_1^{(p)},\omega_1^{(p')}) + \mathrm{CRB}(\omega_2^{(p)},\omega_2^{(p')}) - 2\mathrm{CRB}(\omega_1^{(p)},\omega_2^{(p')})),$$

in which $g_p = \mathrm{sgn}\left(\omega_1^{(p)} - \omega_2^{(p)}\right)$.

Proof see Appendix 1.

Remark 1: It is worth noting that the hypothesis test (8) is a binary one-sided test and that the MLE used is

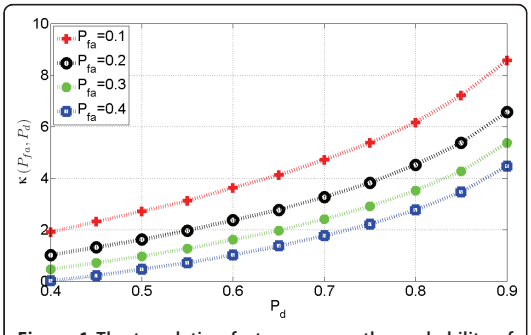


Figure 1 The translation factor κ versus the probability of detection P_{d} and P_{fa}. One can notice that increasing P_{d} or decreasing P_{fa} has the effect to increase the value of the translation factor κ. This is expected since increasing P_{d} or decreasing P_{fa} leads to a more selective decision [28,35].

an unconstrained estimator. Thus, one can deduce that the GLRT, used to derive the asymptotic MSRL [13,35]: (i) is the asymptotically uniformly most powerful test among all invariant statistical tests, and (ii) has an asymptotic constant false-alarm rate (CFAR). Which is, in the asymptotic case, considered as the *strongest* statement of optimality that one could expect to obtain [28].

- *Existence of the MSRL*: It is natural to assume that the CRB is a non-increasing (i.e., decreasing or constant) function on $\mathbb{R}^+$ w.r.t. δ since it is more difficult to estimate two closely spaced signals than two largely-spaced ones. In the same time the left hand side of (15) is a monotonically increasing function w. r.t. δ on $\mathbb{R}^+$. Thus for a fixed couple $(P_{\text{fa}}, P_{\text{d}})$, the solution of the implicit equation given by (15) always exists. However, theoretically, there is no assurance that the solution of equation (15) is unique.
- Note that, in practical situation, the case where $\text{CRB}(\delta)$ is not a function of δ is important since in this case, $\text{CRB}(\delta)$ is constant w.r.t. δ and thus the solution of (15) exists and is unique (see section "MSRL closed-form expression").

In the following section, we study the explicit effect of this so-called translation factor.

The relationship between the MSRL based on the CRB and the hypothesis test approaches

In this section, we link the asymptotic MSRL (derived using the hypothesis test approach, see Result 1) to a new proposed extension of the SRL based on the Smith criterion [11]. First, we recall that the Smith criterion defines the SRL in the case of $P = 1$ only. Then, we extend this criterion to $P \geq 1$ (i.e., the case of the multidimensional harmonic model). Finally, we link the MSRL based on the hypothesis test approach (see Result 1) to the MSRL based on the CRB approach (i.e., the extended SRL based on the Smith criterion).

The Smith criterion: Since the CRB expresses a lower bound on the covariance matrix of any unbiased estimator, then it expresses also the ultimate estimation accuracy. In this context, Smith proposed the following criterion for the case of two source signals parameterized each one by only one frequency [11]: *two signals are resolvable if the difference between their frequency*, $\delta_{\omega^{(1)}} = \omega_2^{(1)} - \omega_1^{(1)}$, *is greater than the standard deviation of the frequency difference estimation*. Since, the standard deviation can be approximated by the CRB, then, the SRL, in the Smith criterion sense, is defined as the limit of $\delta_{\omega^{(1)}}$ for which $\delta_{\omega^{(1)}} < \sqrt{\text{CRB}(\delta_{\omega^{(1)}})}$ is achieved. This means that, the SRL is the solution of the following implicit equation

$$\delta^2_{\omega^{(1)}} = \text{CRB}(\delta_{\omega^{(1)}}).$$

The extension of the Smith criterion to the case of $P \geq 1$: Based on the above framework, a straightforward extension of the Smith criterion to the case of $P \geq 1$ for the multidimensional harmonic model is as follows: *two multidimensional harmonic retrieval signals are resolvable if the distance between C_1 and C_2, is greater than the standard deviation of the δ_{CRB} estimation*. Consequently, assuming that the CRB exists and is well defined, the MSRL δ_{CRB} is given as the solution of the following implicit equation

$$\begin{cases} \delta^2_{\text{CRB}} = \text{CRB}(\delta_{\text{CRB}}) \\ \text{s.t.} \quad \delta_{\text{CRB}} = \sum_{p=1}^{P} |\omega_2^{(p)} - \omega_1^{(P)}|. \end{cases} \tag{17}$$

Comparison and link between the MSRL based on the CRB approach and the MSRL based on the hypothesis test approach: The MSRL based on the hypothesis test approach is given as the solution of

$$\begin{cases} \delta = \kappa(P_{\text{fa}}, P_{\text{d}})\sqrt{\text{CRB}(\delta)}, \\ \text{s.t.} \quad \delta = \sum_{p=1}^{P} \left|\omega_2^{(p)} - \omega_1^{(p)}\right|, \end{cases}$$

whereas the MSRL based on the CRB approach is given as the solution of (17). Consequently, one has the following result:

Result 2: Upon to a translation factor, the asymptotic MSRL based on the hypothesis test approach (i.e., using the binary one-sided hypothesis test given in (8)) is equivalent to the proposed MSRL based on the CRB approach (i. e., using the extension of the Smith criterion). Consequently, the criterion given in (17) is equivalent to an asymptotically uniformly most powerful test among all invariant statistical tests for $\kappa(P_{\text{fa}}, P_{\text{d}}) = 1$ (see Figure 2 for the values of $(P_{\text{fa}}, P_{\text{d}})$ such that $\kappa\ (P_{\text{fa}}, P_{\text{d}}) = 1$).

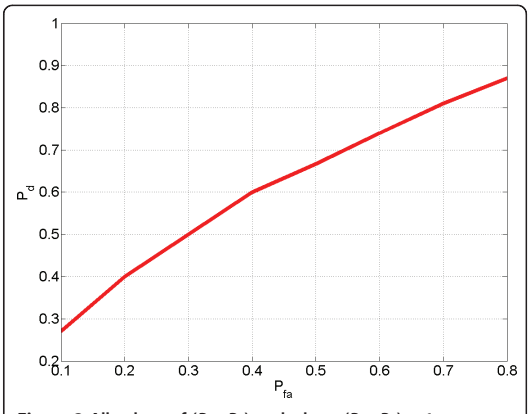


Figure 2 All values of $(P_{\text{fa}}, P_{\text{d}})$ such that $\kappa(P_{\text{fa}}, P_{\text{d}}) = 1$.

The following section is dedicated to the analytical computation of closed-form expression of the MSRL. In section "Assumptions," we introduce the assumptions used to compute the MSRL in the case of a Gaussian random noise and orthogonal waveforms. Then, we derive non matrix closed-form expressions of the CRB (note that to the best of our knowledge, no closed-form expressions of the CRB for such model is available in the literature). In "MSRL derivation" and thanks to these expressions, the MSRL wil be deduced using (16). Finally, the MSRL analysis is given.

MSRL closed-form expression

in section "Determination of the MSRL for two sources" we have defined the general model of the multidimensional harmonic model. To derive a closed-form expression of the MSRL, we need more assumptions on the covariance noise matrix and/or on the signal sources.

Assumptions

- The noise is assumed to be a complex circular white Gaussian random process i.i.d. with zero-mean and unknown variance $\sigma^2\mathbf{I}_{LN_1\dots N_P}$.
- We consider a multidimensional harmonic model due to the superposition of two harmonics each of them of dimension $P \geq 1$. Furthermore, for sake of simplicity and clarity, the sources have been assumed known and orthogonal (e.g., [7,38]). In this case, the unknown parameter vector is fixed and does not grow with the number of snapshots. Consequently, the CRB is an achievable bound [36].
- Each parameter of interest w.r.t. to the first signal, $\omega_1^{(p)}$ $p = 1 \dots P$, can be as close as possible to the parameter of interest w.r.t. to the second signal $\omega_2^{(p)}$ $p = 1 \dots P$, but not equal. This is not really a restrictive assumption, since in most applications, having two or more identical parameters of interest is a zero probability event [[9], p. 53].

Under these assumptions, the joint probability density function of the noisy observations y for a given unknown deterministic parameter vector $\boldsymbol{\xi}$ is as follows:

$$p(\mathbf{y}|\boldsymbol{\xi}) = \prod_{t=1}^{L} p(\mathrm{vec}(\boldsymbol{\mathcal{Y}}(t))|\boldsymbol{\xi}) = \frac{1}{(\pi\sigma^2)^{LN}} e^{\frac{-1}{\sigma^2}(\mathbf{y}-\mathbf{x})^{\mathrm{H}}(\mathbf{y}-\mathbf{x})},$$

where $N = \prod_{p=1}^{P} N_p$. The multidimensional harmonic retrieval model with known sources is considered herein, and thus, the parameter vector is given by

$$\boldsymbol{\xi} = \left[\boldsymbol{\omega}^{\mathrm{T}} \sigma^2\right]^{\mathrm{T}}, \tag{18}$$

where

$$\boldsymbol{\omega} = \left[(\boldsymbol{\omega}^{(1)})^{\mathrm{T}} \cdots (\boldsymbol{\omega}^{(P)})^{\mathrm{T}}\right]^{\mathrm{T}},$$

in which

$$\boldsymbol{\omega}^{(p)} = \left[\omega_1^{(p)} \omega_2^{(p)}\right]^{\mathrm{T}}. \tag{19}$$

CRB for the multidimensional harmonic model with orthogonal known signal sources

The Fisher information matrix (FIM) of the noisy observations $\mathbf{y}$ w.r.t. a parameter vector $\boldsymbol{\xi}$ is given by [39]

$$\mathbf{FIM}(\boldsymbol{\xi}) = \mathbb{E}\left\{\frac{\partial \ln p(\mathbf{y}|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \left(\frac{\partial \ln p(\mathbf{y}|\boldsymbol{\xi})}{\partial \boldsymbol{\xi}}\right)^{\mathrm{H}}\right\}.$$

For a complex circular Gaussian observation model, the (ith, kth) element of the FIM for the parameter vector $\boldsymbol{\xi}$ is given by [34]

$$[\mathbf{FIM}(\boldsymbol{\xi})]_{i,k} = \frac{LN}{\sigma^4}\frac{\partial\sigma^2}{\partial[\boldsymbol{\xi}]_i}\frac{\partial\sigma^2}{\partial[\boldsymbol{\xi}]_k} + \frac{2}{\sigma^2}\Re\left\{\frac{\partial \mathbf{x}^{\mathrm{H}}}{\partial[\boldsymbol{\xi}]_i}\frac{\partial \mathbf{x}}{\partial[\boldsymbol{\xi}]_k}\right\} \quad (i,k) = \{1,\dots,2P+1\}^2. \tag{20}$$

Consequently, one can state the following lemma.

Lemma 1: The FIM for the sum of two P-order harmonic models with orthogonal known sources, has a block diagonal structure and is given by

$$\mathbf{FIM}(\boldsymbol{\xi}) = \frac{2}{\sigma^2}\begin{bmatrix} \mathbf{F}_{\boldsymbol{\omega}} & \mathbf{0}_{2P\times 1} \\ \mathbf{0}_{1\times 2P} & \times \end{bmatrix}, \tag{21}$$

where, the $(2P) \times (2P)$ matrix $\mathbf{F}_{\boldsymbol{\omega}}$ is also a block diagonal matrix given by

$$\mathbf{F}_{\boldsymbol{\omega}} = LN(\boldsymbol{\Delta} \otimes \mathbf{G}), \tag{22}$$

in which $\boldsymbol{\Delta} = \mathrm{diag}\{||\boldsymbol{\alpha}_1||^2, ||\boldsymbol{\alpha}_2||^2\}$ where

$$\boldsymbol{\alpha}_m = \begin{bmatrix}\alpha_m(1) & \dots & \alpha_m(L)\end{bmatrix}^{\mathrm{T}} \quad \text{for} \quad m \in \{1,2\}, \tag{23}$$

and

$$[\mathbf{G}]_{k,l} = \begin{cases} \dfrac{(2N_k-1)(N_k-1)}{6} & \text{for} \quad k = l, \\ \dfrac{(N_k-1)(N_l-1)}{2} & \text{for} \quad k \neq l. \end{cases}$$

Proof see Appendix 2.

After some calculation and using Lemma 1, one can state the following result.

Result 3: The closed-form expressions of the CRB for the sum of two P-order harmonic models with orthogonal known signal sources are given by

$$\mathrm{CRB}(\omega_m^{(p)}) = \frac{6}{LN\mathrm{SNR}_m} C_p, \quad m \in \{1,2\}, \tag{24}$$

where $\mathrm{SNR}_m = \frac{||\boldsymbol{\alpha}_m||^2}{\sigma^2}$ denotes the SNR of the mth source and where

$$C_p = \frac{N_p(1-3V_P)+3V_P+1}{(N_p+1)(N_p^2-1)} \quad \text{in which} \quad V_P = \frac{1}{1+3\sum_{p=1}^{P}\frac{N_p-1}{N_p+1}}.$$

Furthermore, the cross-terms are given by

$$\mathrm{CRB}(\omega_m^{(p)}, \omega_{m'}^{(p')}) = \begin{cases} 0 & \text{for } m \neq m', \\ \dfrac{-6}{LN\mathrm{SNR}_m}\tilde{C}_{p,p'} & \text{for } m = m' \text{ and } p \neq p', \end{cases} \quad (25)$$

where

$$\tilde{C}_{p,p'} = \frac{3V_P}{(N_p+1)(N_{p'}+1)}.$$

Proof see Appendix 3.

MSRL derivation

Using the previous result, one obtains the unique solution of (16), thus, the MSRL for model (1) is given by the following result:

Result 4: The MSRL for the sum of P-order harmonic models with orthogonal known signal sources, is given by

$$\delta = \sqrt{\frac{6}{LN\mathrm{ESNR}}\left(\sum_{p=1}^{P} C_p - \sum_{\substack{p,p'=1 \\ p\neq p'}}^{P} g_p g_{p'} \tilde{C}_{p,p'}\right)}, \quad (26)$$

where the so-called extended SNR is given by $\mathrm{ESNR} = \frac{\mathrm{SNR}_1\mathrm{SNR}_2}{\mathrm{SNR}_1+\mathrm{SNR}_2}$.

Proof see Appendix 4.

Numerical analysis

Taking advantage of the latter result, one can analyze the MSRL given by (26):

- First, from Figure 3 note that the numerical solution of the MSRL based on (12) is in good agreement with the analytical expression of the MSRL (23), which validate the closed-form expression given in (23). On the other hand, one can notice that, for $P_d = 0.37$ and $P_{fa} = 0.1$ the MSRL based on the CRB is exactly equal to the MSRL based on hypothesis test approach derived in the asymptotic case. From the case $P_d = 0.49$ and $P_{fa} = 0.3$ or/and $P_d = 0.32$ and $P_{fa} = 0.1$, one can notice the influence of the translation factor $\kappa(P_{fa}, P_d)$ on the MSRL.

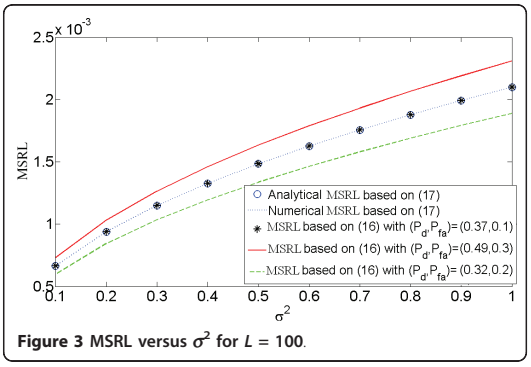


Figure 3 MSRL versus σ^2 for $L = 100$.

- The MSRL^g is $O\left(\sqrt{\frac{1}{\mathrm{ESNR}}}\right)$ which is consistent with some previous results for the case $P = 1$ (e.g., [12,14,24]).
- From (26) and for a large number of sensors $N_1 = N_2 = \ldots = N_P = N \gg 1$, one obtains a simple expression

$$\delta = \sqrt{\frac{12}{LN^{P+1}\mathrm{ESNR}}\frac{P}{1+3P}},$$

meaning that, the SRL is $O\left(\sqrt{\frac{1}{N^{P+1}}}\right)$.

- Furthermore, since P ≥ 1, one has

$$\frac{(P+1)(3P+1)}{P(3P+4)} < 1,$$

and consequently, the ratio between the MSRL of a multidimensional harmonic retrieval with P parameters of interest, denoted by δ_P and the MSRL of a multidimensional harmonic retrieval with $P + 1$ parameters of interest, denoted by δ_{P+1}, is given by

$$\frac{\delta_{P+1}}{\delta_P} = \sqrt{\frac{(P+1)(3P+1)}{NP(3P+4)}}, \quad (27)$$

meaning that the MSRL for $P + 1$ parameters of interest is less than the one for P parameters of interest (see Figure 4). This, can be explained by the estimation additional parameter and also by an increase of the received noisy data thanks to the additional dimension. One should note that this property is proved theoretically thanks to (27) using the assumption of an equal and large number of sensors. However, from Figure 4 we notice that, in practice, this can be verified even for a

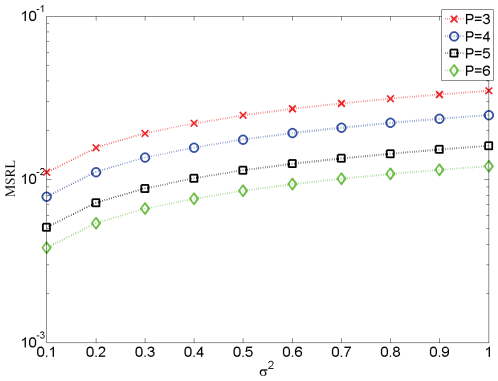


Figure 4 The SRL for multidimensional harmonic retrieval with orthogonal known sources for *M* equally powered sources, where *P* = 3, 4, 5, 6, *L* = 100, and the numbers of sensors are given by $N_1 = 3$, $N_2 = 5$, $N_3 = 4$, $N_4 = 4$, $N_5 = 4$, and $N_6 = 3$.

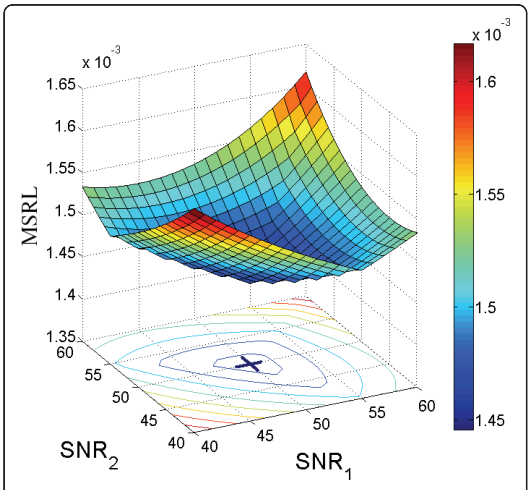


Figure 5 MSRL versus SNR_1, the SNR of the first source, and SNR_2, the SNR of the second source. One can notice that the optimal distribution of the SNR (which corresponds to the lowest MSLR) corresponds to $\mathrm{SNR}_1 = \mathrm{SNR}_2 = \frac{\mathrm{SNR}_{\mathrm{total}}}{2}$ as predicted by Corollary 1.

small number of sensors (e.g., in Figure 4 one has $3 \leq N_p \leq 5$ for $p = 3, ..., 6$).

- Furthermore, since

$$\sqrt{\frac{4}{LN^{P+1}\mathrm{ESNR}}} \leq \delta_P < \delta_{P-1} < \cdots < \delta_1$$

one can note that, the SRL is lower bounded by $\sqrt{\frac{4}{LN^{P+1}\mathrm{ESNR}}}$.

- One can address the problem of finding the optimal distribution of power sources making the SRL the smallest as possible (s.t. the constraint of constant total source power). In this issue, one can state the following corollary: *Corollary 1:* The optimal power's source distribution that ensures the smallest MSRL is obtained only for the equi-powered sources case.

Proof see Appendix 5.

This result was observed numerically for $P = 1$ in [12] (see Figure 5 for the multidimensional harmonic model). Moreover, it has been shown also by simulation for the case $P = 1$ that the so-called maximum likelihood breakdown (i.e., when the mean square error of the MLE increases rapidly) occurs at higher SNR in the case of different power signal sources than in the case of equi-powered signal sources [40]. The authors explained it by the fact that one source grabs most of the total power, then, this latter will be estimated more accurately, whereas the second one, will take an arbitrary parameter estimation which represents an outlier.

- In the same way, let us consider the problem of the optimal placement of the sensors[h] $N_1, ..., N_P$, making the minimum MSRL s.t. the constraint that the total number of sensors is constant (i.e., $N_{\mathrm{total}} = \sum_{p=1}^{P} N_p$ in which we suppose that N_{total} is a multiple of P).

Corollary 2: If the total number of sensors N_{total}, is a multiple of P, then an optimal placement of the sensors that ensure the lowest MSRL is (see Figure 6 and 7)

$$N_1 = \cdots = N_P = \frac{N_{\mathrm{total}}}{P}. \tag{28}$$

Proof see Appendix 6.

Remark 3: Note that, in the case where N_{total} is not a multiple of P, one expects that the optimal MSRL is given in the case where the sensors distribution approaches the equi-sensors distribution situation given in corollary 3. Figure 7 confirms that (in the case of $P = 3$, $N_1 = 8$ and a total number of sensors $N = 22$). From Figure 7, one can notice that the optimal distribution of the number of sensors corresponds to $N_2 = N_3 = 7$ and $N_1 = 8$ which is the nearest situation to the equi-sensors distribution.

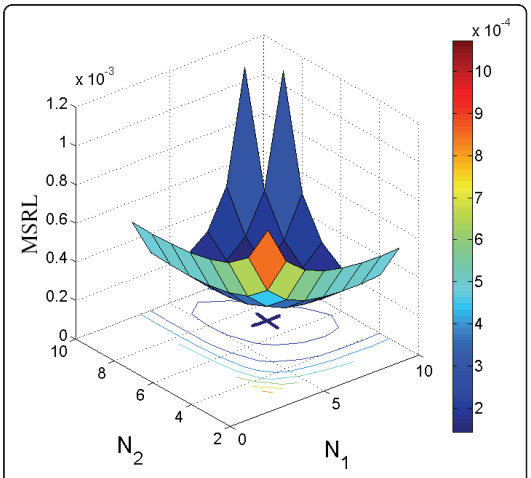


Figure 6 The MSRL versus N_1 and N_2 in the case of $P = 3$ and a total number of sensors $N_{\text{total}} = 21$. One can notice that the optimal distribution of the number of sensors (which corresponds to the lowest SLR) corresponds to $N_1 = N_2 = N_3 = \frac{N_{\text{total}}}{3}$ as predicted by (28).

Conclusion

In this article, we have derived the MSRL for the multi-dimensional harmonic retrieval model. Toward this end, we have extended the concept of SRL to multiple parameters of interest per signal. First, we have used a hypothesis test approach. The applied test is shown to be asymptotically an uniformly most powerful test which is the *strongest* statement of optimality that one could hope to obtain. Second, we have linked the asymptotic MSRL based on the hypothesis test approach to a new extension of the SRL based on the Cramér-Rao bound approach. Using the Cramér-Rao bound and a proper change of variable formula, closed-form expression of the MSRL are given.

Finally, note that the concept of the MSRL can be used to optimize, for example, the waveform and/or the array geometry for a specific problem.

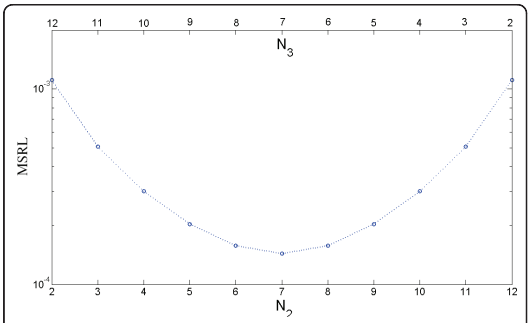


Figure 7 The plot of the MSRL versus N_2 in the case of $P = 3$, $N_1 = 8$ and a total number of sensors $N = 22$.

Appendix 1

The proof of Result 1

Appendix 1.1: In this appendix, we derive the MSRL using the l_1 norm.

From $\mathbf{CRB}(\boldsymbol{\xi})$ where $\boldsymbol{\xi} = [\boldsymbol{\omega}^{\mathrm{T}}\ \boldsymbol{\rho}^{\mathrm{T}}]^{\mathrm{T}}$ in which $\boldsymbol{\omega} = [\omega_1^{(1)}\omega_2^{(1)}\omega_1^{(2)}\omega_2^{(2)}\cdots\omega_1^{(P)}\omega_2^{(P)}]^{\mathrm{T}}$, one can deduce $\mathbf{CRB}(\breve{\boldsymbol{\xi}})$ where $\breve{\boldsymbol{\xi}} = \mathbf{g}(\boldsymbol{\xi}) = [\delta\ \boldsymbol{\vartheta}^{\mathrm{T}}]^{\mathrm{T}}$ in which $\boldsymbol{\vartheta} \triangleq [\omega_2^{(1)}(\boldsymbol{\omega}^{(2)})^{\mathrm{T}}\cdots(\boldsymbol{\omega}^{(P)})^{\mathrm{T}}]^{\mathrm{T}}$. Thanks to the Jacobian matrix given by

$$\frac{\partial \mathbf{g}(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} = \begin{bmatrix} \mathbf{h}^{\mathrm{T}} & \mathbf{0} \\ \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix},$$

where $\mathbf{h} = [g_1 g_2\ \ldots\ g_P\ ]^{\mathrm{T}} \otimes [1\ -\ 1]^{\mathrm{T}}$, in which $g_p = \frac{\partial \delta}{\partial \omega_1^{(p)}} = -\frac{\partial \delta}{\partial \omega_2^{(p)}} = \mathrm{sgn}\left(\omega_1^{(p)} - \omega_2^{(p)}\right)$ and $\mathbf{A} = [\mathbf{0}\ \mathbf{I}]$. Using the change of variable formula

$$\mathbf{CRB}(\breve{\boldsymbol{\xi}}) = \frac{\partial \mathbf{g}(\breve{\boldsymbol{\xi}})}{\partial \breve{\boldsymbol{\xi}}} \mathbf{CRB}(\boldsymbol{\xi}) \left(\frac{\partial \mathbf{g}(\breve{\boldsymbol{\xi}})}{\partial \breve{\boldsymbol{\xi}}}\right)^{\mathrm{T}}, \tag{29}$$

one has

$$\mathbf{CRB}(\breve{\boldsymbol{\xi}}) = \begin{bmatrix} \mathbf{h}^{\mathrm{T}}\mathbf{CRB}(\boldsymbol{\omega})\mathbf{h} & \times \\ \times & \mathbf{I} \end{bmatrix}.$$

Consequently, after some calculus, one obtains

$$\begin{aligned} \mathbf{CRB}(\delta) &\triangleq [\mathbf{CRB}(\breve{\boldsymbol{\xi}})]_{1,1} = \mathbf{h}^{\mathrm{T}}\mathbf{CRB}(\boldsymbol{\omega})\mathbf{h} \\ &= \sum_{p=1}^{2P}\sum_{p'=1}^{2P} [\mathbf{h}]_p [\mathbf{h}]_{p'} [\mathbf{CRB}(\boldsymbol{\omega})]_{p,p'} \\ &= \sum_{p=1}^{P}\sum_{p'=1}^{P} g_p g_{p'} \left([\mathbf{CRB}(\boldsymbol{\xi})]_{2p,2p'} + [\mathbf{CRB}(\boldsymbol{\xi})]_{2p-1,2p'-1} - [\mathbf{CRB}(\boldsymbol{\xi})]_{2p,2p'-1} - [\mathbf{CRB}(\boldsymbol{\xi})]_{2p-1,2p'}\right) \\ &\triangleq A_{\text{direct}} + A_{\text{cross}}, \end{aligned} \tag{30}$$

where

$A_{\text{direct}} = \sum_{p=1}^{P} \mathrm{CRB}(\omega_1^{(p)}) + \mathrm{CRB}(\omega_2^{(p)}) - 2\mathrm{CRB}(\omega_1^{(p)}, \omega_2^{(p)})$

and where $A_{\text{cross}}(k) = \sum_{p=1}^{P}\sum_{\substack{p'=1 \\ p'\neq p}}^{P} g_p g_{p'} \left(\mathrm{CRB}(\omega_1^{(p)}, \omega_1^{(p')}) + \mathrm{CRB}(\omega_2^{(p)}, \omega_2^{(p')}) - 2\mathrm{CRB}(\omega_1^{(p)}, \omega_2^{(p')})\right)$

Finally using (30) one obtains (16)

Appendix 1.2: In this part, we derive the MSRL using the l_k norm for a given integer $k \geq 1$. The aim of this part is to support the endnote a, which stays that using the l_1 norm computing the MSRL using the l_1 norm is for the calculation convenience.

Once again, from $\mathbf{CRB}(\boldsymbol{\xi})$, one can deduce $\mathbf{CRB}(\breve{\boldsymbol{\xi}}_k)$ where $\breve{\boldsymbol{\xi}}_k = \mathbf{g}_k(\boldsymbol{\xi}) = [\delta(k)\ \boldsymbol{\vartheta}^{\mathrm{T}}]^{\mathrm{T}}$ in which the distance between C_1 and C_2 using the l_k norm is given by $\delta(k) \triangleq$

k-norm distance$(\mathcal{C}_1, \mathcal{C}_2) = \left(\sum_{p=1}^{P} \delta_p^k\right)^{1/k}$ and where $\boldsymbol{\vartheta} \triangleq [\omega_2^{(1)}(\boldsymbol{\omega}^{(2)})^{\mathrm{T}} \dots (\boldsymbol{\omega}^{(P)})^{\mathrm{T}}]^{\mathrm{T}}$. The Jacobian matrix is given by

$$\frac{\partial \mathbf{g}(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} = \begin{bmatrix} \mathbf{h}_k^{\mathrm{T}} & \mathbf{0} \\ \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix},$$

where $\mathbf{h}_k = [1 \; -1]^{\mathrm{T}} \otimes [g_1(k) g_2(k) \dots g_P(k)]^{\mathrm{T}}$, in which $g_p(k) = \dfrac{\partial \delta(k)}{\partial \omega_1^{(p)}} = -\dfrac{\partial \delta(k)}{\partial \omega_2^{(p)}}$ and $\mathbf{A} = [\mathbf{0}\mathbf{I}]$. Since $|x|^k$ can be written as $\sqrt{x^{2k}}$. Thus, for $\times \neq 0$, one has

$$g_p(k) = \frac{\partial\left(\sum_{p'=1}^{P}\sqrt{\left(\omega_1^{(p')}-\omega_2^{(p')}\right)^{2k}}\right)^{1/k}}{\partial\omega_1^{(p)}} = \frac{1}{k}\left(\sum_{i=1}^{P}\sqrt{\left(\omega_1^{(i)}-\omega_2^{(i)}\right)^{2k}}\right)^{\frac{1}{k}-1}\frac{\partial\sqrt{\left(\omega_1^{(i)}-\omega_2^{(i)}\right)^{2k}}}{\partial\omega_1^{(i)}}$$
$$= \operatorname{sgn}(\omega_1^{(p)}-\omega_2^{(p)})\left(\sum_{p=1}^{P}\sqrt{\left(\omega_1^{(p')}-\omega_2^{(p')}\right)^{2k}}\right)^{\frac{1}{k}-1}\sqrt{\left(\omega_1^{(p)}-\omega_2^{(p)}\right)^{2(k-1)}} = \operatorname{sgn}(\omega_1^{(p)}-\omega_2^{(p)})\delta^{1-k}\delta_p^{k-1}. \quad (31)$$

Again, using the change of variable formula (29), one has

$$\mathbf{CRB}(\breve{\boldsymbol{\xi}}_k) = \begin{bmatrix} \mathbf{h}_k^{\mathrm{T}}\mathbf{CRB}(\boldsymbol{\omega})\mathbf{h}_k & \times \\ \times & \mathbf{I} \end{bmatrix}.$$

Consequently, after some calculus, one obtains

$$\begin{aligned}\mathrm{CRB}(\delta(k)) &\triangleq [\mathbf{CRB}(\boldsymbol{\xi}_\delta)]_{1,1} \\ &= \sum_{p=1}^{P}\sum_{p'=1}^{P} g_{p'}(k) g_{p'}(k)([\mathbf{CRB}(\boldsymbol{\xi})]_{2p,2p'} + [\mathbf{CRB}(\boldsymbol{\xi})]_{2p-1,2p'-1} - [\mathbf{CRB}(\boldsymbol{\xi})]_{2p,2p'-1} - [\mathbf{CRB}(\boldsymbol{\xi})]_{2p-1,2p'}) \\ &= (\delta(k))^{2(1-k)}(A_{\mathrm{direct}}(k) + A_{\mathrm{cross}}(k)),\end{aligned} \quad (32)$$

where

$A_{\mathrm{direct}}(k) = \sum_{p=1}^{P} \delta_p^{2(k-1)}\left(\mathrm{CRB}(\omega_1^{(p)}) + \mathrm{CRB}(\omega_2^{(p)}) - 2\mathrm{CRB}(\omega_1^{(p)}, \omega_2^{(p)})\right)$ and where $A_{\mathrm{cross}}(k) = \sum_{p=1}^{P}\sum_{p'\neq p}^{P} \delta_p^{k-1}\delta_{p'}^{k-1}\operatorname{sgn}(\omega_1^{(p)}-\omega_2^{(p)})\operatorname{sgn}(\omega_1^{(p')}-\omega_2^{(p')})\left(\mathrm{CRB}(\omega_1^{(p)}, \omega_1^{(p')}) + \mathrm{CRB}(\omega_2^{(p)}, \omega_2^{(p')}) - 2\mathrm{CRB}(\omega_1^{(p)}, \omega_2^{(p')})\right)$.

Consequently, note that resolving analytically the implicit equation (32) w.r.t. $\delta(k)$ is intractable (aside from some special cases). Whereas, resolving analytically the implicit equation (30) can be tedious but feasible (see section "MSRL closed form expression").

Furthermore, denoting $g_p(1) = g_p$, $A_{\mathrm{cross}}(1) \triangleq A_{\mathrm{cross}}$ and $A_{\mathrm{direct}}(1) \triangleq A_{\mathrm{direct}}$ and using (32) one obtains (16).

Appendix 2

Proof of Lemma 1

From (20) one can note the well-known property that the model signal parameters are decoupled from the noise variance [42]. Consequently, the block-diagonal structure in (21) is self-evident.

Now, let us prove (22). From (4), one obtains

$$\frac{\partial \mathrm{vec}(\boldsymbol{\mathcal{X}}(t))}{\partial \omega_m^{(p)}} = j s_m(t)\left(\boldsymbol{a}(\omega_m^{(1)}) \otimes \boldsymbol{a}(\omega_m^{(2)}) \otimes \cdots \otimes \boldsymbol{a}'(\omega_m^{(p)}) \otimes \cdots \otimes \boldsymbol{a}(\omega_m^{(P)})\right),$$

where

$$\boldsymbol{a}'(\omega_m^{(p)}) = \begin{bmatrix} 0 & e^{j\omega_m^{(p)}} & \dots & (Np-1)e^{j(Np-1)\omega_m^{(p)}} \end{bmatrix}^{\mathrm{T}}.$$

Thus,

$$\frac{\partial \mathbf{x}}{\partial \omega_m^{(p)}} = j s_m \otimes \left(\boldsymbol{a}(\omega_m^{(1)}) \otimes \boldsymbol{a}(\omega_m^{(2)}) \otimes \cdots \otimes \boldsymbol{a}'(\omega_m^{(p)}) \otimes \cdots \otimes \boldsymbol{a}(\omega_m^{(P)})\right),$$

where $s_m = [s_m(1) \dots s_m(L)]^{\mathrm{T}}$. Using the distributivity of the Hermitian operator over the Kronecker product and the mixed-product property of the Kronecker product [43] and assuming, without loss of generality that $p' < p$, one obtains

$$\begin{aligned}(\frac{\partial \mathbf{x}}{\partial \omega_m^{(p)}})^{\mathrm{H}}\frac{\partial \mathbf{x}}{\partial \omega_m^{(p')}} &= \left(\boldsymbol{s}_m^{\mathrm{H}} \otimes \left[\boldsymbol{a}^{\mathrm{H}}(\omega_m^{(1)}) \otimes \boldsymbol{a}^{\mathrm{H}}(\omega_m^{(2)}) \otimes \cdots \otimes \boldsymbol{a}'^{\mathrm{H}}(\omega_m^{(p')}) \otimes \cdots \otimes \boldsymbol{a}^{\mathrm{H}}(\omega_m^{(P)})\right]\right) \\ &\quad \times \left(\boldsymbol{s}_m \otimes \left[\boldsymbol{a}(\omega_m^{(1)}) \otimes \boldsymbol{a}(\omega_m^{(2)}) \otimes \cdots \otimes \boldsymbol{a}'(\omega_m^{(p)}) \otimes \cdots \otimes \boldsymbol{a}(\omega_m^{(P)})\right]\right) \\ &= (\boldsymbol{s}_m^{\mathrm{H}}, \boldsymbol{s}_m) \otimes \left(\boldsymbol{a}^{\mathrm{H}}(\omega_m^{(1)})\boldsymbol{a}(\omega_m^{(1)})\right) \otimes \cdots \otimes \left(\boldsymbol{a}'^{\mathrm{H}}(\omega_m^{(p)})\boldsymbol{a}(\omega_m^{(p)})\right) \otimes \dots \\ &\quad \dots \otimes \left(\boldsymbol{a}^{\mathrm{H}}(\omega_m^{(p')})\boldsymbol{a}'(\omega_m^{(p')})\right) \otimes \cdots \otimes \left(\boldsymbol{a}^{\mathrm{H}}(\omega_m^{(P)})\boldsymbol{a}(\omega_m^{(P)})\right).\end{aligned} \quad (33)$$

On the other hand, one has

$$\boldsymbol{a}^{\mathrm{H}}(\omega_m^{(p)})\boldsymbol{a}(\omega_m^{(p)}) = N_p, \quad (34)$$

whereas

$$\boldsymbol{a}^{\mathrm{H}}(\omega_m^{(p)})\boldsymbol{a}'(\omega_m^{(p)}) = \frac{N_p(N_p-1)}{2} \quad \text{and} \quad \boldsymbol{a}'^{\mathrm{H}}(\omega_m^{(p)})\boldsymbol{a}'(\omega_m^{(p)}) = \frac{N_p(2N_p-1)(N_p-1)}{6} \quad (35)$$

Finally, assuming known orthogonal wavefronts [38] (i.e., $\boldsymbol{s}_m^{\mathrm{H}}\boldsymbol{s}_m = 0$) and replacing (35) and (34) into (33), one obtains

$$\left(\frac{\partial \mathbf{x}}{\partial \omega_m^{(p)}}\right)^{\mathrm{H}}\frac{\partial \mathbf{x}}{\partial \omega_{m'}^{(p')}} = \begin{cases} 0 & \text{for } m \neq m', \\ L||\boldsymbol{\alpha}_m||^2 N\dfrac{(N_p-1)(N_{p'}-1)}{4} & \text{for } m = m' \text{ and } p \neq p', \\ L||\boldsymbol{\alpha}_m||^2 N\dfrac{(2N_p-1)(N_p-1)}{6} & \text{for } m = m' \text{ and } p = p', \end{cases} \quad (36)$$

where $\boldsymbol{\alpha}_m = [\alpha_m(1) \dots \alpha_m(L)]$ for $m \in \{1, 2\}$: Consequently, using (36), $\mathbf{F}_{\boldsymbol{\omega}}$ can be expressed as a block diagonal matrix

$$\mathbf{F}_{\boldsymbol{\omega}} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2 \end{bmatrix}, \quad (37)$$

where each $P \times P$ block $\mathbf{J}_m$ is defined by

$$\mathbf{J}_m = L||\boldsymbol{\alpha}_m||^2 N\mathbf{G}, \quad (38)$$

where

$$\mathbf{G} = \begin{bmatrix} \dfrac{(N_1-1)(2N_1-1)}{6} & \dfrac{(N_1-1)(N_2-1)}{4} & \cdots & \dfrac{(N_1-1)(N_P-1)}{4} \\ \dfrac{(N_2-1)(N_1-1)}{4} & \dfrac{(N_2-1)(2N_2-1)}{6} & \cdots & \dfrac{(N_2-1)(N_P-1)}{4} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{(N_P-1)(N_1-1)}{4} & \dfrac{(N_2P-1)(N_2-1)}{4} & \cdots & \dfrac{(N_P-1)(2N_P-1)}{6} \end{bmatrix}.$$

Consequently, from (37) and (38) one obtains (22).

Appendix 3

Proof of Result 3

Using (22) one obtains

$$\mathbf{CRB}(\boldsymbol{\omega}) = \frac{\sigma^2}{2}\mathbf{F}_{\boldsymbol{\omega}}^{-1} = \frac{\sigma^2}{2LN}(\boldsymbol{\Delta}^{-1} \otimes \mathbf{G}^{-1}) \tag{39}$$

where $\boldsymbol{\Delta}^{-1} = \text{diag}\left\{\frac{1}{||\boldsymbol{\alpha}_1||^2}, \frac{1}{||\boldsymbol{\alpha}_2||^2}\right\}$. In the following, we give a closed-form expression of $\mathbf{G}^{-1}$. One can notice that the matrix $\mathbf{G}$ has a particular structure such that it can be rewritten as the sum of a diagonal matrix and of a rank-one matrix: $\mathbf{G} = \mathbf{Q} + \boldsymbol{\gamma}\boldsymbol{\gamma}^{\mathrm{T}}$ where $\mathbf{Q} = \frac{1}{12}\text{diag}\{N_1^2 - 1, \ldots, N_P^2 - 1\}$ and $\boldsymbol{\gamma} = \frac{1}{2}[N_1 - 1, \ldots, N_P - 1]^{\mathrm{T}}$ Thanks to this particular structure, an analytical inverse of $\mathbf{G}$ can easily be obtained. Indeed, using the matrix inversion lemma

$$\begin{aligned}\mathbf{G}^{-1} &= (\mathbf{Q} + \boldsymbol{\gamma}\boldsymbol{\gamma}^{\mathrm{T}})^{-1} \\ &= \mathbf{Q}^{-1} - \frac{\mathbf{Q}^{-1}\boldsymbol{\gamma}\boldsymbol{\gamma}^{\mathrm{T}}\mathbf{Q}^{-1}}{1 + \boldsymbol{\gamma}^{\mathrm{T}}\mathbf{Q}^{-1}\boldsymbol{\gamma}}.\end{aligned} \tag{40}$$

A straightforward calculus leads to the following results,

$$\mathbf{Q}^{-1}\boldsymbol{\gamma}\boldsymbol{\gamma}^{\mathrm{T}}\mathbf{Q}^{-1} = 36\begin{bmatrix} \frac{1}{(N_1+1)^2} & \frac{1}{(N_1+1)(N_2+1)} & \cdots & \frac{1}{(N_1+1)(N_P+1)} \\ \frac{1}{(N_2+1)(N_1+1)} & \frac{1}{(N_2+1)^2} & \cdots & \frac{1}{(N_2+1)(N_P+1)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{(N_P+1)(N_1+1)} & \frac{1}{(N_P+1)(N_2+1)} & \cdots & \frac{1}{(N_P+1)^2} \end{bmatrix}, \tag{41}$$

and

$$\boldsymbol{\gamma}^{\mathrm{T}}\mathbf{Q}^{-1}\boldsymbol{\gamma} = 3\sum_{p=1}^{P}\frac{N_p - 1}{N_p + 1}. \tag{42}$$

Consequently, replacing (41) and (42) into (40), one obtains

$$[\mathbf{G}^{-1}]_{k,l} = \begin{cases} 12\frac{N_p(1 - 3V_P) + 3V_P + 1}{(N_p + 1)(N_p^2 - 1)} & \text{for} \quad k = l, \\ -\frac{36V_P}{(N_p + 1)(N_{p'} + 1)} & \text{for} \quad k \neq l, \end{cases} \tag{43}$$

where $V_P = \left(1 + 3\sum_{p=1}^{P}\frac{N_p - 1}{N_p + 1}\right)^{-1}$. Finally, replacing (43) into (39) one finishes the proof.

Appendix 4

Proof of Result 4

Using Results 1 and 3, one has

$$\begin{aligned}A_{\text{direct}} &= \sum_{p=1}^{P}\left(\text{CRB}(\omega_1^{(p)}) + \mathit{CRB}\,(\omega_2^{(p)})\right) \\ &= \frac{6\sigma^2}{LN}\left(\frac{1}{||\boldsymbol{\alpha}_1||^2} + \frac{1}{||\boldsymbol{\alpha}_2||^2}\right)\sum_{p=1}^{P}\frac{N_p(1 - 3V_P) + 3V_P + 1}{(N_p + 1)(N_p^2 - 1)},\end{aligned} \tag{44}$$

and

$$\begin{aligned}A_{\text{cross}} &= \sum_{p=1}^{P}\sum_{\substack{p'=1 \\ p' \neq p}}^{P} g_p g_{p'}\left(\text{CRB}(\omega_1^{(p)}, \omega_1^{(p')}) + \text{CRB}(\omega_2^{(p)}, \omega_2^{(p')})\right) \\ &= -\frac{6\sigma^2}{LN}\left(\frac{1}{||\boldsymbol{\alpha}_1||^2} + \frac{1}{||\boldsymbol{\alpha}_2||^2}\right)\sum_{\substack{p,p'=1 \\ p \neq p'}}^{P}\frac{3g_p g_{p'} V_P}{(N_p + 1)(N_{p'} + 1)}.\end{aligned} \tag{45}$$

Consequently, replacing (44) and (45) into (16), one finishes the proof.

Appendix 5

Proof of Corollary 1

In this appendix, we minimize the MSRL under the constraint $\text{SNR}_1 + \text{SNR}_2 = \text{SNR}_{\text{total}}$ (where $\text{SNR}_{\text{total}}$ is a real fixed value). Since, the term $\left(\sum_{p=1}^{P} C_p - \sum_{\substack{p,p'=1 \\ p \neq p'}}^{P} g_p g_{p'}\tilde{C}_{p,p'}\right)$ is independent from SNR_1 and SNR_2, minimizing δ is equivalent to minimize $\mathcal{G}(\text{SNR}_1, \text{SNR}_2)$ where

$$\mathcal{G}(\text{SNR}_1, \text{SNR}_2) = \delta^2\frac{LN}{6}\left(\sum_{p=1}^{P} C_p - \sum_{\substack{p,p'=1 \\ p \neq p'}}^{P} g_p g_{p'}\tilde{C}_{p,p'}\right)^{-1} = \frac{\text{SNR}_1 + \text{SNR}_2}{\text{SNR}_1\text{SNR}_2}.$$

Using the method of Lagrange multipliers, the problem is as follows:

$$\begin{cases} \min_{\text{SNR}_1,\text{SNR}_2}\mathcal{G}(\text{SNR}_1, \text{SNR}_2) \\ \text{s.t.} \\ \text{SNR}_1 + \text{SNR}_2 = \text{SNR}_{\text{total}} \end{cases}$$

Thus, the Lagrange function is given by $\mathcal{F}(\text{SNR}_1, \text{SNR}_2, \lambda) = \mathcal{G}(\text{SNR}_1, \text{SNR}_2) + \lambda(\text{SNR}_1 + \text{SNR}_2 - \text{SNR}_{\text{total}})$ where λ denotes the so-called Lagrange multiplier. A simple derivation leads to,

$$\frac{\partial\mathcal{F}(\text{SNR}_1, \text{SNR}_2)}{\partial\ \text{SNR}_1} = \frac{-1}{\text{SNR}_1^2} + \lambda = 0 \tag{46}$$

$$\frac{\partial\mathcal{F}(\text{SNR}_1, \text{SNR}_2)}{\partial\ \text{SNR}_2} = \frac{-1}{\text{SNR}_2^2} + \lambda = 0 \tag{47}$$

$$\frac{\partial\mathcal{F}(\text{SNR}_1, \text{SNR}_2)}{\partial\lambda} = \text{SNR}_1 + \text{SNR}_2 - \text{SNR}_{\text{total}} = 0. \tag{48}$$

Consequently, from (46) and (47), one obtains $\text{SNR}_1 = \text{SNR}_1$. Using (48), one obtains $\text{SNR}_1 = \text{SNR}_2 = \frac{\text{SNR}_{\text{total}}}{2}$. Using the constraint $\text{SNR}_1 + \text{SNR}_2 = \text{SNR}_{\text{total}}$ one deduces *corollary* 1.

Appendix 6

Minimizing δ w.r.t. N_1, ..., N_P is equivalent to minimizing the function $f(\boldsymbol{N}) = \sum_{p=1}^{P} C_p - \sum_{\substack{p,p'=1 \\ p,\neq p'}}^{P} g_p g_{p'}\tilde{C}_{p,p'}$,

where $\boldsymbol{N} = [N_1 \dots N_P]^{\mathrm{T}}$. However, since the numbers of sensors on each array, $N_1, \dots, N_P$, are integers, the derivation of $f(\boldsymbol{N})$ *w.r.t.* $\boldsymbol{N}$ is meaningless. Consequently, let us define the function $\bar{f}(.)$ exactly as $f(.)$ where the set of definition is $\mathbb{R}^P$ instead of $\mathbb{N}^P$. Consequently,

$$\bar{f}(\bar{\boldsymbol{N}})|_{\bar{N}=N} = f(\boldsymbol{N}), \quad \text{where} \quad \bar{\boldsymbol{N}} = [\bar{N}_1 \dots \bar{N}_P]^{\mathrm{T}},$$

in which $\bar{N}_1, \dots, \bar{N}_P$ are real (continuous) variables.

Using the method of Lagrange multipliers, the problem is as follows:

$$\begin{cases} \min_{\bar{N}} \bar{f}(\bar{\boldsymbol{N}}) \\ \sum_{p=1}^{P} \bar{N}_p = \bar{N}_{\text{total}} \end{cases}$$

where $\bar{N}_{\text{total}}$ is a real positive constant value. Thus, the Lagrange function is given by $\Lambda(\bar{\boldsymbol{N}}, \lambda) = \bar{f}(\bar{\boldsymbol{N}}) + \lambda \left(\sum_{p=1}^{P} \bar{N}_p - \bar{N}_{\text{total}} \right)$ where λ denotes the Lagrange multiplier. For a sufficient number of sensors, the Lagrange function can be approximated by

$$\Lambda(\bar{\boldsymbol{N}}, \lambda) \approx \sum_{p=1}^{P} \frac{\bar{N}_p(1-3V)+3V+1}{\bar{N}_p^3} - \sum_{\substack{p,p'=1 \\ p \neq p'}}^{P} \frac{3 g_p g_{p'} V}{\bar{N}_p \bar{N}_{p'}} + \lambda \left(\sum_{p=1}^{P} \bar{N}_p - \bar{N}_{\text{total}} \right)$$

where $V = \dfrac{1}{1+3P}$. A simple derivation leads to,

$$\frac{\partial \Lambda(\bar{\boldsymbol{N}}, \lambda)}{\partial \bar{N}_1} = \frac{3(V-1)}{\bar{N}_1^3} - \frac{3V+1}{\bar{N}_1^4} + \frac{3V}{\bar{N}_1^2} \sum_{\substack{p,p'=1 \\ p \neq p'}}^{P} \frac{g_p g_{p'}}{\bar{N}_{p'}} + \lambda = 0$$

$$\vdots$$

$$\frac{\partial \Lambda(\bar{\boldsymbol{N}}, \lambda)}{\partial \bar{N}_P} = \frac{3(V-1)}{\bar{N}_P^3} - \frac{3V+1}{\bar{N}_P^4} + \frac{3V}{\bar{N}_P^2} \sum_{\substack{p,p'=1 \\ p \neq p'}}^{P} \frac{g_p g_{p'}}{\bar{N}_{p'}} + \lambda = 0$$

$$\frac{\partial \Lambda(\bar{\boldsymbol{N}}, \lambda)}{\partial \lambda} = \sum_{p=1}^{P} \bar{N}_p - \bar{N}_{\text{total}} = 0.$$

This system of equations seems hard to solve. However, an obvious solution is given by $\bar{N}_1 = \dots = \bar{N}_P = \bar{N}$ and $\lambda = \dfrac{3V+1}{\bar{N}^4} - 3\dfrac{V(P\nu - 1) + V - 1}{\bar{N}^3}$ in which $\nu = \sum_{\substack{p,p'=1 \\ p \neq p'}}^{P} g_p g_{p'}$. Since, $\sum_{p=1}^{P} N_p = \bar{N}_{\text{total}}$, thus the trivial solution is given by $\bar{N}_1 = \dots = \bar{N}_P = \dfrac{\bar{N}_{\text{total}}}{P}$. Consequently, if $\bar{N}_{\text{total}}$ is a multiple of P then, the solution of minimizing the function $\bar{f}(\bar{\boldsymbol{N}})$ in $\mathbb{R}^P$ coincides the solution of minimizing the function $f(\boldsymbol{N})$ in $\mathbb{N}^P$. Thus, the optimal placement minimizing the MSRL is $N_1 = \dots = N_P = \dfrac{\bar{N}_{\text{total}}}{P}$. This conclude the proof.

Endnotes

[a]The notion of distance and closely spaced signals used in the following, is w.r.t. to the metric space (d, C), where $d : C \times C \to \mathbb{R}$ in which d and C denote a metric and the set of the parameters of interest, respectively. [b]See [2-9] for some practical examples for the multidimensional harmonic retrieval model. [c]This study can be straightforwardly extended to other norms. The choice of the l_1 is motivated by its calculation convenience (see the derivation of Result 1 and Appendix 1). Furthermore, since the MSRL is considered to be small (this assumption can be argued by the fact that the high-resolution algorithms have asymptotically an infinite resolving power [44]), thus all continuous p-norms are similar to (i.e., *looks like*) the l_1 norm. More importantly, in a finite dimensional vector space, all continuous p-norms are equivalent [[45], p. 53], thus the choice of a specific norm is free. [d]Note that, due to the specific definition of the SRL in [13] (i.e., using the same notation as in [13], $\delta = \cos(\boldsymbol{u}_1^{\mathrm{T}} \boldsymbol{u}_2)$)and the restrictive assumption in [13] ($\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ belong to the same plan), the SRL as defined in [13] cannot be used in the multidimensional harmonic context. [e]One of the necessary conditions regardless the noise pdf is that $\omega_1^{(p)} \neq \omega_2^{(p)}$. Meaning that each parameter of interest w.r.t. to the first signal $\omega_1^{(p)}$ can be as close as possible to the parameter of interest w.r.t. to the second signal $\omega_2^{(p)}$, but not equal. This is not really a restrictive assumptions, since in most applications, having two or more identical parameters of interest is *a zero probability event* [[9], p. 53]. [f]Note that applying (15) for $P = 1$ and for $\kappa(P_{\text{fa}}, P_{\text{d}}) = 1$, one obtains the Smith criterion [11]. [g]Where $O(.)$ denotes the Landau notation [46]. [h]One should note, that we assumed a uniform linear multi-array, and the problem is to find the optimal distribution of the number of sensors on each array. The more general case, i.e., where the optimization problem considers the non linearity of the multi-way array, is beyond the scope of the problem addressed herein.

Abbreviations

CRB: Cramér-Rao Bound; DOAs: direction of arrivals; FIM: Fisher information matrix; GLRT: generalized likelihood ratio test; MLE: maximum likelihood estimates; MSRL: multidimensional SRL; PARAFAC: PARallel FACtor; pdf: probability density function; SNR: signal-to-noise ratio; SRL: statistical resolution limit.

Acknowledgements

This project is funded by region Île de France and Digiteo Research Park. This work has been partially presented in communication [41].

Competing interests

The authors declare that they have no competing interests.

Received: 10 November 2010 Accepted: 13 June 2011
Published: 13 June 2011

References

1. T Jiang, N Sidiropoulos, J ten Berge, Almost-sure identifiability of multidimensional harmonic retrieval. IEEE Trans. Signal Processing. **49**(9), 1849–1859 (2001). doi:10.1109/78.942615

Annexe G

A Cramér-Rao bounds based analysis of 3D antenna array geometries made from ULA branches

A paraître dans Multidimensional Systems and Signal Processing, Springer

Multidimensional Systems and Signal Processing manuscript No.
(will be inserted by the editor)

A Cramér Rao bounds based analysis of 3D antenna array geometries made from ULA branches

Dinh Thang VU · Alexandre RENAUX · Rémy BOYER · Sylvie MARCOS

Received: date / Accepted: date

Abstract In the context of passive sources localization using antenna array, the estimation accuracy of elevation, and azimuth are related not only to the kind of estimator which is used, but also to the geometry of the considered antenna array. Although there are several available results on the linear array, and also for planar arrays, other geometries existing in the literature, such as 3D arrays, have been less studied. In this paper, we study the impact of the geometry of a family of 3D models of antenna array on the estimation performance of elevation, and azimuth. The Cramér-Rao Bound (CRB), which is widely spread in signal processing to characterize the estimation performance will be used here as a useful tool to find the optimal configuration. In particular, we give closed-form expressions of CRB for a 3D antenna array under both conditional, and unconditional observation models. Thanks to these explicit expressions, the impact of the third dimension to the estimation performance is analyzed. Particularly, we give criterions to design an isotropic 3D array depending on the considered observation model. Several 3D particular geometry antennas made from uniform linear array (ULA) are analyzed, and compared with 2D antenna arrays. The isotropy condition of such arrays is analyzed. The presented framework can be used for further studies of other types of arrays.

Keywords Array geometry optimization · direction of arrival estimation · performance bound.

1 Introduction

In the context of passive sources localization by an array of sensors, the Direction-Of-Arrival (DOA) estimation performance is not only linked to the kind of estimator used but also to the array geometry, *i.e.*, the sensors location in the space.

Dinh Thang VU · Alexandre RENAUX · Rémy BOYER · Sylvie MARCOS
University Paris-Sud 11, Laboratory of Signals and Systems, Supélec, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette cedex, France
Tel.: +331 6985 1763
Fax: +331 6985 1765
E-mail: {Vu,Renaux,Remy.Boyer,Marcos}@lss.supelec.fr

For an array of sensors, the meaning of "performance" can be seen from different points of view: beampattern properties, ambiguities of the array, isotropy, localization estimation in terms of mean square error (MSE), etc. A huge amount of works is available in the literature concerning the study of array ambiguities (see, *e.g.*, [1–5]) the beampattern (see, *e.g.*, [6, 7]), and the isotropic properties of arrays (see, *e.g.*, [8]).

In this paper, we are interested in the optimal array geometry leading to the best performance in terms of MSE. More particularly, we will focus on three dimensional (3D) array geometries less studied in the literature. Indeed, although there are already many available results on planar arrays (2D), there exists other geometries such as 3D arrays. There are many applications where the sensors are scattered in space leading to an arbitrary shape of the antenna (network of telescopes on the Earth's surface, networks of electrodes on the skull of a patient, networks of sensors in a room or in a small space for robotics functions, networks of buoys on the surface of the sea, etc). Moreover, compared to the 2D antenna, the 3D antenna have some intuitive advantages, such as the 3D antenna overcomes the ambiguity of the 2D antenna in some unambiguous cases. For example, one can imagine that in the radar application problem, the targets are located the 3D space and which would be hidden by certain types of landscape (hills, forests, etc.). Therefore, the targets would be "invisible" for a simple planar antenna. However, the 3D antenna could provide a better detection in this situation. The limited number of results in 3D geometry antenna is perhaps due to its complexity leading to more complex expressions.

The analysis already provided in literature deal with two kinds of geometries: geometries based on circular arrays [6] or spherical arrays [9], and geometries based on linear branches (such as the well known Uniform Linear Array (ULA), the V-shaped arrays, the cross arrays or rectangular arrays). More particular attention has been paid on uniform arrays. This paper follows the context of arrays made with ULA branches.

In order to study the performance in terms of MSE, the most popular tool is clearly the Cramér-Rao bound (CRB) [10], probably because it can generally be achieved by the variance of localization estimators for a high number of snapshots [11] or at high Signal-to-Noise Ratio (SNR) [12, 13]. The CRB has already been widely used in the literature to describe the fundamental properties of arrays. Through a simple form of the CRB expression, [14] shows the impact of the sensors location on the DOA estimation accuracy in the case of 2D arrays. Concerning DOA estimation, in [15], [8], [16], some conditions on sensors positions to ensure the isotropy are studied for 2D, and 3D arrays, by way of the off-diagonal entries of the CRB, where, the arrays have the same estimation accuracy over the whole field of view. In [17], [18], the CRB for the source position estimation based on the time difference of arrival method (TDOA) is used to prove that the best geometry which minimize the trace of the CRB matrix is the uniform angular array (UAA). A Bayesian CRB approach for the case where the source is coplanar with the antenna and the DOA is modeled as a random variable is introduced in [19]. In [20], a deep study of the CRB for 2D antenna and a source anywhere in the space has been provided, leading to interesting results concerning the so-called V-shaped array in terms of isotropy, and MSE performance. Then, based on the work of [20], a novel planar geometry called the optimum ambiguity-free planar antenna array with a closed-form of V-shaped array has been introduced in [21]. Finally, in a

recent work [22], the authors showed that the both conditional, and unconditional CRB, jointly with the variance of DOA obtained from MUSIC algorithm, can be expressed in the same term depending on the sensors location, and this kind of CRB expressions can be used as a tool in order to optimize the array.

Note that, in array processing, the source signal is generally modeled as a Gaussian random process or as a deterministic sequence. These models are referred to as the unconditional model, and the conditional model respectively [23]. Particularly, under conditional model, the incoming signal waveforms can be assumed as either known or unknown parameters. Consequently, the computational cost of the estimation problem varies *w.r.t* the signal waveforms assumption. The unknown signal waveforms always leads to the increase of the parameter dimension compared to the unconditional model. On the contrary, there exist in literature several applications where the signal amplitudes are known such as mobile telecommunication. The knowledge of the signal can improve the estimation performance, also reduce the problem complexity. We can cite here several works concerning the context of known waveform signal (see, *e.g.*, [24–28]).

Of course, since the observation model can change, there are two different CRB associated to each model called unconditional CRB (UCRB), and conditional CRB (CCRB). It has been proved that the UCRB can be achieved for a high number of snapshots [11], however, it is not achievable at high SNR (for a fixed number of snapshot) [13]. On the other hand, the CCRB is achieved at high SNR [12] but it is not achieved for a large number of snapshots [11]. Surprisingly, to the best of our knowledge, all the previously proposed results are conducted in the framework of the unconditional observation model, and consequently, in the framework of the UCRB. We will show in this paper that in the framework of the conditional model, some results concerning the array geometry differ significantly from the unconditional observation model.

In this paper, both conditional, and unconditional observation models are considered to study 3D geometries. First, we detail the Fisher Information Matrix (FIM) expressions concerning the azimuth, and elevation in the case of a general 3D array. The Fisher information represents the way to measure the information about the parameter contained in the observations via its likelihood function. Secondly, closed form expressions of the CRB are provided when one adds an orthogonal branch to a planar array with any geometry. This model is the first step to analyze the contribution of the third dimension where an intuitive advantage of 3D antenna arrays *w.r.t* 2D antenna arrays is the overcoming of the ambiguity problem in elevation estimation. Third, to analyze the impact of the array geometry on estimation, we propose several closed-form expressions of the CRB for classical array shapes made with the well known ULA branches. Note that these kinds of geometry (namely the L-shaped, and V-shaped arrays) have already been investigated in the 2D case, which are seen as particular cases of our proposed expressions. In [29], the L-shaped antenna arrays has been proved to have 37% better accuracy than the cross array. In [30], the author introduce the isotropic conditions for the sensors positions, and for the opening angle between the two branches of the uniform/ nonuniform V-shaped planar antenna under unconditional assumption. Our goal is to extend these geometries in the 3D case to analyze the impact of a 3D additional branch in terms of MSE. These results are then analyzed to describe the performance of these arrays in terms of MSE, isotropy, and the decoupling properties. Finally, the comparison between the 3D,

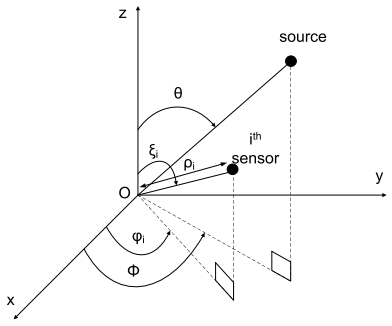


Fig. 1 Coordinate system for the source, and the sensors

and 2D antenna arrays, and also the comparison between 3D, and uniform circular antenna arrays (UCA), which have the same number of sensors, are analyzed to illustrate the impact of the third dimension. However, with a constant number of sensors, in order to add the third dimension to the antenna, the aperture of the antenna must be reduced. Therefore, the estimation accuracy will be affected.

The notational convention adopted in this paper is listed as follows: italic indicates a scalar quantity, bold lower case indicates a vector, bold upper case indicates a matrix. $\widehat{\Theta}$ indicates the estimated value of Θ. $\mathbf{A}^T$ is the transpose of $\mathbf{A}$. $\mathbf{A}^*$ is the conjugate of $\mathbf{A}$. $\mathbf{A}^H$ is the transpose, and conjugate of $\mathbf{A}$. $\Re\{.\}$ denotes the real part. $\Im\{.\}$ denotes the imaginary part. $|.|$ denotes the absolute value of complex scalar. $tr(\mathbf{A})$ denotes the trace of matrix $\mathbf{A}$. $[\mathbf{A}]_{i,j}$ is the $\{i,j\}^{th}$ element of matrix $\mathbf{A}$. $\det(\mathbf{A})$ is the determinant of the square matrix $\mathbf{A}$. $\mathbf{I}_K$ denotes the $K \times K$ identity matrix. And a_i denotes the i^{th} element of vector $\mathbf{a}$.

The paper is organized as follows. In Sec. II, we introduce the model, then we resume briefly the notion of CRB depending on the two hypotheses about the source signal. In Sec. III, the expressions of general CRB under the two assumptions are listed. In Sec. IV, we develop the explicit expression CRB for a family of 3D antenna arrays. Next, in Sec. V, we consider the applications of the CRB to some particular 3D antenna models to find isotropic, and uncoupling conditions. In Sec. VI, we give some typical simulations to illustrate the performance comparison between the different antenna arrays. Finally, in Sec. VII, a conclusion is presented .

2 Model setup

In this paper, we are interested in using an unbiased estimator to localize a single source emitting narrow-band signal in the far-field area by using a three dimensional array containing M identical, and omnidirectional sensors. The source position is characterized by its spherical coordinates, such as the bearing angle vector $\mathbf{\Theta} = \left[\phi\ \theta\right]^T$ where ϕ is the azimuth, and θ the elevation of the source. The i^{th} sensor position is characterized by the triple parameters $(\rho_i,\ \xi_i,\ \varphi_i)$ (see Fig. 1). θ, and ξ_i are measured clockwise from the z axis, while ϕ, and φ_i are measured

counter-clockwise from the x axis. Letting $s(t)$, $y(t)$, $n(t)$ denote the source signal, the output signal at the array of sensors, and the additive noise respectively, for $t = 1, \ldots, T$, where T is the number of snapshots. At the t^{th} observation, the output signal at the array of sensors is then given by:

$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ \vdots \\ y_M(t) \end{bmatrix} = \begin{bmatrix} e^{\left(\frac{j2\pi}{\lambda}\mathbf{v}_1^T\mathbf{r}(\Theta)\right)} \\ \vdots \\ e^{\left(\frac{j2\pi}{\lambda}\mathbf{v}_M^T\mathbf{r}(\Theta)\right)} \end{bmatrix} s(t) + \mathbf{n}(t) = \mathbf{a}(\boldsymbol{\Theta})s(t) + \mathbf{n}(t), \quad (1)$$

where λ denotes the wavelength. The vector $\mathbf{a}(\boldsymbol{\Theta})$ is the $M \times 1$ steering vector with its i^{th} element given by $[\mathbf{a}(\boldsymbol{\Theta})]_i = \exp\left(\frac{j2\pi}{\lambda}\mathbf{v}_i^T\mathbf{r}(\boldsymbol{\Theta})\right)$, where $\mathbf{r}(\boldsymbol{\Theta}) = \begin{bmatrix}\sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta\end{bmatrix}^T$ is the unit vector pointing toward the source, and $\mathbf{v}_i = \begin{bmatrix}\rho_i\sin\xi_i\cos\varphi_i & \rho_i\sin\xi_i\sin\varphi_i & \rho_i\cos\xi_i\end{bmatrix}^T$ is the position of the i^{th} sensor. In the spherical coordinate system, the i^{th} element of the steering vector is given by:

$$[\mathbf{a}(\boldsymbol{\Theta})]_i = e^{\left(\frac{2j\pi\rho_i}{\lambda}(\sin\theta\sin\xi_i\cos(\phi-\varphi_i)+\cos\xi_i\cos\theta)\right)}. \quad (2)$$

The noise vector $\mathbf{n}(t) \in \mathbb{C}^M$ is assumed to be Gaussian, circular, independent, and identically distributed (*i.i.d.*), zero mean with covariance matrix $\sigma_n^2\mathbf{I}_M$.

Concerning the source, the two following alternative hypotheses can be assumed:

- H_1: $s(t)$ is complex, deterministic, and assumed to be known at the receiver.
- H_2: $s(t)$ is assumed circular, Gaussian, zero-mean with variance σ_s^2 known at the receiver ($s(t) \sim \mathbb{C}\mathcal{N}(0, \sigma_s^2)$), *i.i.d*, and independent of the noise.

Depending on the assumption H_1 or H_2 which is used, both mean or covariance matrix of the output signal may depend on $\boldsymbol{\Theta}$. To be more general, let us first assume that $\mathbf{y}|\boldsymbol{\Theta} \sim \mathbb{C}\mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\Theta}), \mathbf{R}(\boldsymbol{\Theta}))$, where $\boldsymbol{\mu}(\boldsymbol{\Theta})$ is the $M \times 1$ mean vector, and $\mathbf{R}(\boldsymbol{\Theta})$ is the $M \times M$ covariance matrix. From the Schwarz inequality, the variance of any unbiased estimator $\widehat{\boldsymbol{\Theta}}$ will satisfy: $\mathbf{var}(\widehat{\boldsymbol{\Theta}}_i) \geq \left[\mathbf{FIM}^{-1}(\Theta)\right]_{ii}$ which is known as the CRB, where $\mathbf{FIM}(\Theta)$ is the $M \times M$ Fisher Information Matrix (FIM). For *i.i.d* observations, the FIM is given by [6]:

$$[\mathbf{FIM}(\boldsymbol{\Theta})]_{i,j} = -E\left\{\frac{\partial^2 \ln p(\mathbf{Y}|\boldsymbol{\Theta})}{\partial\Theta_i\partial\Theta_j}\right\} = -\sum_{t=1}^{T} E\left\{\frac{\partial^2 \ln p(\mathbf{y}(t)|\boldsymbol{\Theta})}{\partial\Theta_i\partial\Theta_j}\right\}, \quad (3)$$

where $\mathbf{Y} = [\mathbf{y}(1) \quad \ldots \quad \mathbf{y}(T)]$. The likelihood function is given by: $p(\mathbf{Y}|\boldsymbol{\Theta}) = \prod_{t=1}^{T} p(\mathbf{y}(t)|\boldsymbol{\Theta})$, where

$$p(\mathbf{y}(\mathbf{t})|\boldsymbol{\Theta}) = \frac{1}{\pi^M \det[\mathbf{R}(\boldsymbol{\Theta})]}\mathbf{exp}\left(-(\mathbf{y}(t) - \boldsymbol{\mu}(\boldsymbol{\Theta}))^H\mathbf{R}^{-1}(\boldsymbol{\Theta})(\mathbf{y}(t) - \boldsymbol{\mu}(\boldsymbol{\Theta}))\right).$$

A general expression of the FIM for circular Gaussian complex observations can be deduced from [31], and [32]:

$$[\mathbf{FIM}(\boldsymbol{\Theta})]_{i,j} = tr\left(\mathbf{R}^{-1}(\Theta)\frac{\partial\mathbf{R}(\boldsymbol{\Theta})}{\partial\Theta_i}\mathbf{R}^{-1}(\boldsymbol{\Theta})\frac{\partial\mathbf{R}(\boldsymbol{\Theta})}{\partial\Theta_j}\right) + 2\Re\left(\left[\frac{\partial\boldsymbol{\mu}(\boldsymbol{\Theta})}{\partial\Theta_i}\right]^H\mathbf{R}^{-1}(\boldsymbol{\Theta})\frac{\partial\boldsymbol{\mu}(\boldsymbol{\Theta})}{\partial\Theta_j}\right). \quad (4)$$

Consequently, under H_1, since $\mathbf{y}(t)|\boldsymbol{\Theta} \sim \mathbb{C}\mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\Theta}), \sigma^2\mathbf{I})$ then (4) is reduced to

$$[\mathbf{FIM}(\boldsymbol{\Theta})]_{i,j} = 2\Re\left(\left[\frac{\partial\boldsymbol{\mu}(\boldsymbol{\Theta})}{\partial\Theta_i}\right]^H \mathbf{R}^{-1}(\boldsymbol{\Theta})\frac{\partial\boldsymbol{\mu}(\boldsymbol{\Theta})}{\partial\Theta_j}\right).$$

And under H_2, since $\mathbf{y}(t)|\boldsymbol{\Theta} \sim \mathbb{C}\mathcal{N}(\mathbf{0}, \mathbf{R}(\boldsymbol{\Theta}))$, then (4) is reduced to

$$[\mathbf{FIM}(\boldsymbol{\Theta})]_{i,j} = tr\left(\mathbf{R}^{-1}(\Theta)\frac{\partial\mathbf{R}(\boldsymbol{\Theta})}{\partial\Theta_i}\mathbf{R}^{-1}(\boldsymbol{\Theta})\frac{\partial\mathbf{R}(\boldsymbol{\Theta})}{\partial\Theta_j}\right).$$

The parameters of interest are the azimuth, and elevation angles, *i.e.*, the vector $\boldsymbol{\Theta}$ which are assumed deterministic. Therefore, the CRB, denoted $\mathbf{C}(\boldsymbol{\Theta})$, is a 2×2 matrix which can be defined as:

$$\mathbf{C}(\boldsymbol{\Theta}) = \mathbf{FIM}(\boldsymbol{\Theta})^{-1} = \begin{bmatrix} C_{\theta\theta}(\boldsymbol{\Theta}) & C_{\theta\phi}(\boldsymbol{\Theta}) \\ C_{\phi\theta}(\boldsymbol{\Theta}) & C_{\phi\phi}(\boldsymbol{\Theta}) \end{bmatrix}, \tag{5}$$

where, $C_{\theta\theta}$, and $C_{\phi\phi}$ represent the CRBs of elevation, and azimuth, respectively. $C_{\theta\phi} = C_{\phi\theta}$ represents the coupling between parameters θ, and ϕ.

3 FIM expressions for a general 3D array

In this section, we will detail the CRB expressions under both the conditional, and unconditional observation models.

3.1 Conditional observation model (H_1 assumption)

Under H_1, since the parameters only appear in the mean $\boldsymbol{\mu}(\boldsymbol{\Theta})$, *i.e.*, $\mathbf{R}(\boldsymbol{\Theta})$ is not a function of $\boldsymbol{\Theta}$ in this case, from Eqn. (4), the FIM can be simplified as follows:

$$[\mathbf{FIM}(\boldsymbol{\Theta})]_{i,j} = 2\Re\left(\frac{\partial\boldsymbol{\mu}(\boldsymbol{\Theta})^H}{\partial\Theta_i}\mathbf{R}^{-1}(\boldsymbol{\Theta})\frac{\partial\boldsymbol{\mu}(\boldsymbol{\Theta})}{\partial\Theta_j}\right). \tag{6}$$

In this case, the mean vector is given by: $\boldsymbol{\mu}(\boldsymbol{\Theta}) = (\mathbf{I}_T \otimes \mathbf{a}(\boldsymbol{\Theta}))\mathbf{s}$, where $\otimes$ denotes the Kronecker product, $\mathbf{s}$ denotes the source signal vector $\boldsymbol{s} = [s(1)\dots s(T)]^T$. The covariance matrix is given by: $\mathbf{R} = \sigma_n^2\mathbf{I}_{MT}$. Therefore, (6) becomes:

$$[\mathbf{FIM}(\boldsymbol{\Theta})]_{i,j} = \frac{2||\boldsymbol{s}||^2}{\sigma_n^2}\Re\left(\frac{\partial\boldsymbol{a}(\boldsymbol{\Theta})}{\partial\Theta_i}^H\frac{\partial\boldsymbol{a}(\boldsymbol{\Theta})}{\partial\Theta_j}\right), \tag{7}$$

where $i,j \in \{1,2\}^2$, and $\Theta_1 = \theta$, and $\Theta_2 = \phi$, and where $||\boldsymbol{s}||^2 = \boldsymbol{s}^H\boldsymbol{s}$. The derivation of the steering vector *w.r.t.* θ, and ϕ is

$$\begin{aligned}\frac{\partial[\mathbf{a}(\boldsymbol{\Theta})]_i}{\partial\theta} = \frac{2j\pi\rho_i}{\lambda}&\left(\cos\theta\sin\xi_i\cos\left(\phi-\varphi_i\right)-\cos\xi_i\sin\theta\right)\\ &\times e^{\left(\frac{2j\pi\rho_i}{\lambda}\left(\sin\theta\sin\xi_i\cos\left(\phi-\varphi_i\right)+\cos\xi_i\cos\theta\right)\right)},\end{aligned} \tag{8}$$

and

$$\frac{\partial[\mathbf{a}(\mathbf{\Theta})]_i}{\partial\phi} = \frac{2j\pi\rho_i}{\lambda}\left(-\sin\theta\sin\xi_i\sin\left(\phi-\varphi_i\right)\right)e^{\left(\frac{2j\pi\rho_i}{\lambda}\left(\sin\theta\sin\xi_i\cos\left(\phi-\varphi_i\right)+\cos\xi_i\cos\theta\right)\right)}. \tag{9}$$

Then, (7) becomes

$$\begin{cases} \frac{[\mathbf{FIM}]_{1,1}}{C_{SNR}} = \sum_{i=1}^{M}\rho_i^2\left(\cos\theta\sin\xi_i\cos\left(\phi-\varphi_i\right)-\cos\xi_i\sin\theta\right)^2, \\ \frac{[\mathbf{FIM}]_{2,2}}{C_{SNR}} = \sum_{i=1}^{M}\rho_i^2\left(\sin\theta\sin\xi_i\sin\left(\phi-\varphi_i\right)\right)^2, \\ \frac{[\mathbf{FIM}]_{1,2}}{C_{SNR}} = -\sum_{i=1}^{M}\rho_i^2\left(\sin\theta\sin\xi_i\sin\left(\phi-\varphi_i\right)\right)\left(\cos\theta\sin\xi_i\cos\left(\phi-\varphi_i\right)-\cos\xi_i\sin\theta\right), \end{cases} \tag{10}$$

where $C_{SNR} = \frac{8\pi^2||\boldsymbol{s}||^2}{\sigma_n^2\lambda^2}$. And the determinant of the **FIM** is given by

$$\det\left(\mathbf{FIM}(\mathbf{\Theta})\right) = [\mathbf{FIM}]_{1,1}[\mathbf{FIM}]_{2,2} - [\mathbf{FIM}]_{1,2}[\mathbf{FIM}]_{2,1}. \tag{11}$$

3.2 Unconditional observation model (H_2 assumption)

Under H_2 assumption, since the parameters only appear in the covariance $\mathbf{R}(\mathbf{\Theta})$, from Eqn. (4), the FIM becomes:

$$[\mathbf{FIM}(\mathbf{\Theta})]_{i,j} = tr\left(\mathbf{R^{-1}}(\mathbf{\Theta})\frac{\partial\mathbf{R}(\mathbf{\Theta})}{\partial\Theta_i}\mathbf{R^{-1}}(\mathbf{\Theta})\frac{\partial\mathbf{R}(\mathbf{\Theta})}{\partial\Theta_j}\right). \tag{12}$$

Because $\mathbf{R}(\mathbf{\Theta}) = \sigma_s^2\mathbf{I}_T\otimes(\mathbf{a}(\mathbf{\Theta})\mathbf{a^H}(\mathbf{\Theta})) + \sigma_n^2\mathbf{I}_{MT}$, and from [33, eq. (39)], (12) can be written as follows:

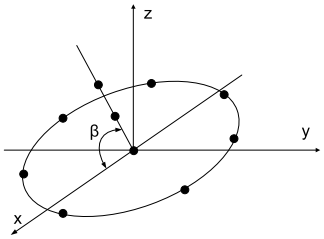
$$[\mathbf{FIM}(\mathbf{\Theta})]_{i,j} = \frac{2TM\sigma_s^4}{\sigma_n^2(\sigma_n^2+M\sigma_s^2)}\left(\frac{\partial\boldsymbol{a}(\boldsymbol{\Theta})}{\partial\Theta_i}^H\frac{\partial\boldsymbol{a}(\boldsymbol{\Theta})}{\partial\Theta_j} - \frac{1}{M}\frac{\partial\boldsymbol{a}(\boldsymbol{\Theta})}{\partial\Theta_i}^H\boldsymbol{a}(\boldsymbol{\Theta})\boldsymbol{a}(\boldsymbol{\Theta})^H\frac{\partial\boldsymbol{a}(\boldsymbol{\Theta})}{\partial\Theta_j}\right), \tag{13}$$

where $i,j = \{1,2\}^2$. $\frac{\partial\mathbf{a}(\mathbf{\Theta})}{\partial\Theta_1} = \frac{\partial\mathbf{a}(\mathbf{\Theta})}{\partial\theta}$ is given by Eqn. (8), and $\frac{\partial\mathbf{a}(\mathbf{\Theta})}{\partial\Theta_2} = \frac{\partial\mathbf{a}(\mathbf{\Theta})}{\partial\phi}$ is given by Eqn. (9). Then, (13) leads to (14), shown at the bottom of the page, where $U_{SNR} = \frac{8\pi^2TM\sigma_s^4}{\sigma_n^2\lambda^2(\sigma_n^2+M\sigma_s^2)}$.

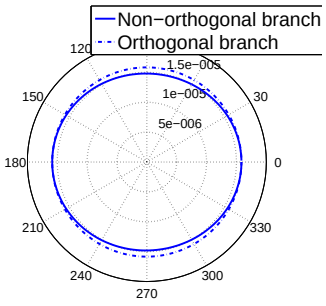
$$\begin{aligned} \frac{[\mathbf{FIM}]_{1,1}}{U_{SNR}} &= \sum_{i=1}^{M}\rho_i^2\left(\cos\theta\sin\xi_i\cos\left(\phi-\varphi_i\right)-\cos\xi_i\sin\theta\right)^2 \\ &\quad -\tfrac{1}{M}\left(\sum_{i=1}^{M}\rho_i\left(\cos\theta\sin\xi_i\cos\left(\phi-\varphi_i\right)-\cos\xi_i\sin\theta\right)\right)^2, \\ \frac{[\mathbf{FIM}]_{2,2}}{U_{SNR}} &= \sum_{i=1}^{M}\rho_i^2\left(\sin\theta\sin\xi_i\sin\left(\phi-\varphi_i\right)\right)^2 - \tfrac{1}{M}\left(\sum_{i=1}^{M}\rho_i\sin\theta\sin\xi_i\sin\left(\phi-\varphi_i\right)\right)^2, \\ \frac{[\mathbf{FIM}]_{1,2}}{U_{SNR}} &= -\sum_{i=1}^{M}\rho_i^2\left(\cos\theta\sin\xi_i\cos\left(\phi-\varphi_i\right)-\cos\xi_i\sin\theta\right)\left(\sin\theta\sin\xi_i\sin\left(\phi-\varphi_i\right)\right) \\ &\quad +\tfrac{1}{M}\sum_{i=1}^{M}\rho_i\left(\cos\theta\sin\xi_i\cos\left(\phi-\varphi_i\right)-\cos\xi_i\sin\theta\right)\sum_{i=1}^{M}\rho_i\sin\theta\sin\xi_i\sin\left(\phi-\varphi_i\right), \end{aligned}$$

And the determinant of the **FIM** is given by (11). In order to simplify the analysis of the general CRBs expressions, we will consider in the following section the CRBs expressions in case of several 3D geometries arrays based on ULA branches.

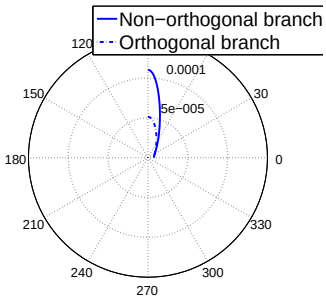
4 Planar array + ULA orthogonal branch



(a) Orthogonal branch antenna, and non orthogonal branch antenna



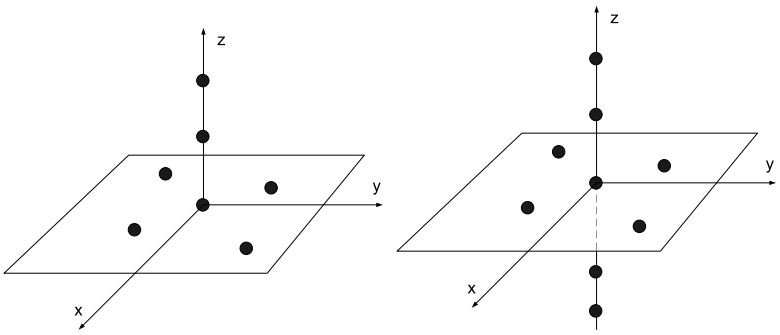
(b) $C_{\phi\phi}$ $w.r.t$ ϕ, with $\theta = 45°$



(c) $C_{\theta\theta}$ $w.r.t$ θ, with $\phi = 90°$

Fig. 2 Orthogonal branch versus non-orthogonal branch antenna.

In considering the combination of planar antenna, and the linear antenna branch in order to make an 3D antenna, there are two cases: either orthogonal branch or non-orthogonal branch. In order to analyze the impact of the array branch position to the estimation performance of the antenna, let us consider a numerical simulation about the DOA estimation performance of the antenna made from an uniform circular antenna with 7 sensors, and an uniform linear branch with 2 sensors (see Fig. 2.(a)). Let β denotes the angle between the branch and the circular antenna plane. The inter-sensors spacing is a half-wavelength. We then compare the estimation performance between the antenna with $\beta = 90°$, *i.e.*, orthogonal branch, and the antenna with $\beta = 45°$. Fig. 2.(b) shows the polar representation of the CRB of azimuth w.r.t azimuth angle with the elevation



(a) Planar array with an ULA orthogonal branch

(b) Planar array with two symetric ULA orthogonal branches

Fig. 3 Planar extension array

angle, $\theta = 45^\circ$. Fig. 2.(c) represents the polar representation of the CRB of elevation w.r.t elevation angle with the azimuth angle $\phi = 90^\circ$. The smaller the CRB is, the better estimation performance we have. One can observe that the antenna with orthogonal branch is better in elevation estimation accuracy than the non-orthogonal branch antenna. By several others simulations, and analyses, we observe that the compromise about the DOA estimation performance between the orthogonal and non-orthogonal branch antenna depend to the source position, i.e., the DOA. Consequently, we can not find the optimal branch position for the whole field of source position. However, let us remind that the estimation performance is strongly dependent on the aperture of the antenna. The larger the aperture antenna is, the better will be the estimation accuracy. Therefore, in order to improve the contribution of the 3^{rd} dimension (Oz direction), we have to find the sensor positioning maximizing the aperture of the antenna in the 3^{rd} dimension for the same aperture of the branch. It is clear that the orthogonal branch antenna is the solution. Therefore, in this paper, we consider only the orthogonal branch antenna case.

In this section, we consider an extension of an arbitrary planar array consisting of N_1 sensors when an (or two opposite) ULA orthogonal branch(es) are added. The number of sensors located on the orthogonal branch(es) is denoted by N_2. Therefore, the total number of sensors is given by $M = N_1 + N_2$. Without loss of generality, let us assume that the z axis is a ULA branch, while the xOy plane coincides with the planar array. In order to analyze the impact of the third dimension to the estimation performance, let us denote $\rho_{k,i}$ the distance of a sensor to the origin where the index $k = 1$ means that the sensor is located on the plane otherwise $k = 2$ means that the sensor is located on the orthogonal branch. For this reason, $\rho_{2,i}$ represents the distance from the origin to a sensor located on the orthogonal branch, and $\rho_{1,i}$ represents the distance from the origin to a sensor

located on the plane xOy. Then, let us set:

$$\begin{aligned}
S_{12} &= \sum_{i=1}^{N_1} \rho_{1,i}^2 e^{2j\varphi_i}, \\
S_{10} &= \sum_{i=1}^{N_1} \rho_{1,i}^2, \\
S_{11} &= \sum_{i=1}^{N_1} \rho_{1,i} e^{j\varphi_i}, \\
S_{13} &= \sum_{i=1}^{N_1} \rho_{1,i}, \\
S_{20} &= \sum_{i=N_1+1}^{M} \rho_{2,i}^2, \\
S_{23} &= \sum_{i=N_1+1}^{M} \rho_{2,i}.
\end{aligned} \tag{14}$$

Note that the parameters $S_{k,i}$, with $k \in \{1,2\}, i \in \{0,\ldots,3\}$ depends only on the array geometry.

4.1 Planar array with an ULA orthogonal branch

The antenna geometry is illustrated in Fig 3(a).

– *Conditional observation model*

Under H_1, the CRB has the following compact expression (see A for the proof):

$$\begin{aligned}
C_{\theta\theta} &= \frac{2}{C_{SNR}} \frac{\left(S_{10} - \Re\{S_{12}e^{-2j\phi}\}\right)}{\left(\cos^2\theta(S_{10}^2 - |S_{12}|^2) + \sin^2\theta S_{20}\left(2S_{10} - 2\Re\{S_{12}e^{-2j\phi}\}\right)\right)}, \\
C_{\phi\phi} &= \frac{4}{C_{SNR}\sin^2\theta} \frac{\left[\frac{1}{2}\cos^2\theta\left(\Re\{S_{12}e^{-2j\phi}\} + S_{10}\right) + \sin^2\theta S_{20}\right]}{\left(\cos^2\theta(S_{10}^2 - |S_{12}|^2) + \sin^2\theta S_{20}\left(2S_{10} - 2\Re\{S_{12}e^{-2j\phi}\}\right)\right)}, \\
C_{\theta\phi} &= \frac{-\cos\theta}{C_{SNR}\sin\theta} \frac{\Im\{S_{12}e^{-2j\phi}\}}{\left(\cos^2\theta(S_{10}^2 - |S_{12}|^2) + \sin^2\theta S_{20}\left(2S_{10} - 2\Re\{S_{12}e^{-2j\phi}\}\right)\right)}
\end{aligned} \tag{15}$$

– *Unconditional observation model*

The elements of the CRB is given by : $C_{ij} = \frac{Num_{ij}}{Den}$ where $i,j = \{\phi,\theta\}^2$. The denominator of CRB is given by:

$$\begin{aligned}
\frac{Den}{(U_{SNR})^2\sin^2\theta} = \frac{\cos^2\theta}{4} &\left(\left(S_{10} - \frac{|S_{11}|^2}{M}\right)^2 - \left|S_{12} - \frac{S_{11}^2}{M}\right|^2\right) \\
&+ \frac{\sin 2\theta}{2M} S_{23}\left(S_{10}\Re\{e^{-j\phi}S_{11}\} - \Re\{e^{-j\phi}S_{12}S_{11}^*\}\right) \\
&+ \sin^2\theta\frac{S_{20}}{2}\left(S_{10} - \frac{|S_{11}|^2}{M} - \Re\left\{e^{-2j\phi}\left(S_{12} - \frac{S_{11}^2}{M}\right)\right\}\right) \\
&+ \sin^2\theta\frac{S_{23}^2}{2M}\left(\Re\{e^{-2j\phi}S_{12}\} - S_{10}\right).
\end{aligned} \tag{16}$$

The numerators of the CRB elements are given by:

$$\begin{aligned}
\frac{Num_{\phi\phi}}{U_{SNR}} &= \frac{\cos^2\theta}{2}\left(S_{10}-\frac{|S_{11}|^2}{M}+\Re\left\{e^{-2j\phi}\left(S_{12}-\frac{S_{11}^2}{M}\right)\right\}\right)\\
&\quad+\sin^2\theta\left(S_{20}-\frac{S_{23}^2}{M}\right)+\frac{\sin 2\theta S_{23}}{M}\Re\left\{e^{-j\phi}S_{11}\right\},\\
\frac{Num_{\theta\theta}}{U_{SNR}} &= \frac{\sin^2\theta}{2}\left(S_{10}-\frac{|S_{11}|^2}{M}-\Re\left\{e^{-2j\phi}\left(S_{12}-\frac{S_{11}^2}{M}\right)\right\}\right),\\
\frac{Num_{\theta\phi}}{U_{SNR}} &= \frac{\sin 2\theta}{4}\Im\left\{e^{-2j\phi}\left(S_{12}-\frac{S_{11}^2}{M}\right)\right\}+\frac{\sin^2\theta S_{23}}{M}\Im\left\{e^{-j\phi}S_{11}\right\}.
\end{aligned} \tag{17}$$

The proof are shown in B.

4.2 Planar array with two symmetric orthogonal branches

If the antenna structure has two symmetric orthogonal branches in such a way that the orthogonal branches centroid is located on xOy plane (see Fig. 3(b)), then, a simpler CRB expression can be deduced for the unconditional model.

– *Conditional observation model*

In this case, the CRB has the same expressions as the previous antenna model given in Eqn. (15).

– *Unconditional observation model*

Let N_2' be the number of sensors located on the opposite orthogonal branch. Hence, the CRB expression can be deduced from the geometry described on Fig. 3(a) by letting $S_{23}=0$. This leads to:

$$\begin{aligned}
C_{\phi\phi} &= \frac{\frac{\cos^2\theta\left(S_{10}-\frac{|S_{11}|^2}{M}+\Re\left\{e^{-2j\phi}\left(S_{12}-\frac{S_{11}^2}{M}\right)\right\}\right)}{2U_{SNR}\sin^2\theta}+\frac{\sin^2\theta S_{20}}{U_{SNR}\sin^2\theta}}{\left(\begin{array}{l}\frac{\cos^2\theta}{4}\left(\left(S_{10}-\frac{|S_{11}|^2}{M}\right)^2-\left|S_{12}-\frac{S_{11}^2}{M}\right|^2\right)\\+\frac{S_{20}\sin^2\theta}{2}\left(S_{10}-\frac{|S_{11}|^2}{M}-\Re\left\{e^{-2j\phi}\left(S_{12}-\frac{S_{11}^2}{M}\right)\right\}\right)\end{array}\right)},\\
C_{\theta\theta} &= \frac{\frac{1}{2U_{SNR}}\left(S_{10}-\frac{|S_{11}|^2}{M}-\Re\left\{e^{-2j\phi}\left(S_{12}-\frac{S_{11}^2}{M}\right)\right\}\right)}{\left(\begin{array}{l}\frac{\cos^2\theta}{4}\left(\left(S_{10}-\frac{|S_{11}|^2}{M}\right)^2-\left|S_{12}-\frac{S_{11}^2}{M}\right|^2\right)\\+\frac{S_{20}\sin^2\theta}{2}\left(S_{10}-\frac{|S_{11}|^2}{M}-\Re\left\{e^{-2j\phi}\left(S_{12}-\frac{S_{11}^2}{M}\right)\right\}\right)\end{array}\right)},\\
C_{\theta\phi} &= \frac{-\frac{\cos\theta}{2U_{SNR}\sin\theta}\Im\left\{e^{-2j\phi}\left(S_{12}-\frac{S_{11}^2}{M}\right)\right\}}{\left(\begin{array}{l}\frac{\cos^2\theta}{4}\left(\left(S_{10}-\frac{|S_{11}|^2}{M}\right)^2-\left|S_{12}-\frac{S_{11}^2}{M}\right|^2\right)\\+\frac{S_{20}\sin^2\theta}{2}\left(S_{10}-\frac{|S_{11}|^2}{M}-\Re\left\{e^{-2j\phi}\left(S_{12}-\frac{S_{11}^2}{M}\right)\right\}\right)\end{array}\right)}.
\end{aligned} \tag{18}$$

See C for the proof.

4.3 Planar array

Due to the fact that planar array (2D) is a particular case of a 3D array ($N_2=0$), the CRB for an arbitrary planar array are obtained by letting $S_{20}=S_{23}=0$,

which leads to the following equations. Under H_1 assumption:

$$
\begin{aligned}
C_{\theta\theta} &= \frac{2\left(S_{10} - \Re\{S_{12}e^{-2j\phi}\}\right)}{C_{SNR}\cos^2\theta(S_{10}^2 - |S_{12}|^2)},\\
C_{\phi\phi} &= \frac{2\left(\Re\{S_{12}e^{-2j\phi}\} + S_{10}\right)}{C_{SNR}\sin^2\theta(S_{10}^2 - |S_{12}|^2)},\\
C_{\theta\phi} &= -\frac{\Im\{S_{12}e^{-2j\phi}\}}{C_{SNR}\sin\theta\cos\theta(S_{10}^2 - |S_{12}|^2)}.
\end{aligned} \tag{19}
$$

and under H_2 assumption, the CRB leads to the results of [20]:

$$
\begin{aligned}
C_{\phi\phi} &= \frac{S_{10} - \frac{|S_{11}|^2}{M} + \Re\left\{e^{-2j\phi}\left(S_{12} - \frac{S_{11}^2}{M}\right)\right\}}{U_{SNR}\frac{\sin^2\theta}{2}\left(\left(S_{10} - \frac{|S_{11}|^2}{M}\right)^2 - \left|S_{12} - \frac{S_{11}^2}{M}\right|^2\right)},\\
C_{\theta\theta} &= \frac{S_{10} - \frac{|S_{11}|^2}{M} - \Re\left\{e^{-2j\phi}\left(S_{12} - \frac{S_{11}^2}{M}\right)\right\}}{U_{SNR}\frac{\cos^2\theta}{2}\left(\left(S_{10} - \frac{|S_{11}|^2}{M}\right)^2 - \left|S_{12} - \frac{S_{11}^2}{M}\right|^2\right)},\\
C_{\theta\phi} &= -\frac{\Im\left\{e^{-2j\phi}\left(S_{12} - \frac{S_{11}^2}{M}\right)\right\}}{U_{SNR}\frac{\sin 2\theta}{4}\left(\left(S_{10} - \frac{|S_{11}|^2}{M}\right)^2 - \left|S_{12} - \frac{S_{11}^2}{M}\right|^2\right)}.
\end{aligned} \tag{20}
$$

4.4 Analysis

4.4.1 Isotropy, and uncoupling properties

One of several interests from the obtained closed-form expressions of the CRB is to design the array antenna in terms of isotropy, directivity, uncoupled parameters estimation... An array antenna is called isotropic if it has an uniform estimation accuracy, *i.e.*, the CRB is not a function of the parameter of interest over the whole field of view. The uncoupled property is a desired criterion to have azimuth, and elevation estimation errors mutually independent, and hence, to avoid the degradation of the CRB. In [1, 15, 16, 20], the isotropy condition, and uncoupled parameters estimation for planar antenna was introduced. It showed that we can achieve both isotropic, and uncoupled properties with some particular array geometry. In the literature, considering isotropic property, the CRB is used only for the planar array as a criterion [8], [20], while mean square angular error (MSAE) is used for studying 3D array [8]. The CRB closed-form expressions previously derived are used here to find the array's configuration where isotropic, and/or uncoupled properties are attained.

– *Conditional observation model*

Because in both cases: single orthogonal branch, and two symmetric orthogonal branches, we always have the same expression for the CRB under H_1, the isotropic, and uncoupling conditions in these cases are similar. From the definition

of isotropy, and from Eqn. (15), both isotropic (only in terms of azimuth), and uncoupling are obtained if

$$S_{12} = 0. \tag{21}$$

Since S_{12} represents the sensors located on the plane xOy, we can deduce a criterion for the sensors positioning which respects to Eqn. (21):

$$\begin{cases} \sum\limits_{i=1}^{N_1} \rho_{1,i}^2 \cos 2\varphi_i = 0, \\ \sum\limits_{i=1}^{N_1} \rho_{1,i}^2 \sin 2\varphi_i = 0. \end{cases} \tag{22}$$

The L-shaped array extension is an example that can achieve criterion (22), and it will be detailed in the next section.

– *Unconditional observation model*

For the planar antenna with a single symmetric orthogonal branch, from Eqn. (16), and (17), isotropy, and uncoupled properties can be achieved if the following expressions are both satisfied:

$$\begin{cases} S_{12} = 0, \\ S_{11} = 0. \end{cases} \tag{23}$$

The expression $S_{11} = 0$ leads to:

$$\begin{cases} \sum\limits_{i=1}^{N_1} \rho_{1,i} \cos \varphi_i = 0, \\ \sum\limits_{i=1}^{N_1} \rho_{1,i} \sin \varphi_i = 0, \end{cases} \tag{24}$$

i.e., the line containing the ULA branch must pass through the centroid of the planar array. Some examples of the arrays satisfying condition (23) are shown in Fig. 4.

Contrary to the single ULA orthogonal branch case, for the planar antenna with two symmetric orthogonal branches, from Eqn. (18), isotropic, and uncoupling estimation are met if

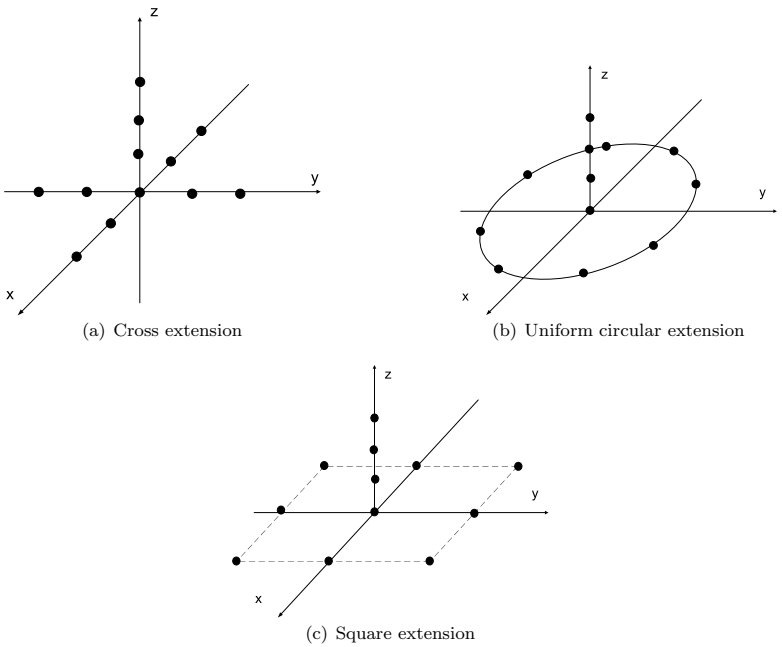
$$S_{12} = \frac{S_{11}^2}{M}. \tag{25}$$

It leads to the same solution of the planar arrays [20], where (23) is a particular solution. Hence, the sensors positions located on the xOy plane must satisfy the following criteria:

$$\begin{cases} \sum\limits_{i=1}^{N_1} \rho_{1,i}^2 \cos 2\varphi_i = \dfrac{\left(\sum\limits_{i=1}^{N_1} \rho_{1,i} \cos \varphi_i\right)^2 - \left(\sum\limits_{i=1}^{N_1} \rho_{1,i} \sin \varphi_i\right)^2}{M}, \\ \sum\limits_{i=1}^{N_1} \rho_{1,i}^2 \sin 2\varphi_i = \frac{2}{M} \sum\limits_{i=1}^{N_1} \rho_{1,i} \cos \varphi_i \sum\limits_{i=1}^{N_1} \rho_{1,i} \sin \varphi_i. \end{cases} \tag{26}$$

An intuitive solution of (25) is given by $S_{12} = S_{11} = 0$ with some antenna models shown in Fig. 4 (with two symmetric orthogonal axes).

From these analysis, we can conclude here:



(a) Cross extension (b) Uniform circular extension

(c) Square extension

Fig. 4 Various 3D isotropic array satisfying (23)

- Under H_1, by adding an orthogonal branch to the planar antenna, or under H_2 with two symmetric branches added, the conditions of isotropy, and decoupling do not change.
- However under H_2, in the case where only one orthogonal branch is added, only the particular solution $S_{11} = S_{12} = 0$ leads to the isotropy, and decoupling.

4.4.2 Conditional versus unconditional models

Intuitively, one can observe that the CRB expressions under H_1 are generally more compact than under H_2. Surprisingly, by comparing Eqn. (15), (16), and (17) for the 3D model, and Eqn. (19), and (20) for the planar antenna, it can be noted that: the CCRB, and the UCRB can be expressed in the same term *w.r.t* the sensors' location, if the following condition is satisfied:

$$S_{11} = S_{23} = 0. \tag{27}$$

In other words, the arrays will have the same behavior under both conditional, and unconditional observation models if the two ULA branches are symmetric, and the line containing these branches must pass through the centroid of the planar

antenna. Moreover, by considering the ratio between CCRB, and UCRB for this family of arrays:

$$\frac{UCRB}{CCRB} = \frac{C_{SNR}}{U_{SNR}} = 1 + \frac{1}{M\frac{\sigma_s^2}{\sigma_n^2}} = 1 + \frac{1}{M \times SNR}, \tag{28}$$

it is clear that for a large number of sensors or a high signal to noise ratio, this family of arrays has the identical estimation accuracy under both H_1, and H_2. This is consistent with the results presented in [11].

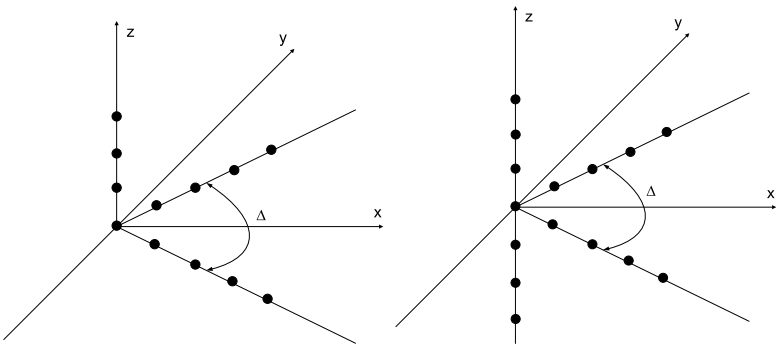
4.5 Summary

From these aforementioned results , some remarks can be done:

- The analytic, and compact expressions of the CRB under both conditional, and unconditional observation model for a family of 3D antenna arrays, and arbitrary 2D antenna arrays are derived.
- The CRB of azimuth, and elevation of the 2D models are a cosine or a sine function of the source elevation. This has been already noticed in [20] for the unconditional case, but, to the best of our knowledge, was not known in the conditional observation case. They vary in opposite ways: when the azimuth CRB is minimum, the elevation CRB is maximum, and conversely. Moreover, one can see that the CRBs of azimuth (respectively elevation) tends to infinity when elevation tends to 0° (respectively 90°). However, the CRB of elevation of the 3D arrays is no longer a sine function of elevation, and has a finite value at $\theta = 90^\circ$. Consequently, the 3D arrays model overcomes the ambiguity problem case of the 2D arrays.
- We found the conditions on the array geometry, with which we obtain the same estimation accuracy under both H_1, and H_2 assumptions.
- The isotropic, and decoupling criterions are introduced. We found that, under H_1, adding an orthogonal branch to the planar array does not change the conditions of isotropy, and decoupling. While under H_2, depending to the number of branches added (single branch or two symmetric branches), the conditions of isotropy, and decoupling may be modified then leading to a particular solution.

5 Particular cases

In the previous Section, an array geometry consisting of a single orthogonal branch (or two symmetric orthogonal branches) added to an arbitrary planar array has been considered, and closed-forms expressions of CRB have been introduced. In this Section, we will detail these CRB expressions for several important particular cases of planar antennas, and their 3D extensions in order to simplify the antenna design problem. These antenna array geometries have been widely studied in several works but almost all of them are limited to the 2D geometry arrays. In particular, the 3D extension of the V-shaped antenna array will be used here to analyze the impact of the third dimension on the estimation accuracy.



(a) V-shaped with an orthogonal branch (b) V-shaped with two orthogonal symmetric branch

Fig. 5 V-shaped extension array

5.1 3D extension of the V-shaped array

First of all, we study the V-shaped array extension consisting of a 2D V-shaped array made from two ULA branches separated by an angle denoted Δ, and from one or two opposite ULA orthogonal branches (Fig. 5). Without loss of generality, we assume that the V-shaped array is located on the xOy plane, while its ULA orthogonal branch(es) coincide(s) with the z axis. The opening angle Δ is used as a degree of freedom to find the optimal geometry. Note that in [20], the V-shaped 2D array has been studied only under the unconditional observation model. Consequently, a condition on Δ leading to an isotropic array when the number of sensors M tends to infinity was found: ($\Delta_{iso} = 2\arctan(1/2)$). The authors proved also that the V-shaped 2D array has better performance than the classical uniform circular array for the same number of sensors.

Consequently, we here extend the work of [20] to the 3D case under both conditional, and unconditional models. For this array, under both assumptions H_1, and H_2, the parameters S_{12}, S_{11}, S_{10} can be expressed as:

$$\begin{cases} S_{12} = S_{10}\cos\Delta, \\ S_{11} = S_{13}\cos\frac{\Delta}{2}, \\ \Im\{S_{12}e^{-2j\phi}\} = -S_{10}\cos\Delta\sin 2\phi, \\ \Re\{S_{12}e^{-2j\phi}\} = S_{10}\cos\Delta\cos 2\phi, \\ \Re\{S_{11}e^{-j\phi}\} = S_{13}\cos\frac{\Delta}{2}\cos\phi, \\ \Im\{S_{11}e^{-j\phi}\} = -S_{13}\cos\frac{\Delta}{2}\sin\phi, \\ \Re\Big\{e^{-2j\phi}\left(S_{12} - \frac{S_{11}^2}{M}\right)\Big\} = \left(S_{10}\cos\Delta - \frac{S_{13}^2\cos^2\frac{\Delta}{2}}{M}\right)\cos 2\phi, \\ \Re\{e^{-j\phi}S_{12}S_{11}^*\} = S_{10}S_{13}\cos\Delta\cos\frac{\Delta}{2}\cos\phi. \end{cases} \tag{29}$$

These parameters will be then applied into Eqn. (15), (16), (17), and (18) in order to find closed-form expressions of the CRB of the V-shaped 3D array extension.

5.1.1 V-shaped 2D array with an orthogonal branch

The geometry of this antenna model is presented in Fig. 5(a).

– *Conditional observation model*

The CRB is easily derived from Eqn. (15), and leads to

$$
\begin{aligned}
C_{\theta\theta} &= \frac{2}{C_{SNR}} \frac{S_{10}(1-\cos\Delta\cos 2\phi)}{\left(S_{10}^2\sin^2\Delta\cos^2\theta+\sin^2\theta 2S_{10}S_{20}(1-\cos\Delta\cos 2\Phi)\right)},\\
C_{\phi\phi} &= \frac{4}{C_{SNR}\sin^2\theta} \frac{\left(\frac{1}{2}\cos^2\theta S_{10}(\cos\Delta\cos 2\phi+1)+\sin^2\theta S_{20}\right)}{\left(S_{10}^2\sin^2\Delta\cos^2\theta+\sin^2\theta 2S_{10}S_{20}(1-\cos\Delta\cos 2\Phi)\right)},\\
C_{\theta\phi} &= \frac{1}{C_{SNR}\tan\theta} \frac{S_{10}\cos\Delta\sin 2\phi}{\left(S_{10}^2\sin^2\Delta\cos^2\theta+\sin^2\theta 2S_{10}S_{20}(1-\cos\Delta\cos 2\Phi)\right)}.
\end{aligned} \tag{30}
$$

– *Unconditional observation model*

By applying Eqn. (29) into Eqn. (16), and (17), the CRB is given by: $C_{ij}=\frac{Num_{ij}}{Den}$ where $(i,j)=\{\theta,\phi\}$, and where the denominator Den is given by

$$
\begin{aligned}
\frac{Den}{(U_{SNR})^2\sin^2\theta} &= \cos^2\theta S_{10}\left(\frac{S_{10}\sin^2\Delta}{4}+\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{2M}(\cos\Delta-1)\right)\\
&\quad+\frac{\sin 2\theta S_{23}S_{10}S_{13}}{2M}\cos\frac{\Delta}{2}(1-\cos\Delta)\cos\phi\\
&\quad+\sin^2\theta\cos 2\phi\left(\frac{S_{23}^2S_{10}\cos\Delta}{2M}-\frac{S_{20}}{2}\left(S_{10}\cos\Delta-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{M}\right)\right)\\
&\quad+\sin^2\theta\left(\frac{S_{20}S_{10}}{2}-\frac{S_{20}S_{13}^2\cos^2\frac{\Delta}{2}}{2M}-\frac{S_{10}S_{23}^2}{2M}\right).
\end{aligned} \tag{31}
$$

and where the numerators are given by

$$
\begin{aligned}
\frac{Num_{\phi\phi}}{U_{SNR}} &= \cos^2\theta\left(\frac{S_{10}}{2}-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{2M}\right)+\sin^2\theta(S_{20}-\frac{S_{23}^2}{M})\\
&\quad+\frac{\cos^2\theta\cos 2\phi}{2}\left(S_{10}\cos\Delta-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{M}\right)\\
&\quad+\frac{1}{M}\sin 2\theta S_{23}S_{13}\cos\frac{\Delta}{2}\cos\phi,\\
\frac{Num_{\theta\theta}}{U_{SNR}} &= \sin^2\theta\left(\frac{S_{10}}{2}-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{2M}\right)-\frac{\sin^2\theta\cos 2\phi}{2}\left(S_{10}\cos\Delta-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{M}\right),\\
\frac{Num_{\theta\phi}}{U_{SNR}} &= -\frac{\sin 2\theta}{4}\sin 2\phi\left(S_{10}\cos\Delta-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{M}\right)-\frac{\sin^2\theta}{M}S_{23}S_{13}\cos\frac{\Delta}{2}\sin\phi.
\end{aligned} \tag{32}
$$

The analysis of these expressions will be detailed in the next section.

5.1.2 V-shaped 2D array with two symmetric orthogonal branches

The geometry of this antenna model is presented in Fig. 5(b).

– *Conditional observation model*

The expressions of CRB under H_1 are the same as Eqn. (30).

– *Unconditional observation model*

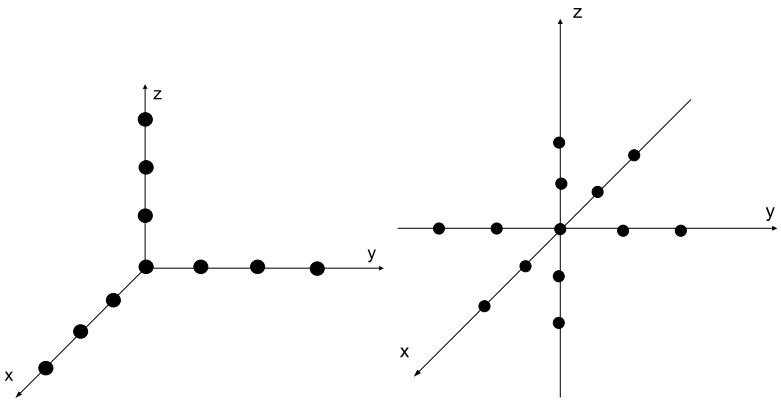
Similarly to the above section, if the 3D array is built from a planar array, and two orthogonal symmetric branches (Fig. 5(b)), by applying Eqn. (29) into Eqn. (18), we have a more compact CRB expressions given by:

$$
\begin{aligned}
C_{\phi\phi} &= \frac{\frac{\cos^2\theta\left(\frac{S_{10}}{2}-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{2M}\right)+\sin^2\theta S_{20}+\frac{1}{2}\cos^2\theta\cos 2\phi\left(S_{10}\cos\Delta-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{M}\right)}{U_{SNR}\sin^2\theta}}{\begin{pmatrix} S_{10}\cos^2\theta\left(\frac{S_{10}\sin^2\Delta}{4}+\frac{S_{13}^2\cos^2\frac{\Delta}{2}(\cos\Delta-1)}{2M}\right) \\ -S_{20}\sin^2\theta\left(\cos 2\phi\left(S_{10}\cos\Delta-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{M}\right)-\left(\frac{S_{10}}{2}-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{2M}\right)\right)\end{pmatrix}}, \\
C_{\theta\theta} &= \frac{\frac{S_{10}-\frac{1}{M}S_{13}^2\cos^2\frac{\Delta}{2}-\cos 2\phi\left(S_{10}\cos\Delta-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{M}\right)}{2U_{SNR}}}{\begin{pmatrix} S_{10}\cos^2\theta\left(\frac{S_{10}\sin^2\Delta}{4}+\frac{S_{13}^2\cos^2\frac{\Delta}{2}(\cos\Delta-1)}{2M}\right) \\ -S_{20}\sin^2\theta\left(\cos 2\phi\left(S_{10}\cos\Delta-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{M}\right)-\left(\frac{S_{10}}{2}-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{2M}\right)\right)\end{pmatrix}}, \\
C_{\theta\phi} &= \frac{\frac{\sin^2\theta\cos\theta\sin 2\phi\left(S_{10}\cos\Delta-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{M}\right)}{2U_{SNR}}}{\begin{pmatrix} S_{10}\cos^2\theta\left(\frac{S_{10}\sin^2\Delta}{4}+\frac{S_{13}^2\cos^2\frac{\Delta}{2}(\cos\Delta-1)}{2M}\right) \\ -S_{20}\sin^2\theta\left(\cos 2\phi\left(S_{10}\cos\Delta-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{M}\right)-\left(\frac{S_{10}}{2}-\frac{S_{13}^2\cos^2\frac{\Delta}{2}}{2M}\right)\right)\end{pmatrix}}.
\end{aligned}
\tag{33}
$$

These expressions concerning the V-shaped 3D array under conditional, and unconditional observation models will be analyzed in the next section.

5.2 L-shaped 3D array extension

We call "L-shaped 3D array extension" a particular case of the V-shaped 3D array where the parameter Δ is fixed to be $\Delta = \frac{\pi}{2}$. The L-shaped (2D) array has already been studied in [29] where it is shown that the L-shaped (2D) array is 37% better in terms of estimation accuracy than the cross array. Without loss of generality, let us suppose that the three branches of the array coincides with the coordinate system axes (see Fig. 6(a)).



(a) L-shaped 3D array extension (b) 3D uniform angular array

Fig. 6 Orthogonal arrays

5.2.1 Conditional observation model

Under H_1, expression (30) leads to

$$\begin{aligned} C_{\theta\theta} &= \frac{2}{C_{SNR}(S_{10}\cos^2\theta + 2S_{20}\sin^2\theta)}, \\ C_{\phi\phi} &= \frac{2}{C_{SNR}S_{10}\sin^2\theta}, \\ C_{\theta\phi} &= 0. \end{aligned} \tag{34}$$

We can notice that, in this case, the parameters θ, and ϕ are decoupled. The CRB becomes very compact.

5.2.2 Unconditional observation model

Under H_2, by letting $\Delta = 90^\circ$, Eqn. (31), and (32) become

$$\begin{aligned}
C_{\phi\phi} &= \frac{\left(\cos^2\theta\left(\frac{S_{10}}{2} - \frac{S_{13}^2(\cos 2\phi+1)}{4M}\right) + \sin^2\theta(S_{20} - \frac{S_{23}^2}{M}) + \frac{1}{M\sqrt{2}}\sin 2\theta\cos\phi S_{23}S_{13}\right)}{U_{SNR}\sin^2\theta} \\
&\quad \Bigg/ \begin{pmatrix} S_{10}\cos^2\theta\left(\frac{S_{10}}{4} - \frac{S_{13}^2}{4M}\right) + \frac{S_{23}S_{10}S_{13}\cos\phi\sin 2\theta}{2\sqrt{2}M} \\ -\frac{1}{4M}\sin^2\theta\cos 2\phi S_{20}S_{13}^2 + \sin^2\theta\left(\frac{S_{20}S_{10}}{2} - \frac{S_{20}S_{13}^2}{4M} - \frac{S_{10}S_{23}^2}{2M}\right) \end{pmatrix}, \\
C_{\theta\theta} &= \frac{\frac{1}{2U_{SNR}}\left(S_{10} + \frac{1}{2M}S_{13}^2(\cos 2\phi - 1)\right)}{\begin{pmatrix} S_{10}\cos^2\theta\left(\frac{S_{10}}{4} - \frac{S_{13}^2}{4M}\right) + \frac{S_{23}S_{10}S_{13}\cos\phi\sin 2\theta}{2\sqrt{2}M} \\ -\frac{1}{4M}\sin^2\theta\cos 2\phi S_{20}S_{13}^2 + \sin^2\theta\left(\frac{S_{20}S_{10}}{2} - \frac{S_{20}S_{13}^2}{4M} - \frac{S_{10}S_{23}^2}{2M}\right) \end{pmatrix}}, \\
C_{\theta\phi} &= \frac{\frac{1}{\sqrt{2}MU_{SNR}\sin^2\theta}\left(\frac{1}{4\sqrt{2}}S_{13}^2\sin 2\theta\sin 2\phi - \sin^2\theta\sin\phi S_{23}S_{13}\right)}{\begin{pmatrix} S_{10}\cos^2\theta\left(\frac{S_{10}}{4} - \frac{S_{13}^2}{4M}\right) + \frac{S_{23}S_{10}S_{13}\cos\phi\sin 2\theta}{2\sqrt{2}M} \\ -\frac{1}{4M}\sin^2\theta\cos 2\phi S_{20}S_{13}^2 + \sin^2\theta\left(\frac{S_{20}S_{10}}{2} - \frac{S_{20}S_{13}^2}{4M} - \frac{S_{10}S_{23}^2}{2M}\right) \end{pmatrix}}.
\end{aligned} \tag{35}$$

5.3 3D uniform angular array

A natural variant of "L-shaped 3D extension array", presented in Fig. 6(b), can be considered. This array is called 3D uniform angular antenna array (UAA). In [17], the UAA has been proved that it minimizes the CRB for the case of source position's estimation. Thanks to its special structure, which is totally symmetric, its CRB becomes more compact due to the fact that $S_{11} = S_{12} = S_{13} = S_{23} = 0$, so we obtain:

5.3.1 Conditional observation model

The CRB is the same as Eqn. (34).

5.3.2 Unconditional observation model

$$\begin{cases} C_{\phi\phi} = \frac{2}{U_{SNR}S_{10}\sin^2\theta}, \\ C_{\theta\theta} = \frac{2}{U_{SNR}(S_{10}\cos^2\theta + 2S_{20}\sin^2\theta)}, \\ C_{\theta\phi} = 0. \end{cases}$$

From Eqn. (34), and (36), we observe that under H_1, and H_2, the CRB of the UAA has identical expressions except the terms C_{SNR} under H_1, and U_{SNR} under H_2. Therefore, we conclude that the UAA has the similar behaviors under both conditional, and unconditional observation assumptions.

Moreover, if we choose the array structure such as $S_{20} = \frac{S_{10}}{2}$, *i.e.*, the number of sensors of the six branches are equal, or $N_2 = N_2' = \frac{N_1 - 1}{4}$, then the CRB of elevation is independent to both the azimuth, and elevation, *i.e.*, to DOA.

5.4 Analysis

In this Section, the aforementioned results for the particular antenna models are analyzed in order to find the isotropy, uncoupling condition, and also to compare their behavior under the conditional, and unconditional assumptions.

5.4.1 Isotropy, and uncoupling properties

In this case, our purpose is to find the value of the degree of freedom Δ_{iso} with which, the V-shaped extension arrays attain isotropy, and/ or decoupling.

– *Conditional observation model*

The condition of isotropy, and decoupling (21) leads Eqn. (30) to $\Delta_{iso} = 90^\circ$ for both V-shaped with a single orthogonal branch or with two symmetric orthogonal branches antenna. It can be noted that this case is in contradiction with the results mentioned in [20] for the unconditional model, and 2D array, and with the results obtained below.

– *Unconditional observation model*

Concerning the V-shaped array with an orthogonal branch, from condition (23), the isotropic property is achieved if $S_{11} = 0$ is satisfied, *i.e.*, the line containing the ULA branch must pass through the centroid of the planar antenna. Given the fact that the line containing the ULA branch does not pass through the centroid of the planar part of the V-shaped 3D extension, therefore, there does not exist any value of Δ satisfying the isotropic condition.

Concerning the V-shaped array with two symmetric orthogonal branches, from Eqn. (33), and (25), we can see that Δ_{iso} is the solution of equation $S_{12} - \frac{S_{11}^2}{M} = 0$. Consequently, depending on the method used to make the branches of the antenna array (ULA, minimum redundancy [34], D-optimal [35], etc.) we might obtain different values of Δ_{iso}. In the case where the antenna array is made from ULA, then from (26), it easily leads to:

$$\Delta_{iso} = \arccos\left(\frac{3(N_1^2 - 1)}{8MN_1 - 3N_1^2 + 3}\right). \tag{36}$$

Let us set the positive $\alpha = \frac{N_1}{M} \leq 1$. The value of α associated to a planar antenna will be equal to 1, while that one associated to a 3D antenna array is strictly lower than 1. Then, Δ_{iso} can be expressed as $\Delta_{iso} = \arccos\left(\frac{3(\alpha^2 - 1/M^2)}{8\alpha - 3\alpha^2 + 3/M^2}\right)$. We are interested also to define the range of Δ_{iso} *w.r.t.* α in this case. It is clear that:

$$\begin{array}{l} \text{If } \alpha \to 1, \text{ and } M \gg 1 \Rightarrow \Delta_{iso} \simeq \arccos\left(\frac{3}{5}\right) = 53.13^\circ \\ \text{If } \alpha \to 0, \text{ and } M \gg 1 \Rightarrow \Delta_{iso} \simeq \arccos\left(0\right) = 90^\circ \end{array} \tag{37}$$

In Fig. 7, when α tends to 0, *i.e*, the number of sensors located on the orthogonal axis is much larger than the number of sensors located on the planar array, then, the value of Δ_{iso} tends to 90°. On the contrary, if α tends to 1, *i.e*, the number of sensors located on the planar array is much larger than those located on the orthogonal axis, then, the value of Δ_{iso} tends to $\arccos(3/5)$. In particular, in the case where $\alpha = 1$, we obtain exactly the same result ($\Delta_{iso} = 53.13^\circ$) for the

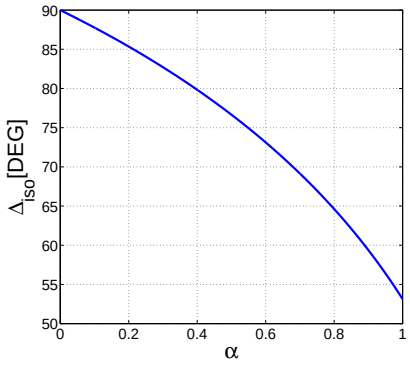


Fig. 7 Variation of Δ_{iso} *w.r.t.* α with $M = 1000$.

planar antenna array as in [20]. Therefore, Δ_{iso} in this case will vary from 53.13° to 90°.

A remark can be done here that under H_2, adding two symmetric orthogonal branches does not modify the conditions of isotropic, and decoupling ($S_{12} = S_{11}^2/M$) *w.r.t* the planar array, but it changes the arrangement of the sensors located on the planar part because of the intervention of N_2 to S_{12}, and S_{11}.

5.4.2 Conditional versus unconditional models

Since the V-shaped 3D extension array does not satisfy condition (27) because the the line containing the ULA branch does not pass through the centroid of the planar part of the antenna, then it is impossible to find an optimal value of Δ, with which, the CCRB, and the UCRB have the same expressions. The CCRB in this case is always more compact than the UCRB. Contrary to the V-shaped 3D extension array, the 3D UAA satisfies well condition (27), therefore the UCRB, and CCRB will have the identical compact expression at high SNR or for a large number of sensors.

5.5 Summary

Thanks to the degree of freedom Δ of the V-shaped family arrays, the analysis of the impact of the array geometry on the estimation performance is simplified. We here can make some remarks:

- In almost cases, CCRB has a more compact expression than UCRB.
- Under H_1, the value of Δ_{iso} is constant ($\Delta_{iso} = 90°$), while it takes a range of values under H_2, depending to the antenna array configuration. In particular, when $\alpha = 1$, we find the same results ($\Delta_{iso} = 53.13°$) for the V-shaped (2D) antenna as in [20].

– The 3D uniform angular array has several advantages: isotropy, uncoupling, minimization of the CRB in case of the source position's location using TDOA method, and the same estimation accuracy under both the H_1 or H_2 assumptions.

6 Comparison of the estimation accuracy

In this section, we will use the closed form expressions of the CRB calculated in the previous section to compare the estimation performance between the above studied arrays with other classical arrays. In order to simplify the array design problem, we only consider the behavior of the CRB of the V-shaped antenna array and its 3D extension. Its closed form CRB will be analyzed *w.r.t.* the opening angle Δ. For the simulation, all branches of the antenna array being 2D (two branches) either 3D (three branches or four branches) are made from ULAs with the inter-sensor space of half the wavelength. The simulations are performed with a signal to noise ratio equal to 10 dB and a number of snapshots $T = 100$.

6.1 Comparison of the estimation performance between the V-shaped 3D extension antenna array and the planar circular antenna array

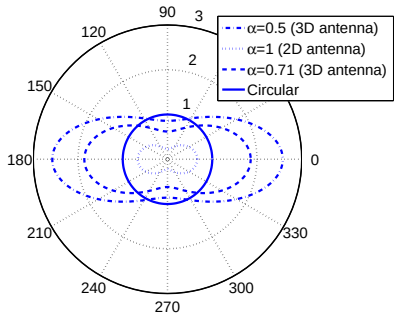
We here compare the estimation performance between the V-shaped antenna array with an isotropic classic antenna: the uniform circular antenna (UCA). For this comparison, the antenna arrays will have the same number of sensors. The sensors of UCA are half-wavelength inter-element spaced, thus, the value of its radius is given by $r = \frac{\lambda}{4 \sin \frac{\pi}{M}}$. Figs. 8 and 9 represent respectively the CRB of azimuth and elevation normalized by the CRB of the UCA ($C_{\theta\theta}/C_{\theta\theta}^{(UCA)}$, $C_{\phi\phi}/C_{\phi\phi}^{(UCA)}$) *w.r.t.* the aforementioned coefficient α, at the opening angle $\Delta = 60^\circ$ and at the elevation $\theta = 45^\circ$ under both conditional and unconditional observation models.

– *Conditional observation model*

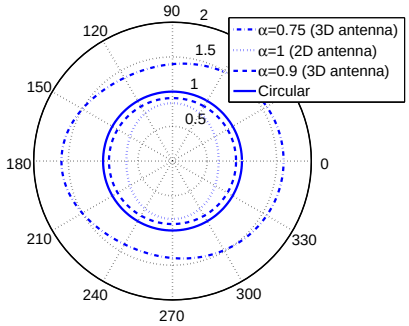
In Fig 9(a), the accuracy concerning the elevation estimation of the V-shaped antenna is always lower, *i.e.*, better than the UCA. In Fig 8(a), it is shown that the performance concerning the azimuth estimation is strictly linked to the number of sensors located on the orthogonal branch, *i.e.*, on the coefficient α. We observed that when the ratio α varies, the estimation performance concerning azimuth and elevation varies differently. When the one improves, the other deteriorates. For the value of α close to 1, *i.e.*, almost of the sensors located on the planar antenna, the estimation accuracy in terms of both the azimuth and elevation of the V-shaped family is better than the one of the UCA.

– *Unconditional observation model*

Figs. 8(b) and 9(b) show that the performance concerning estimation of both azimuth and elevation are strongly dependent on the number of sensors located on the orthogonal branch, *i.e.*, the coefficient α. The link between α and the CRB under H_2 is more complicated than under H_1. When α decreases, then the CRB concerning azimuth estimation deteriorates, while the CRB concerning elevation estimation varies differently according to the DOA: it improves in some zone of



(a) Conditional observation model



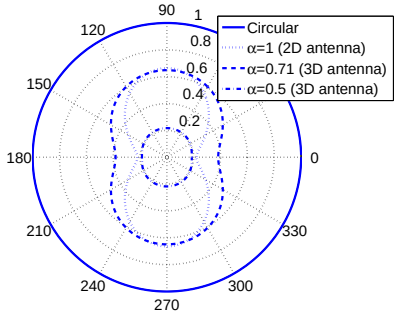
(b) Unconditional observation model

Fig. 8 Polar representation of the normalized CRB of azimuth for all values of azimuth angle, with different values of α, $\Delta = 60°$, and $\theta = 45°$. The array has a single orthogonal branch.

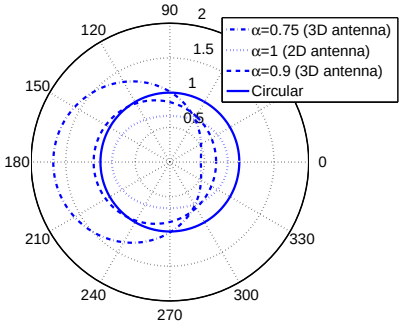
DOA while worsens in the other zones. For the value of α close to 1, the V-shaped family performs better in terms of both azimuth and elevation estimation than the UCA.

6.2 Comparison of the estimation performance of the isotropic antennas

We are interested in considering the case where our array attains the isotropic and uncoupling properties. We here compare the V-shaped isotropic array ($\Delta_{iso} = \frac{\pi}{2}$ under H_1 and $\Delta_{iso} = \arccos\left(\frac{3(\alpha^2 - 1/M^2)}{8\alpha - 3\alpha^2 + 3/M^2}\right)$ under H_2) with the classical isotropic UCA. As mentioned in the previous section, under H_2, the 3D V-shaped extension



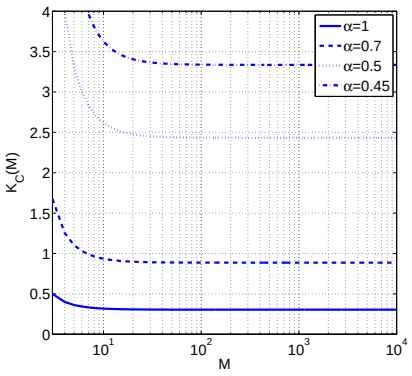
(a) Conditional observation model



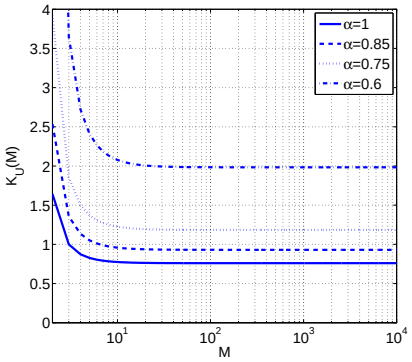
(b) Unconditional observation model

Fig. 9 Polar representation of the normalized CRB of elevation for all values of azimuth angle, with different values of α, $\Delta = 60°$, and $\theta = 45°$. The array has a single orthogonal branch.

array becomes an isotropic array if there are two symmetric orthogonal branches. Therefore, under H_1, a 3D V-shaped array with a single orthogonal branch is used while under H_2, a 3D V-shaped array with two symmetric orthogonal branches is used. We consider the ratio $K_C(M)$ (under H_1) or $K_U(M)$ (under H_2) between the CRB concerning the estimation of the azimuth of a family of V-shaped isotropic arrays and the UCA array. Thus, we have $K_C(M) = K_U(M) = \frac{C^{2D}_{\phi\phi}}{C^{UCA}_{\phi\phi}}$ if $\alpha = 1$ and $K_C(M) = K_U(M) = \frac{C^{3D}_{\phi\phi}}{C^{UCA}_{\phi\phi}}$ if $\alpha < 1$. Therefore, this fraction shows the gain in estimation of azimuth accuracy of the family of V-shaped arrays *w.r.t.* to the UCA array.



(a) Conditional observation model



(b) Unconditional observation model

Fig. 10 Fraction $K(M)$ in term of the number of sensors M

– *Conditional observation model*

From (34), the ratio of CRB concerning azimuth of these antenna arrays is given by:

$$\begin{array}{l} K_C(M) = \frac{3}{\alpha(\alpha^2 M^2 - 1)\sin^2\frac{\pi}{M}}. \\ \text{If } \alpha M >> 1 \rightarrow K_C(M) = \frac{3}{\pi^2\alpha^3}. \end{array} \tag{38}$$

We can say that the V-shaped antenna array is better than the UCA array in terms of the estimation of azimuth if and only if the fraction $K_C(M)$ is smaller than 1. Fig. 10(a) shows that the 3D V-shaped isotropic antenna array is better than UCA array provided that the value of α satisfies: $0.76 < \alpha < 1$ and $M > 6$.

– *Unconditional observation model*

From (33), after some calculations, the ratio $K_U(M)$ is given by:

$$\begin{array}{l} K_U(M) = \frac{3(8\alpha M^2 - 3\alpha^2 M^2 + 3)}{\sin^2 \frac{\pi}{M} \alpha(\alpha^2 M^2 - 1)(8\alpha M^2 - 6\alpha^2 M^2 + 6)}. \\ \text{If } \alpha M >> 1 \rightarrow K_U(M) = \frac{3(8-3\alpha)}{\alpha^3(8-6\alpha)\pi^2}. \end{array} \quad (39)$$

Fig. 10(b) shows that the 3D V-shaped isotropic antenna array is better than the UCA array if: $0.84 < \alpha < 1$ and $M > 7$.

Table 1 'The azimuth estimation performance gain of 3D V-shaped isotropic antenna according to UCA'

α	1	0.9	0.8	0.7	0.6
H_1 assumption	0.6959	0.5829	0.4060	0.1133	-0.4081
H_2 assumption	0.2399	0.1498	-0.0393	-0.3765	-0.9838

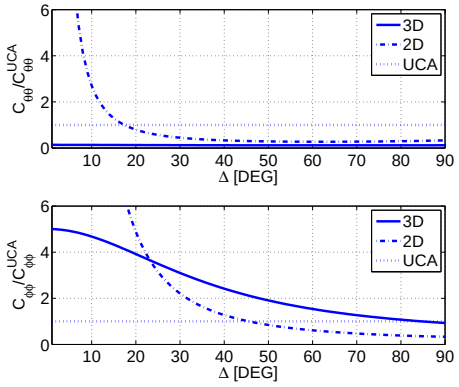
Tab. 1 shows the value of $1 - K_U(M)$ and $1 - K_C(M)$ *w.r.t.* α. These values represent the gain concerning the azimuth estimation of the 3D V-shaped isotropic antenna array to the UCA array for a large number of sensors. We here want to find the value of α, with which $1 - K_C(M) > 0$ under H_1 or $1 - K_U(M) > 0$ under H_2 *i.e.*, the 3D V-shaped antenna array has the better azimuth estimation accuracy than the UCA array. Under both H_1 and H_2 assumption, it is clear that, for all $\alpha > 0.85$, the 3D V-shaped isotropic array is always better than the UCA. Moreover, if $\alpha = 1$ then the azimuth estimation accuracy of the V-shaped isotropic planar array is at least 20% better than the UCA array.

6.3 Comparison of the estimation performance between 2D and 3D antenna arrays

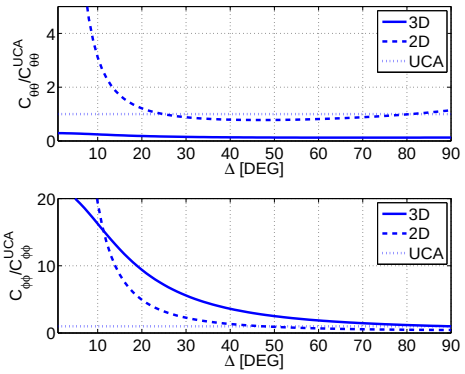
In the following, we compare the performance of estimation between the 3D and 2D arrays. The V-shaped 2D antenna array has $M = 7$ sensors (one at the origin with three other sensors on each branch). The V-shaped 3D extension antenna array consisting of a single orthogonal branch is made also from $M = 7$ sensors (one at the origin and two sensors on every three branches). It should be noted that taking some sensors from the planar array of the 2D antenna array to make the 3D antenna array will decrease the aperture and hence, reduce its performance. Therefore, using non ULA such as minimum redundancy, D-optimal, etc. instead of using ULA can maintain the aperture and also, the performance.

Fig. 11 shows the behaviors of $C^{3D}_{\theta\theta}$, $C^{3D}_{\phi\phi}$, $C^{2D}_{\theta\theta}$, $C^{2D}_{\phi\phi}$ in terms of the opening angle Δ varying from 0° to 90° under H_1 and H_2, and at $\phi = 20°$ and $\theta = 70°$. This is the scenario where the source is close to the plane of the array. Under both two assumptions, for the estimation of elevation, θ, we can see that the 3D antenna array has always the better performance compared to the 2D antenna. However, concerning the azimuth estimation, the 3D array only has better performance than the 2D array if $\Delta < 20°$ under H_1 or $\Delta < 12°$ under H_2.

Fig. 12 shows the same curves, but values of ϕ and θ are respectively equal to 50° and 30°. This is the scenario where the source is far from the plane of the



(a) Conditional observation model

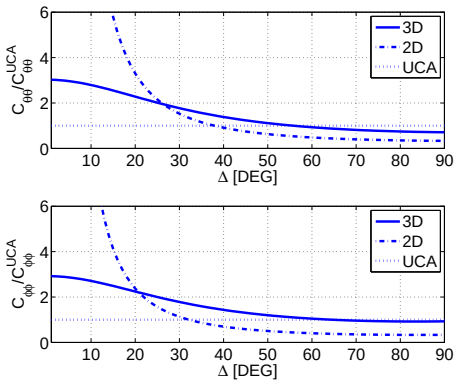


(b) Unconditional observation model

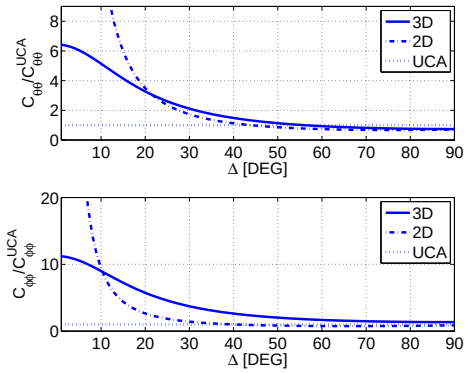
Fig. 11 The behavior of $C^{3D}_{\theta\theta}$, $C^{2D}_{\theta\theta}$, $C^{3D}_{\phi\phi}$ and $C^{2D}_{\phi\phi}$ normalized by the CRB of the UCA according to Δ at $\phi = 20°$ and $\theta = 70°$

antenna array. In this case, for both H_1 and H_2 assumptions, it should be better, contrary to intuition, to choose the 2D antenna array over a limited opening angle obtained numerically by solving $\max\left(C^{3D}_{\phi\phi} = C^{2D}_{\phi\phi}, C^{3D}_{\theta\theta} = C^{2D}_{\theta\theta}\right)$ as a function of Δ.

Finally, by an exhaustive research over all three parameters: elevation, azimuth and opening angle, we found that concerning the elevation estimation, the 3D antenna array is always the better than the 2D antenna if the elevation is larger than a certain threshold θ_0.



(a) Conditional observation model



(b) Unconditional observation model

Fig. 12 The behavior of $C^{3D}_{\theta\theta}$, $C^{2D}_{\theta\theta}$, $C^{3D}_{\phi\phi}$ and $C^{2D}_{\phi\phi}$ normalized by the CRB of the UCA according to Δ at $\phi = 50°$ and $\theta = 30°$

– *Conditional observation model*

We can prove in this case that the threshold θ_0 is about $62.2°$ by solving

$$\frac{C^{3D}_{\theta\theta}}{C^{2D}_{\theta\theta}} < 1 \Leftrightarrow \theta > \arctan\sqrt{\max_{\Delta,\phi}\{\Gamma\}}, \tag{40}$$

where $\Gamma = \frac{\sin^2\Delta\big((M^2-1)-\alpha(\alpha^2M^2-1)\big)}{(1-\cos\Delta\cos 2\phi)4(1-\alpha)((1-\alpha)M+1)(2(1-\alpha)M+1)}$, $\alpha = \frac{N_1}{M} = \frac{5}{7}, M = 7, \theta \in [0°, 90°], \Delta \in (0°, 180°), \phi \in [0°, 360°]$.

– *Unconditional observation model*

By numerical calculus, in the case where $\alpha = 5/7$ and $M = 7$, we obtain the threshold $\theta_0 \simeq 65^\circ$.

7 Conclusion

In this paper, we derived the closed form expressions of the CRB for the estimation of azimuth and elevation of a far field, single source in both conditional and unconditional observation models where a planar array or its 3D extension is used. The 3D array extension here is made by adding one or two orthogonal branches to an arbitrary planar array. These CRB closed form expressions are used here as a useful tool in order to find the isotropy, uncoupling conditions, and the contribution of the third dimension to the estimation accuracy and also to introduce a comparison between conditional and unconditional observation models. Consequently, we showed that the 3D array overcomes the ambiguity problem of the planar (2D) array. Moreover, we found that there exists a family of array geometries with which the CRB can be expressed in the same term under both conditional and unconditional assumptions. Furthermore, at high signal to noise ratio or with a large number of sensors, the CRB expressions under the two assumptions become identical. In the following step, the CRB closed form expressions are then applied into several particular well-known array geometries such as: the V-shaped/ L-shaped array 3D extension, the uniform angular array. It is shown that the isotropy and uncoupling conditions of the 3D array under conditional and unconditional assumptions are different from each other. In particular, for the V-shaped arrays family, under the unconditional observation model, the opening angle Δ_{iso} depends on the number of sensors located on the orthogonal branches while $\Delta_{iso} = 90^\circ$ is the desired value under the conditional assumption. Finally, through several simulations, we conclude that the performance of estimation of the 3D array strongly depends on the rate between the number of sensors located on the orthogonal branches and the total number of sensors (α). When this rate varies, the estimation concerning azimuth and elevation varies differently. In the other hand, by choosing a suitable rate (α close to 1), the 3D array has the better performance than the classical UCA concerning both azimuth and elevation estimation for the same number of sensors. It should be noted that, for a constant number of sensors, adding the 3D branch will decrease the aperture of the antenna, therefore, deteriorate the estimation performance.

Appendix

A Proof of Eqn. (15)

Let us note that the sensors located on the xOy plane are such that $\xi_i = \frac{\pi}{2}$, while the sensors located on the orthogonal axe are such that $\xi_i = 0$. After some calculation, from (7), it easy

to obtain the elements of the Fisher Information Matrix:

$$\begin{aligned}\frac{[\mathbf{FIM}(\mathbf{\Theta})]_{11}}{C_{SNR}} &= \sum_{i=1}^{M}\rho_i^2\left(\cos\theta\sin\xi_i\cos\left(\phi-\varphi_i\right)-\cos\xi_i\sin\theta\right)^2\\ &= \sum_{i=1}^{N_1}\rho_{1,i}^2\cos^2\theta\cos^2\left(\phi-\varphi_i\right)+\sum_{i=N_1+1}^{M}\rho_{2,i}^2\sin^2\theta\\ &= \frac{\cos^2\theta}{4}\left(e^{2j\phi}S_{12}^*+e^{-2j\phi}S_{12}+2S_{10}\right)+\sin^2\theta S_{20}\\ &= \frac{1}{2}\cos^2\theta\left(\Re\{e^{-2j\phi}S_{12}\}+S_{10}\right)+\sin^2\theta S_{20},\end{aligned} \tag{41}$$

$$\begin{aligned}\frac{[\mathbf{FIM}(\mathbf{\Theta})]_{22}}{C_{SNR}} &= \sum_{i=1}^{M}\rho_i^2\left(\sin\theta\sin\xi_i\sin\left(\phi-\varphi_i\right)\right)^2\\ &= \sin^2\theta\sum_{i=1}^{N_1}\rho_{1,i}^2\sin^2\left(\phi-\varphi_i\right)\\ &= -\frac{\sin^2\theta}{4}\left(e^{2j\phi}S_{12}^*+e^{-2j\phi}S_{12}-2S_{10}\right)\\ &= -\frac{1}{2}\sin^2\theta\left(\Re\{e^{-2j\phi}S_{12}\}-S_{10}\right),\end{aligned} \tag{42}$$

and

$$\begin{aligned}\frac{[\mathbf{FIM}(\mathbf{\Theta})]_{12}}{C_{SNR}} &= -\sum_{i=1}^{M}\rho_i^2\sin\theta\sin\xi_i\sin\left(\phi-\varphi_i\right)\left(\cos\theta\sin\xi_i\cos\left(\phi-\varphi_i\right)-\cos\xi_i\sin\theta\right)\\ &= -\sin\theta\cos\theta\sum_{i=1}^{N_1}\rho_{1,i}^2\sin\left(\phi-\varphi_i\right)\cos\left(\phi-\varphi_i\right)\\ &= -\frac{1}{8j}\sin 2\theta\left(e^{2j\phi}S_{12}^*-e^{-2j\phi}S_{12}\right)\\ &= \frac{1}{4}\sin 2\theta\Im\{e^{-2j\phi}S_{12}\}.\end{aligned} \tag{43}$$

The FIM determinant is given by:

$$\begin{aligned}&\frac{\det\left[\mathbf{FIM}(\mathbf{\Theta})\right]}{C_{SNR}^2} = \frac{[\mathbf{FIM}(\mathbf{\Theta})]_{11}[\mathbf{FIM}(\mathbf{\Theta})]_{22}-[\mathbf{FIM}(\mathbf{\Theta})]_{12}[\mathbf{FIM}(\mathbf{\Theta})]_{21}}{C_{SNR}^2}\\ &= \left(\tfrac{\cos^2\theta}{4}\left(e^{2j\phi}S_{12}^*+e^{-2j\phi}S_{12}+2S_{10}\right)+\sin^2\theta S_{20}\right)\left(-\tfrac{\sin^2\theta}{4}\left(e^{2j\phi}S_{12}^*+e^{-2j\phi}S_{12}-2S_{10}\right)\right)\\ &\quad-\left(-\tfrac{1}{8j}\sin 2\theta\left(e^{2j\phi}S_{12}^*-e^{-2j\phi}S_{12}\right)\right)^2\\ &= \tfrac{\sin^2 2\theta}{64}\left(4S_{10}^2-\left(e^{2j\phi}S_{12}^*+e^{-2j\phi}S_{12}\right)^2\right)-\tfrac{\sin^4\theta}{4}S_{20}\left(e^{2j\phi}S_{12}^*+e^{-2j\phi}S_{12}-2S_{10}\right)\\ &\quad+\tfrac{\sin^2 2\theta}{64}\left(e^{2j\phi}S_{12}^*-e^{-2j\phi}S_{12}\right)^2\\ &= \tfrac{\sin^2 2\theta}{64}\left(4S_{10}^2-4|S_{12}|^2\right)+\tfrac{\sin^4\theta}{4}S_{20}\left(2S_{10}-2\Re\{e^{-2j\phi}S_{12}\}\right)\\ &= \tfrac{\sin^2\theta}{4}\left(\cos^2\theta(S_{10}^2-|S_{12}|^2)+2\sin^2\theta S_{20}\left(S_{10}-\Re\{e^{-2j\phi}S_{12}\}\right)\right).\end{aligned} \tag{44}$$

B Proof of Eqn. (16) and Eqn. (17)

In the same way as for conditional case, from (13), we have

$$
\begin{aligned}
&\tfrac{[\mathbf{FIM}(\boldsymbol{\Theta})]_{1,1}}{U_{SNR}} = \\
&= \sum_{i=1}^{M} \rho_i^2 \left(\cos\theta \sin\xi_i \cos\left(\phi-\varphi_i\right) - \cos\xi_i \sin\theta\right)^2 \\
&\quad - \tfrac{1}{M}\left(\sum_{i=1}^{M} \rho_i \left(\cos\theta \sin\xi_i \cos\left(\phi-\varphi_i\right) - \cos\xi_i \sin\theta\right)\right)^2 \\
&= \cos^2\theta \sum_{i=1}^{N_1} \rho_{1,i}^2 \cos^2\left(\phi-\varphi_i\right) + \sin^2\theta \sum_{i=N_1+1}^{M} \rho_{2,i}^2 \\
&\quad - \tfrac{1}{M}\left(\cos\theta \sum_{i=1}^{N_1} \rho_{1,i} \cos\left(\phi-\varphi_i\right) - \sin\theta \sum_{i=N_1+1}^{M} \rho_{2,i}\right)^2 \\
&= \tfrac{1}{4}\cos^2\theta \left(S_{12} e^{-2j\phi} + S_{12}^* e^{2j\phi} + 2S_{10}\right) + \sin^2\theta S_{20} \\
&\quad - \tfrac{1}{M}\left(\tfrac{\cos^2\theta}{4}\left(S_{11}^2 e^{-2j\phi} + S_{11}^{2*} e^{2j\phi} + 2|S_{11}|^2\right)\right. \\
&\qquad \left. + \sin^2\theta S_{23}^2 - \tfrac{\sin 2\theta}{2}\left(S_{11} e^{-j\phi} + S_{11}^* e^{j\phi}\right) S_{23}\right) \\
&= \tfrac{1}{2}\cos^2\theta \left(S_{10} - \tfrac{|S_{11}|^2}{M} + \Re\left\{e^{-2j\phi}\left(S_{12} - \tfrac{S_{11}^2}{M}\right)\right\}\right) \\
&\quad + \sin^2\theta \left(S_{20} - \tfrac{S_{23}^2}{M}\right) + \tfrac{\sin 2\theta}{M} S_{23} \Re\{e^{-j\phi} S_{11}\},
\end{aligned}
\tag{45}
$$

$$
\begin{aligned}
&\tfrac{[\mathbf{FIM}(\boldsymbol{\Theta})]_{2,2}}{U_{SNR}} = \sum_{i=1}^{M} \rho_i^2 \sin^2\theta \sin^2\xi_i \sin^2\left(\phi-\varphi_i\right) - \tfrac{1}{M}\left(\sum_{i=1}^{M} \rho_i \sin\theta \sin\xi_i \sin\left(\phi-\varphi_i\right)\right)^2 \\
&= \sin^2\theta \sum_{i=1}^{N_1} \rho_{p,i}^2 \sin^2\left(\phi-\varphi_i\right) - \tfrac{\sin^2\theta}{M}\left(\sum_{i=1}^{N_1} \rho_{1,i} \sin\left(\phi-\varphi_i\right)\right)^2 \\
&= -\tfrac{1}{4}\sin^2\theta \left(e^{-2j\phi} S_{12} + e^{2j\phi} S_{12}^* - 2S_{10}\right) + \tfrac{\sin^2\theta}{4M}\left(e^{2j\phi} S_{11}^{2*} + e^{-2j\phi} S_{11}^2 - 2|S_{11}|^2\right) \\
&= \tfrac{1}{2}\sin^2\theta \left(S_{10} - \tfrac{|S_{11}|^2}{M} - \Re\left\{e^{-2j\phi}\left(S_{12} - \tfrac{S_{11}^2}{M}\right)\right\}\right),
\end{aligned}
\tag{46}
$$

and

$$
\begin{aligned}
&\tfrac{[\mathbf{FIM}(\boldsymbol{\Theta})]_{1,2}}{U_{SNR}} = -\sum_{i=1}^{M} \rho_i \sin\theta \sin\xi_i \sin^2\left(\phi-\varphi_i\right)\left(\cos\theta \sin\xi_i \cos\left(\phi-\varphi_i\right) - \cos\xi_i \sin\theta\right) \\
&\quad + \tfrac{1}{M}\sum_{i=1}^{M} \rho_i \left(\cos\theta \sin\xi_i \cos\left(\phi-\varphi_i\right) - \cos\xi_i \sin\theta\right) \sum_{i=1}^{M} \rho_i \sin\theta \sin\xi_i \sin\left(\phi-\varphi_i\right) \\
&= -\sin\theta\cos\theta \sum_{i=1}^{N_1} \rho_{1,i}^2 \sin\left(\phi-\varphi_i\right)\cos\left(\phi-\varphi_i\right) \\
&\quad + \tfrac{1}{M}\sin\theta \left(\cos\theta \sum_{i=1}^{N_1} \rho_{1,i} \cos\left(\phi-\varphi_i\right) - \sin\theta \sum_{i=N_1+1}^{M} \rho_{2,i}\right) \sum_{i=1}^{N_1} \rho_{1,i} \sin\left(\phi-\varphi_i\right) \\
&= -\tfrac{\sin\theta\cos\theta}{4j}\left(e^{2j\phi} S_{12}^* - e^{-2j\phi} S_{12}\right) \\
&\quad + \tfrac{\sin\theta}{2jM}\left(\tfrac{1}{2}\cos\theta (e^{j\phi} S_{11} + e^{-j\phi} S_{11}^*) - \sin\theta S_{23}\right)\left(e^{j\phi} S_{11}^* - e^{-j\phi} S_{11}\right) \\
&= \tfrac{\sin\theta\cos\theta}{2}\Im\left\{e^{-2j\phi}\left(S_{12} - \tfrac{S_{11}^2}{M}\right)\right\} + \tfrac{\sin^2\theta}{M} S_{23} \Im\{e^{-j\phi} S_{11}\}.
\end{aligned}
\tag{47}
$$

The FIM determinant is given by:

$$
\begin{aligned}
&\frac{\det\left[\mathbf{FIM}(\boldsymbol{\Theta})\right]}{U_{SNR}^2} = \frac{[\mathbf{FIM}(\boldsymbol{\Theta})]_{1,1}[\mathbf{FIM}(\boldsymbol{\Theta})]_{2,2}-[\mathbf{FIM}(\boldsymbol{\Theta})]_{1,2}[\mathbf{FIM}(\boldsymbol{\Theta})]_{2,1}}{U_{SNR}^2}\\
&= \left(\tfrac{1}{2}\cos^2\theta\left(S_{10}-\tfrac{|S_{11}|^2}{M}+\Re\left\{e^{-2j\phi}\left(S_{12}-\tfrac{S_{11}^2}{M}\right)\right\}\right)\right.\\
&\quad\left.+\sin^2\theta\left(S_{20}-\tfrac{S_{23}^2}{M}\right)+\tfrac{\sin 2\theta}{M}S_{23}\Re\{e^{-j\phi}S_{11}\}\right)\\
&\quad\times\left(\tfrac{1}{2}\sin^2\theta\left(S_{10}-\tfrac{|S_{11}|^2}{M}-\Re\left\{e^{-2j\phi}\left(S_{12}-\tfrac{S_{11}^2}{M}\right)\right\}\right)\right)\\
&\quad-\left(\tfrac{1}{2}\sin\theta\cos\theta\Im\left\{e^{-2j\phi}\left(S_{12}-\tfrac{S_{11}^2}{M}\right)\right\}+\tfrac{1}{M}\sin^2\theta S_{23}\Im\{e^{-j\phi}S_{11}\}\right)^2\\
&= \tfrac{1}{4}\sin^2\theta\cos^2\theta\left(\left(S_{10}-\tfrac{|S_{11}|^2}{M}\right)^2-\Re^2\left\{e^{-2j\phi}\left(S_{12}-\tfrac{S_{11}^2}{M}\right)\right\}\right)\\
&\quad+\tfrac{1}{2}\sin^2\theta\left(S_{10}-\tfrac{|S_{11}|^2}{M}-\Re\left\{e^{-2j\phi}\left(S_{12}-\tfrac{S_{11}^2}{M}\right)\right\}\right)\\
&\quad\times\left(\sin^2\theta\left(S_{20}-\tfrac{S_{23}^2}{M}\right)+\tfrac{\sin 2\theta}{M}S_{23}\Re\{e^{-j\phi}S_{11}\}\right)\\
&\quad-\tfrac{1}{4}\sin^2\theta\cos^2\theta\Im^2\left\{e^{-2j\phi}\left(S_{12}-\tfrac{S_{11}^2}{M}\right)\right\}-\tfrac{1}{M^2}\sin^4\theta S_{23}^2\Im^2\{e^{-j\phi}S_{11}\}\\
&\quad-\tfrac{1}{M}\sin^3\theta\cos\theta\Im\left\{e^{-2j\phi}\left(S_{12}-\tfrac{S_{11}^2}{M}\right)\right\}S_{23}\Im\{e^{-j\phi}S_{11}\}\\
&= \tfrac{1}{4}\sin^2\theta\cos^2\theta\left(\left(S_{10}-\tfrac{|S_{11}|^2}{M}\right)^2-\left|S_{12}-\tfrac{S_{11}^2}{M}\right|^2\right)\\
&\quad+\sin^4\theta\left(\tfrac{1}{2}\left(S_{20}-\tfrac{S_{23}^2}{M}\right)\left(S_{10}-\tfrac{|S_{11}|^2}{M}-\Re\left\{e^{-2j\phi}\left(S_{12}-\tfrac{S_{11}^2}{M}\right)\right\}\right)-\tfrac{1}{M^2}S_{23}^2\Im^2\{e^{-j\phi}S_{11}\}\right)\\
&\quad+\tfrac{S_{23}}{M}\sin^3\theta\cos\theta\\
&\quad\times\left(\Re\{e^{-j\phi}S_{11}\}\left(S_{10}-\tfrac{|S_{11}|^2}{M}-\Re\left\{e^{-2j\phi}\left(S_{12}-\tfrac{S_{11}^2}{M}\right)\right\}\right)\right.\\
&\quad\left.-\Im\{e^{-j\phi}S_{11}\}\Im\left\{e^{-2j\phi}\left(S_{12}-\tfrac{S_{11}^2}{M}\right)\right\}\right)\\
&= \tfrac{1}{4}\sin^2\theta\cos^2\theta\left(\left(S_{10}-\tfrac{|S_{11}|^2}{M}\right)^2-\left|S_{12}-\tfrac{S_{11}^2}{M}\right|^2\right)\\
&\quad+\sin^4\theta\left(\tfrac{1}{2}S_{20}\left(S_{10}-\tfrac{|S_{11}|^2}{M}-\Re\left\{e^{-2j\phi}\left(S_{12}-\tfrac{S_{11}^2}{M}\right)\right\}\right)-\tfrac{S_{23}^2}{2M}\left(S_{10}-\Re\{e^{-2j\phi}S_{12}\}\right)\right)\\
&\quad+\tfrac{S_{23}\sin^2\theta\sin 2\theta}{2M}\left(S_{10}\Re\{e^{-j\phi}S_{11}\}-\Re\{e^{-j\phi}S_{12}S_{11}^*\}\right).
\end{aligned}
\tag{48}
$$

C Proof of Eqn. (18)

Note that the sensors located on the xOy plane are such that $\xi_i = \frac{\pi}{2}$, while the sensors located on the first orthogonal axe have $\xi_i = 0$, and the sensors located on the second orthogonal axe are such that $\xi_i = \pi$. In the same way as we prove Eqn. (16) and (17), with the assumption that the two orthogonal branches are symmetric, it leads to:

$$
\sum_{i=1}^{N_2}\rho_{2,i}\cos\xi_i = \sum_{i=1}^{\frac{N_2}{2}}\rho_{2,i}\cos 0+\sum_{i=\frac{N_2}{2}+1}^{N_2}\rho_{2,i}\cos\pi = \sum_{i=1}^{\frac{N_2}{2}}\rho_{2,i}-\sum_{i=\frac{N_2}{2}+1}^{N_2}\rho_{2,i}=0. \tag{49}
$$

Finally it is easy obtain (18) from Eqn. (16) and (17) by letting $S_{23}=0$.

Acknowledgements This project was funded by both région Île-de-France and Digiteo Research Park

References

1. A. Manikas, Differential Geometry in Array Processing, Imperial College Press, 2004.
2. L. C. Godara, A. Cantoni, Uniqueness and linear independence of sterring vectors in array space, J. Acoust. Soc. Amer. 70 (2) (1981) 467– 475.
3. K. C. Tan, S. S. Goh, E. C. Tan, A study of the rank-ambiguilty issues in direction-of-arrival estimation, IEEE Trans. Signal Processing 44 (4) (1996) 880– 887.
4. J. T. H. Lo, S. L. M. Jr, Observability conditions for multiple signal direction finding and array sensor localization, IEEE Trans. Signal Processing 40 (11) (1992) 2641– 2650.
5. M. Gavish, A. J. Weiss, Array geometry for ambiguity resolution in direction finding, IEEE Trans. Antennas Propagat. 44 (6) (1991) 143– 146.
6. H. L. VanTrees, Detection, Estimation and Modulation Theory: Optimum Array Processing, Vol. 4, Wiley, New York, 2002.
7. P. Stoica, R. Moses, Spectral Analysis of Signals, Prentice Hall, NJ, 2005.
8. U. Baysal, R. L. Moses, On the geometry of isotropic arrays, IEEE Trans. Signal Processing 51 (6) (2003) 1469–1477.
9. D. Sengupta, T. Smith, R. Larson, Radiation characteristics of a spherical array of circularly polarized elements, IEEE Trans. Antennas Propagat. 16 (1) (1968) 2– 7.
10. P. Stoica, A. Nehorai, MUSIC, maximum likelihood and the Cramér-Rao bound, IEEE Trans. Acoust., Speech, Signal Processing 37 (1989) 720–741.
11. P. Stoica, A. Nehorai, Performances study of conditional and unconditional direction of arrival estimation, IEEE Trans. Acoust., Speech, Signal Processing 38 (1990) 1783–1795.
12. A. Renaux, P. Foster, E. Chaumette, P. Larzabal, On the High-SNR Conditional Maximum-Likelihood estimator full statistical characterization, IEEE Trans. Signal Processing 54 (12) (2006) 4840– 4843.
13. A. Renaux, P. Foster, E. Boyer, P. Larzabal, Unconditional Maximum Likelihood performance at finite number of samples and high Signal-to-Noise Ratio, IEEE Trans. Signal Processing 55 (5) (2007) 2358– 2364.
14. Y. Hua, T. K. Sarkar, A note on the Cramér-Rao bound for 2-D direction finding based on 2-D array, IEEE Trans. Signal Processing 39 (5) (1991) 1215– 1218.
15. R. O. Nielsen, Azimuth and elevation angle estimation with a three dimensional array, IEEE J. Oceanic Eng. 19 (1) (1994) 84–86.
16. A. Mirkin, L. H. Sibul, Cramér-Rao bounds on angle estimation with a two-dimensional array, IEEE Trans. Signal Processing 39 (1991) 515–517.
17. B. Yang, J. Scheuing, Cramér-Rao bound and optimum sensor array for source localization from the differences of arrival, in: Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, Vol. 4, Philadenphia, USA, 2005, pp. 961–964.
18. K. W. K. Lui, H. C. So, A study of two-dmensional sensor placement using time-difference-of-arrival measurements, Digital Signal Processing 19 (2009) 650– 659.
19. U. Oktel, R. L. Moses, A Bayesian approach to array geometry design, IEEE Trans. Signal Processing 53 (5) (2005) 1919– 1923.
20. H. Gazzah, S. Marcos, Cramér-Rao bounds for antenna array design, IEEE Trans. Signal Processing 54 (1) (2006) 336–345.
21. H. Gazzah, K. Abed-Meraim, Optimum ambiguity free directional and omni directional planar antenna arrays for DOA estimation, IEEE Trans. Signal Processing 57 (10) (2009) 3942–3953.
22. A. Ferréol, P. Chevalier, High resolution direction finding: from performance toward antenna array optimization -the mono-source case, in: Proc. European Signal Processing Conference, Glasgow, Scotland, 2009, pp. 1973–1977.
23. B. Ottersten, M. Viberg, P. Stoica, A. Nehorai, Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing, in: S. Haykin, J. Litva, T. J. Shepherd (Eds.), Radar Array Processing, Springer-Verlag, Berlin, 1993, Ch. 4, pp. 99–151.
24. J. Li, R. T. Compton, Maximum likelihood angle estimation for signals with known waveforms, IEEE Trans. Signal Processing 41 (1993) 2850–2862.
25. M. Cedervall, R. L. Moses, Efficient maximum likelihood DOA estimation for signals with known waveforms in presence of multipath, IEEE Trans. Signal Processing 45 (1997) 808–811.
26. J. Li, B. Halder, P. Stoica, M. Viberg, Computationally efficient angle estimation for signals with known waveforms, IEEE Trans. Signal Processing 43 (1995) 2154–2163.

27. A. Leshem, A.-J. van der Veen, Direction-of-arrival estimation for constant modulus signals, IEEE Trans. Signal Processing 47 (11) (1999) 3125 – 3129.
28. Y. H. Choi, Unified approach to Cramer-Rao bounds in direction estimation with known signal structures, Signal Processing 84 (10) (2004) 1875 – 1882.
29. Y. Hua, T. K. Sarkar, D. D. Weiner, An L-shaped array for estimating 2D directions of wave arrival, IEEE Trans. Antennas Propagat. 39 (1991) 143– 146.
30. T. Filik, T. E. Tuncer, Uniform and nonuniform V-shaped isotropic planar arrays, in: Proc. Sensor Array and Multichannel Signal Processing Workshop, Darmstadt, Germany, 2008, pp. 99–103.
31. S. M. Kay, Fundamentals of Statistical Signal Processing, Vol. 1, Prentice Hall, NJ, 1993.
32. A. V. D. Bos, A Cramer Rao lower bound for complex parameters, IEEE Trans. Acoust., Speech, Signal Processing 42 (1994) 2859.
33. B. Porat, B. Friedlander, Analysis of the asymptotic relative efficiency of the MUSIC algorithm, IEEE Trans. Acoust., Speech, Signal Processing 36 (4) (1988) 532–544.
34. A. T. Moffet, Minimum redundancy linear arrays, IEEE Trans. Antennas Propagat. 16 (1968) 172–175.
35. X. Huang, J. P. Reilly, M. Wong, Optimal design of linear array of sensors, in: Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, Vol. 2, Toronto, Ont., Canada, 1991, pp. 1405–1408.

Annexe H

Performance Bounds for the pulse phase estimation of X-ray pulsars

Soumis à *IEEE* Transactions on Signal Processing

Performance Bounds for The Pulse Phase Estimation of X-Ray Pulsars

Nguyen Duy TRAN, *Student Member, IEEE*, Alexandre RENAUX, *Member, IEEE*, Rémy BOYER, *Member, IEEE*, Sylvie MARCOS, Pascal LARZABAL, *Member, IEEE*.

Abstract

The use of X-ray pulsar signals appears to be a potential solution for autonomous deep space navigation. The main challenge in this kind of navigation is to estimate very precisely the initial phase of the pulse arriving at the detector. Previous studies indicate that in the performance of pulse phase estimators, the so-called threshold phenomenon arises when the observation time is below a critical limit. In this correspondence, to provide a prediction of the threshold position, the closed-form expressions of the lower bounds on the mean square error (MSE) are derived and analyzed in both deterministic and Bayesian contexts. Simulations show that the proposed bounds are able to predict the threshold location in both contexts.

Index Terms

X-ray Pulsar, QCL bound, Weiss-Weinstein bound , pulse phase estimation.

I. Introduction

The development of deep space operation requires accurate and autonomous navigation solutions for the purpose of orienting and controlling a spacecraft. The actual ground-based navigation is very accurate but highly depends on the communication with the ground station, and therefore, is not robust to a loss of contact. Besides, large errors can occur in shadowing areas or at large distance from the ground. While satellite navigation systems, such as the Global Positioning System (GPS), are helping devices operating inside the orbit of the GPS constellation to internally determine their location within a few meters or even less, a similar solution for spacecraft is still an open question. In this context, the celestial-based system that uses signals from celestial sources is a potential candidate for autonomous deep space navigation. Among various types of celestial sources, the pulsars, discovered in 1967, are the subset that emits highly regular, stable, and periodic signals. Their behavior has been observed from years,

Nguyen Duy TRAN, Alexandre RENAUX, Rémy BOYER, and Sylvie MARCOS are with University Paris-Sud 11, Laboratory of Signals and Systems, Supélec, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette cedex, France,phone: +331 6985 1763, fax: +331 6985 1765, (e-mail:{NguyenDuy.Tran,Renaux,Remy.Boyer,Marcos}@lss.supelec.fr)

Pascal LARZABAL is with Ecole Normale Supérieure de Cachan, SATIE Laboratory, 61 Avenue du Président Wilson, 94235 Cachan Cedex, France, phone: +331 4740 2116 (e-mail: larzabal@satie.ens-cachan.fr).

This project was funded by région Île-de-France, Digiteo Research Park and NEWCOM ++.

so the shape and period of their pulse profile are known very accurately. This property could be of the utmost interest for navigation objectives. Therefore, in this contribution, we focus on pulsars among other celestial sources. In the literature, two kinds of pulsars were examined for navigation purposes: sources that emit in the radio band and sources that emit in the X-ray band. We here consider the X-ray pulsars for their feasibility in implementation (thanks to the smaller sized detectors compared to those of radio band) and better accuracy [1].

In pulsar-based navigation, the observed signal is the pulse time-of-arrival (TOA) (or the pulse phase) at the detector. Processing this signal with respect to the recorded database gives us the specific information of the location of the spacecraft. The main problem in this kind of navigation is to estimate very precisely the pulse initial phase, and this challenge has been examined in the literature. In [2], the statistical model of the pulse TOA has been developed and the pulse phase estimation is investigated by deriving and analyzing the maximum likelihood estimator (MLE) and the Cramér-Rao bound (CRB). In [3], the nonlinear least-squares (NLS) estimator of the pulse phase is proposed and compared to the MLE in terms of computational complexity and mean square error (MSE) over the observation time. In both papers, one can observe, in terms of MSE performance, the so-called threshold phenomenon which appears as the observation time is below a critical limit. This can be explained by the distorted cost function used by estimators whose global maximum appears at a far point from the true value [4]. This threshold phenomenon is very similar to the one observed in the classical array processing context. Typically, in the classical array processing context, the threshold value can be predicted by using other bounds tighter than the CRB. These bounds on the MSE can be divided into two categories depending on the parameter assumptions [5]. When the unknown parameters are assumed to be deterministic, the so-called deterministic bounds that evaluate the ”locally best” behavior of the estimators have been proposed. The other category, the so-called Bayesian bounds, deals with the case where the parameters are assumed to be random, and particularly, they take into account the support of the parameters throughout an *a priori* pdf so that they can evaluate the ”globally best” performance. The advantage of Bayesian bounds over the deterministic bounds is their capability to give the fundamental limits of an estimator in terms of MSE over all the MSE range. However, the usefulness of the deterministic bound still remains when the parameter is deterministic, and for threshold prediction. For these reasons, in this paper, we study the performance limits in both assumptions on the parameter of interest.

Note that, in classical array processing, observations are typically modeled as Gaussian random variables, while in X-ray pulse phase estimation, observations are modeled with a Poisson distribution. To the best of our knowledge, there are very few results on lower bounds relevant to this kind of scenario. We can cite here the works in [6] and in [7] where the behavior of the CRB and a simple approximation of the Barankin bound are studied respectively in emission tomography. However, those articles do not consider the Bayesian case.

The rest of this paper is organized as follows: Section II presents the mathematical model of the X-ray pulsar signal, and also the likelihood of the observations. In Section III, we exploit the deterministic bound for the pulse phase and we give the closed-form expression of the Quinlan-Chaumette-Larzabal (QCL) bound [5] [8] since the QCL bound is one of the tightest deterministic bounds. In section IV, we derive the closed-form expression of the Weiss-Weinstein bound (WWB) [9] to analyze the behavior of the global MSE. Next, simulations are presented

in Section V to confirm the good ability of the derived bounds to predict the performance of the MLE. Finally, Section VI draws the conclusions.

II. X-ray Signal Model

In this Section, we give a brief background about the mathematical observation model provided and justified in [2]. This will lead to the likelihood function which will be the cornerstone of our analysis. Let us call k the number of photons detected at the detectors in a fixed time interval (a, b). The photon TOAs are modeled as a non-homogeneous Poisson process (NHPP) with a time-varying rate $\lambda(t) \geq 0$. This means that k follows a Poisson distribution $p(k;(a,b))$ with associated parameter $\int_a^b \lambda(t)dt$:

$$p(k;(a,b)) = \frac{\left[\int_a^b \lambda(t)dt\right]^k \exp\left[-\int_a^b \lambda(t)dt\right]}{k!} \tag{1}$$

The rate function $\lambda(t)$ denotes the aggregate rate of all photons arriving at the detector from the X-ray pulsar and background, expressed in photons per second (ph/s). In practice, the rate function $\lambda(t)$ has the following form:

$$\lambda(t) = \lambda_b + \lambda_s h(\phi_{obs}(t)) \qquad \text{(ph/s)} \tag{2}$$

where λ_s and λ_b are called the *effective source rate* and *effective background arrival rate*, respectively; $h(\phi(t))$ is the normalized pulse profile function, and $\phi_{obs}(t)$ is the phase observed at the detector. Note that, thanks to the database obtained from years, the shape and period of the pulse profile are known very accurately [10]. The pulse profile function $h(\phi(t))$ is defined as a periodic function with its period equal to one cycle, i.e., $h(\phi(t))$ is defined on the interval $\phi \in [0, 1)$, and we have $h(m + \phi) = h(\phi(t))$ for all integers m. Besides, the function $h(\phi(t))$ is normalized, i.e, $\int_0^1 h(\phi)d\phi = 1$, and $\min_\phi h(\phi(t)) = 0$.

The observed phase at the detector is given by $\phi_{obs}(t) = \phi_0 + \int_{t_0}^t f(\tau)d\tau$, where ϕ_0 is the initial phase, where t_0 is the start of the observation interval, and where $f(t)$ is the observed signal frequency which depends on the constant source frequency and the variant Doppler frequency shift. Note that, in this paper, we concentrate on the initial phase estimation problem, then, we assume that the observed frequency is a known constant. This is the constant-frequency model as in [3] where the observed phase at the detector can be rewritten as $\phi_{obs} = \phi_0 + (t - t_0)f$. The Poisson rate function can, now, be considered as a function of the only unknown parameter, the initial phase, as below $\lambda(t;\phi_0) = \lambda_b + \lambda_s h(\phi_0 + (t - t_0)f)$. Since λ_b and λ_s are known from the database, then, the remaining challenge here is to estimate the initial phase ϕ_0. This is what has been done in [2] and [3] where two estimators, the MLE and the NLS, are studied and their performance has been compared to the CRB in terms of MSE. In this work, other bounds, such as QCL and WWB, are exploited to have a better benchmark. For this reason, we derive below the likelihood function.

The observation interval $(t_0, t_0 + T_{obs})$ is partitioned into N equal-length segments. We define x_n, $n = 0, 1, \ldots, N - 1$, as the number of photons detected in the n-th segment, and $\Delta t \equiv T_{obs}/N$ as the segment size. If N is large

enough, the Poisson rate $\lambda(t,\phi_0)$ can be assumed constant in the n-th segment, i.e. $\lambda_n(\phi_0)=\lambda(t_n;\phi_0)$, where $t_n=t_0+n\Delta t$. The probability mass function (pmf) for each Poisson random variable x_n, $n=0,1,\ldots,N-1$, can be written as: $p(x_n=x;\phi_0)=\frac{[\lambda_n(\phi_0)\Delta t]^x}{x!}\exp(-\lambda_n(\phi_0)\Delta t)$, where x is a non-negative integer. Under the assumption of independent observations, the likelihood of the full set of observations $\mathbf{x}=[x_0,x_1,\ldots,x_N]$ is given by

$$p(\mathbf{x};\phi_0)=\prod_{n=0}^{N-1}\frac{[\lambda_n(\phi_0)\Delta t]^{x_n}}{x_n!}\exp(-\lambda_n(\phi_0)\Delta t). \tag{3}$$

In Section III and Section IV, we will derive the lower bounds based on (3) in the deterministic context and Bayesian context, respectively.

III. Deterministic Bound

In this Section, we consider the so-called deterministic bounds for pulse phase estimation. Mathematically, the Barankin bound (BB) [11] is known to be tighter than the CRB, however, it is not computable. In classical array processing, to obtain a computable BB, several approximations of BB were proposed [5]. Consequently, in this paper, we derive the QCL bound [5] which is one of the tighest bounds among the Barankin family. This approximation is obtained following the search of an optimum over a set of test points, denoted as $[\theta_0,\ldots,\theta_{N-1}]$.

The Nth-order QCL bound of the unknown parameter ϕ_0 satisfies the following relation $E_{\mathbf{x};\phi_0}\left[(\hat{\phi}-\phi_0)^2\right]\geq B^N_{QCL}(\phi_0)$, where $E_{\mathbf{x};\phi_0}\left[(\hat{\phi}-\phi_0)^2\right]=\sum_{x_0=0}^{\infty}\cdots\sum_{x_{N-1}=0}^{\infty}(\hat{\phi}-\phi_0)^2p(\mathbf{x};\phi_0)$ is the variance of any unbiased estimators $\hat{\phi}$ of ϕ_0. Hereafter, we use, for simplicity, $\sum_{\mathbf{x}=0}^{\infty}$ instead of $\sum_{x_0=0}^{\infty}\cdots\sum_{x_{N-1}=0}^{\infty}$.

The bound B^N_{QCL} is calculated as follows [5]: $B^N_{QCL}=\mathbf{v}^T\mathbf{M}^{-1}_{QCL}\mathbf{v}$ where

$$\begin{cases}\mathbf{v}=[\mathbf{\Phi}^T,1,\ldots,1]^T\in\mathbb{R}^{2N\times 1} \text{ where } \mathbf{\Phi}=[\xi_0\ldots\xi_{N-1}]^T, \text{ where } \xi_n=\theta_n-\phi_0,\ n=0\ldots N-1,\\ \mathbf{M}_{QCL}=\begin{bmatrix}\mathbf{M}_{MS} & \mathbf{H}^T\\ \mathbf{H} & \mathbf{M}_{EFI}\end{bmatrix}\in\mathbb{R}^{2N\times 2N}\end{cases} \tag{4}$$

where

$$\mathbf{M}_{MS}=E_{\mathbf{x};\phi_0}\left[\begin{pmatrix}\frac{p(\mathbf{x};\theta_0)}{p(\mathbf{x};\phi_0)}\\ \vdots\\ \frac{p(\mathbf{x};\theta_{N-1})}{p(\mathbf{x};\phi_0)}\end{pmatrix}\begin{pmatrix}\frac{p(\mathbf{x};\theta_0)}{p(\mathbf{x};\phi_0)}\\ \vdots\\ \frac{p(\mathbf{x};\theta_{N-1})}{p(\mathbf{x};\phi_0)}\end{pmatrix}^T\right], \tag{5}$$

$$\mathbf{M}_{EFI}=E_{\mathbf{x};\phi_0}\left[\begin{pmatrix}\frac{\partial\ln p(\mathbf{x};\theta_0)}{\partial\theta_0}\frac{p(\mathbf{x};\theta_0)}{p(\mathbf{x};\phi_0)}\\ \vdots\\ \frac{\partial\ln p(\mathbf{x};\theta_{N-1})}{\partial\theta_{N-1}}\frac{p(\mathbf{x};\theta_{N-1})}{p(\mathbf{x};\phi_0)}\end{pmatrix}\begin{pmatrix}\frac{\partial\ln p(\mathbf{x};\theta_0)}{\partial\theta_0}\frac{p(\mathbf{x};\theta_0)}{p(\mathbf{x};\phi_0)}\\ \vdots\\ \frac{\partial\ln p(\mathbf{x};\theta_{N-1})}{\partial\theta_{N-1}}\frac{p(\mathbf{x};\theta_{N-1})}{p(\mathbf{x};\phi_0)}\end{pmatrix}^T\right], \tag{6}$$

$$\mathbf{H}=E_{\mathbf{x};\phi_0}\left[\begin{pmatrix}\frac{\partial\ln p(\mathbf{x};\theta_0)}{\partial\theta_0}\frac{p(\mathbf{x};\theta_0)}{p(\mathbf{x};\phi_0)}\\ \vdots\\ \frac{\partial\ln p(\mathbf{x};\theta_{N-1})}{\partial\theta_{N-1}}\frac{p(\mathbf{x};\theta_{N-1})}{p(\mathbf{x};\phi_0)}\end{pmatrix}\begin{pmatrix}\frac{p(\mathbf{x};\theta_0)}{p(\mathbf{x};\phi_0)}\\ \vdots\\ \frac{p(\mathbf{x};\theta_{N-1})}{p(\mathbf{x};\phi_0)}\end{pmatrix}^T\right]. \tag{7}$$

The set $\theta_n, n = 1, \ldots, N-1$ is the so-called set of the test point. After the calculation which is detailed in Appendix one obtains the closed-form expressions of elements (k, l) of matrix $\mathbf{M}_{MS}$ (see Appendix VII-A), $\mathbf{M}_{EFI}$ (see Appendix VII-B), and $\mathbf{H}$ (see Appendix VII-C) as follows:

$$\mathbf{M}_{MS}(k,l) = \exp\left\{ T_{obs} \int_0^1 \lambda(\phi) - \lambda(\xi_k + \phi) - \lambda(\xi_l + \phi) + \frac{\lambda(\xi_k+\phi)\lambda(\xi_l+\phi)}{\lambda(\phi)} d\phi \right\}, \tag{8}$$

$$\begin{aligned}\mathbf{M}_{EFI}(k,l) = \mathbf{M}_{MS}(k,l) \Bigg[& T_{obs}^2 \int_0^1 \frac{\partial\lambda(\phi+\xi_k)}{\partial\xi_k} \frac{\lambda(\phi+\xi_l)}{\lambda(\phi)} d\phi \int_0^1 \frac{\partial\lambda(\phi+\xi_l)}{\partial\xi_l} \frac{\lambda(\phi+\xi_k)}{\lambda(\phi)} d\phi \\ & + T_{obs}^2 \int_0^1 \frac{\partial\lambda(\phi+\xi_k)}{\partial\xi_k} d\phi \int_0^1 \frac{\partial\lambda(\phi+\xi_l)}{\partial\xi_l} d\phi - T_{obs}^2 \int_0^1 \frac{\partial\lambda(\phi+\xi_l)}{\partial\xi_l} d\phi \int_0^1 \frac{\partial\lambda(\phi+\xi_k)}{\partial\xi_k} \frac{\lambda(\phi+\xi_l)}{\lambda(\phi)} d\phi, \\ & - T_{obs}^2 \int_0^1 \frac{\partial\lambda(\phi+\xi_k)}{\partial\xi_k} d\phi \int_0^1 \frac{\partial\lambda(\phi+\xi_l)}{\partial\xi_l} \frac{\lambda(\phi+\xi_k)}{\lambda(\phi)} d\phi + T_{obs} \int_0^1 \frac{\partial\lambda(\phi+\xi_k)}{\partial\xi_k} \frac{\partial\lambda(\phi+\xi_l)}{\partial\xi_l} \frac{1}{\lambda(\phi)} d\phi \Bigg]\end{aligned} \tag{9}$$

and

$$\mathbf{H}(k,l) = \mathbf{M}_{MS}(k,l) \left[T_{obs} \int_0^1 \frac{\partial\lambda(\phi+\xi_k)}{\partial\xi_k} \frac{\lambda(\phi+\xi_l)}{\lambda(\phi)} d\phi - T_{obs} \int_0^1 \frac{\partial\lambda(\phi+\xi_k)}{\partial\xi_k} d\phi \right]. \tag{10}$$

In the expression of the QCL bound, we can see the existence of an integral which can be computed easily and rapidly in a numerical way. Note that, it was also the case in the CRB calculus proposed in [2] and [3]. In Section V, we plot the QCL bound versus several observation times and compare it to the CRB and the MSE of the MLE.

IV. Bayesian bound

As an alternative to the deterministic framework, we propose to handle the problem in the Bayesian framework which will provide a tight minimal bound over all the range of observation time and a good prediction of the observation time threshold. In particular, we assume that the parameter of interest ϕ_0 is random with an *a priori* uniform probability density function (pdf) over the support $[0, 1)$. Note that not all the Bayesian bounds proposed in the literature are able to take into account the case when the parameters of interest are supposed to be uniformly distributed. Therefore, among various types of Bayesian bounds [12], we concentrate, in this Section, on the Weiss-Weinstein bound (WWB) (see [9] [13] [14]), which can deal with the uniformly distributed prior assumption and is one of the tightest bound of the Weiss and Weinstein family [15] [16].

The Weiss-Weinstein bound, denoted WWB, for the unknown parameter ϕ_0 satisfies the following relation $E_{\mathbf{x};\phi_0}\left[(\hat{\phi} - \phi_0)^2\right] \geq WWB$, where $E_{\mathbf{x};\phi_0}\left[(\hat{\phi} - \phi_0)^2\right] = \int_{\mathbf{\Theta}} \sum_{\mathbf{x}=0}^{\infty} (\hat{\phi} - \phi_0)^2 p(\mathbf{x}, \phi_0) d\phi_0$ is the variance of any estimators of ϕ_0 where $p(\mathbf{x}, \phi_0)$ being the joint pdf. Note that, contrary to the deterministic bounds, no assumption is made on the estimator $\hat{\phi}$, e.g., $\hat{\phi}$ can be biased. The WWB is calculated by [9]

$$WWB = \underset{u,s}{sup} \frac{u^2 \exp(2\eta(s,u))}{\exp(\eta(2s,u)) + \exp(\eta(2-2s,-u)) - 2\exp(\eta(s,2u))}, \tag{11}$$

where $s \in [0,1]$, where u is the test point chosen such that $\phi_0 + u \in [0,1)$, and $\eta(\alpha,\beta)$ is defined by

$$\begin{aligned}\eta(\alpha,\beta) &= \ln \int_{\boldsymbol{\Theta}} \sum_{\mathbf{x}=0}^{\infty} p(\mathbf{x},\phi_0+\beta)^{\alpha} p(\mathbf{x},\phi_0)^{1-\alpha} d\phi_0 \\ &= \ln \int_{\boldsymbol{\Theta}} \sum_{\mathbf{x}=0}^{\infty} p(\mathbf{x};\phi_0+\beta)^{\alpha} p(\phi_0+\beta)^{\alpha} p(\mathbf{x};\phi_0)^{1-\alpha} p(\phi_0)^{1-\alpha} d\phi_0 \\ &= \ln \int_{\boldsymbol{\Theta}} \eta'(\alpha,\beta) p(\phi_0+\beta)^{\alpha} p(\phi_0)^{1-\alpha} d\phi_0, \end{aligned} \tag{12}$$

where we define $\eta'(\alpha,\beta) = \sum_{\mathbf{x}=0}^{\infty} p(\mathbf{x};\phi_0+\beta)^{\alpha} p(\mathbf{x};\phi_0)^{1-\alpha}$. The closed-form expression of $\eta'(\alpha,\beta)$ is given by (see Appendix VII-D for details)

$$\eta'(\alpha,\beta) = \exp\left\{ T_{obs} \int_0^1 \left(-\alpha\lambda(\phi+\beta) - (1-\alpha)\lambda(\phi) + \lambda(\phi+\beta)^{\alpha}\lambda(\phi)^{1-\alpha} \right) d\phi \right\}. \tag{13}$$

Note that, the dependance of $\eta'(\alpha,\beta)$ on ϕ_0 is now removed, then, the integral in (12) can be calculated easily. Finally, the Weiss-Weinstein bound for the unknown parameter ϕ_0 is given by

$$WWB = \sup_{u,s} \frac{u^2(1-u)^2\eta'^2(s,u)}{(1-u)\eta'(2s,u) + (1-u)\eta'(2-2s,-u) - 2(1-2u)\eta'(s,2u)}. \tag{14}$$

As it appears in the QCL bound [see (8)-(10)], integrals also exist in the expression of the WWB but they can be numerically integrated.

V. Numerical Results

To evaluate the proposed bounds, we compare them to the performance of the MLE of the pulse phase which is given by $\hat{\phi} = \arg\max_{\phi_0 \in \boldsymbol{\Theta}} \sum_{n=0}^{N-1} [x_n \ln[\lambda_n(\phi_0)\Delta t] - \lambda_n(\phi_0)\Delta t]$. The performance of the MLE is simulated using 1000 Monte Carlo runs. The pulsar period is 33.5 ms. The photon rates are $\lambda_b = 5$ (ph/s) and $\lambda_s = 15$ (ph/s). As the phase is defined on the $[0,1)$ interval, the phase error value is calculated modulo one cycle, i.e., $min[mod(\phi_0 - \hat{\phi},1), mod(\hat{\phi} - \phi_0,1)]$. For example, the error between 0.9 cycle and 0.1 cycle should be 0.2 cycle, and not 0.8 cycle.

In Figure 1, we compare the QCL bound and the MS bound to the CRB and to the MSE of the MLE versus the observation time. The initial phase, $\phi_0 = 0.9$ cycle, is chosen arbitrarily. It can be seen that the QCL provides a good prediction of the threshold location compared to the MSE of the MLE.

Figure 2 shows the WWB, the CRB and the empirical global MSE of the MLE of ϕ_0 versus the observation time. It can be seen that the WWB predicts well not only the threshold position, but also the MSE of the MLE in all range of observation time (asymptotic and threshold regions). Note that the tightest WWB is achieved when $s = 1/2$.

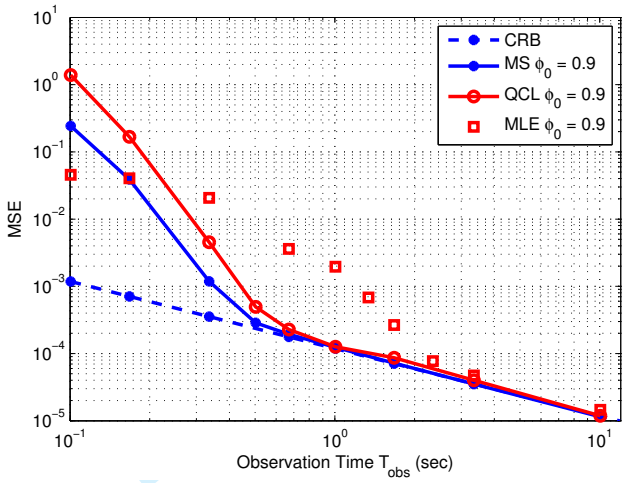


Fig. 1. QCL bound, MS bound, CRB and empirical MSE of the MLE of ϕ_0 versus observation time

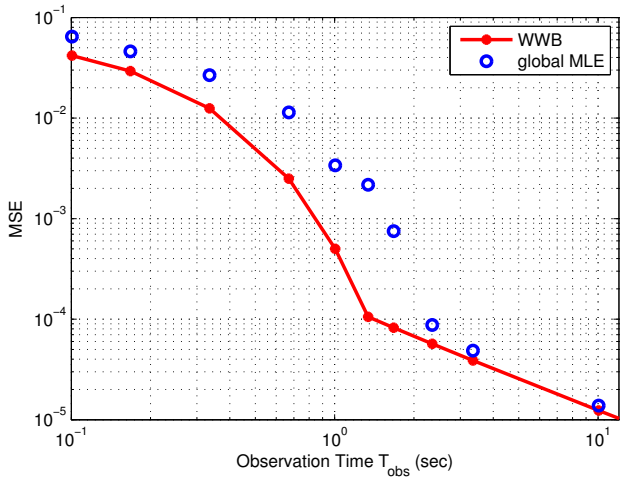


Fig. 2. CRB and empirical global MSE of MLE of ϕ_0 versus observation time

VI. Conclusions

To estimate accurately the initial phase of the pulse arriving at the detector appears to be the key challenge in a system using the X-ray pulsar to perform the autonomous deep space navigating. Therefore, we have derived the closed-form expressions of the lower bounds on the MSE and analyzed their behavior for the problem of X-ray pulse

phase estimation. Indeed, both deterministic (QCL bound) and Bayesian bounds (WWB) have been considered. We have shown that both types of lower bound provide good prediction of the threshold location depending on the estimation framework.

VII. Appendix

A. Derivation of $\mathbf{M}_{MS}(k,l)$

From (5), the element (k,l) of the matrix $\mathbf{M}_{MS}$ is given by

$$\begin{aligned}\mathbf{M}_{MS}(k,l) &= E_{\mathbf{x};\phi_0}\left[\frac{p(\mathbf{x};\theta_k)}{p(\mathbf{x};\phi_0)}\frac{p(\mathbf{x};\theta_l)}{p(\mathbf{x};\phi_0)}\right] \\ &= \sum_{\mathbf{x}=0}^{\infty}\prod_{n=0}^{N-1}\left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x_n}\frac{1}{x_n!}\exp\left(\Delta t[\lambda_n(\phi_0)-\lambda_n(\theta_k)-\lambda_n(\theta_l)]\right) \quad (15)\\ &= \exp\left\{\Delta t\sum_{n=0}^{N-1}\lambda_n(\phi_0)-\lambda_n(\theta_k)-\lambda_n(\theta_l)+\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)}{\lambda_n(\phi_0)}\right\}. \quad (16)\end{aligned}$$

By taking the limit $\Delta t \to 0$, or equivalently $N \to \infty$, and by noting that $\lambda_n(.) = \lambda(t_n;.)$, we can convert the above summation to an integral:

$$\begin{aligned}\mathbf{M}_{MS}(k,l) &= \exp\left\{\int_{t_0}^{t_0+T_{obs}}\left[\lambda(t,\phi_0)-\lambda(t,\theta_k)-\lambda(t,\theta_l)+\frac{\lambda(t,\theta_k)\lambda(t,\theta_l)}{\lambda(t,\phi_0)}\right]dt\right\} \\ &= \exp\left\{\int_{\phi_0}^{\phi_0+fT_{obs}}\frac{1}{f}\left[\lambda(\phi)-\lambda(\phi+\xi_k)-\lambda(\phi+\xi_l)+\frac{\lambda(\phi+\xi_k)\lambda(\phi+\xi_l)}{\lambda(\phi)}\right]d\phi\right\} \\ &= \exp\left\{\int_{0}^{fT_{obs}}\frac{1}{f}\left[\lambda(\phi)-\lambda(\phi+\xi_k)-\lambda(\phi+\xi_l)+\frac{\lambda(\phi+\xi_k)\lambda(\phi+\xi_l)}{\lambda(\phi)}\right]d\phi\right\} \quad (17)\end{aligned}$$

where $\lambda(\phi+\alpha) = \lambda_b + \lambda_s h(\phi+\alpha)$.

When the observation time is an integer number of the pulsar period, i.e., $fT_{obs} \approx N_p$ (cycle), $\mathbf{M}_{MS}(k,l)$ can be rewritten as

$$\begin{aligned}\mathbf{M}_{MS}(k,l) &= \exp\left\{\sum_{n=1}^{N_p}\int_{n-1}^{n}\frac{1}{f}\left[\lambda(\phi)-\lambda(\xi_k+\phi)-\lambda(\xi_l+\phi)+\frac{\lambda(\xi_k+\phi)\lambda(\xi_l+\phi)}{\lambda(\phi)}\right]d\phi\right\} \\ &= \exp\left\{N_p\int_{0}^{1}\frac{1}{f}\left[\lambda(\phi)-\lambda(\xi_k+\phi)-\lambda(\xi_l+\phi)+\frac{\lambda(\xi_k+\phi)\lambda(\xi_l+\phi)}{\lambda(\phi)}\right]d\phi\right\} \\ &= \exp\left\{T_{obs}\int_{0}^{1}\left[\lambda(\phi)-\lambda(\xi_k+\phi)-\lambda(\xi_l+\phi)+\frac{\lambda(\xi_k+\phi)\lambda(\xi_l+\phi)}{\lambda(\phi)}\right]d\phi\right\}. \quad (18)\end{aligned}$$

B. Derivation of $\mathbf{M}_{EFI}(k,l)$

From (6), the element (k,l) of the matrix $\mathbf{M}_{EFI}$ is given by

$$\begin{aligned}\mathbf{M}_{EFI}(k,l) &= E_{\mathbf{x};\phi_0}\left[\frac{\partial \ln p(\mathbf{x};\theta_k)}{\partial\theta_k}\frac{p(\mathbf{x};\theta_k)}{p(\mathbf{x};\phi_0)}\frac{\partial \ln p(\mathbf{x};\theta_l)}{\partial\theta_l}\frac{p(\mathbf{x};\theta_l)}{p(\mathbf{x};\phi_0)}\right]\\ &= \sum_{\mathbf{x}=0}^{\infty}\left[\sum_{p=0}^{N-1} x_p\frac{\partial \ln[\lambda_p(\theta_k)\Delta t]}{\partial\theta_k} - \frac{\partial\lambda_p(\theta_k)\Delta t}{\partial\theta_k}\right]\left[\sum_{q=0}^{N-1} x_q\frac{\partial \ln[\lambda_q(\theta_k)\Delta t]}{\partial\theta_k} - \frac{\partial\lambda_q(\theta_k)\Delta t}{\partial\theta_k}\right]\\ &\qquad \prod_{n=0}^{N-1}\left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x_n}\frac{1}{x_n!}\exp\left(\Delta t[\lambda_n(\phi_0)-\lambda_n(\theta_k)-\lambda_n(\theta_l)]\right)\\ &= C_1 + C_2 + C_3 + C_4, \end{aligned} \tag{19}$$

whose components C_i, $i = 1,2,3,4$, are calculated as follows

$$\begin{aligned}C_1 &= \sum_{\mathbf{x}=0}^{\infty}\sum_{p=0}^{N-1}\sum_{q=0}^{N-1} x_p x_q\frac{\partial \ln[\lambda_p(\theta_k)\Delta t]}{\partial\theta_k}\frac{\partial \ln[\lambda_q(\theta_l)\Delta t]}{\partial\theta_l}\prod_{n=0}^{N-1}\left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x_n}\frac{1}{x_n!}\exp\left(\Delta t[\lambda_n(\phi_0)-\lambda_n(\theta_k)-\lambda_n(\theta_l)]\right)\\ &= \sum_{p=0}^{N-1}\sum_{q=0}^{N-1}\frac{\partial \ln[\lambda_p(\theta_k)\Delta t]}{\partial\theta_k}\frac{\partial \ln[\lambda_q(\theta_k)\Delta t]}{\partial\theta_k}\sum_{\mathbf{x}=0}^{\infty} x_p x_q\prod_{n=0}^{N-1}\left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x_n}\frac{1}{x_n!}\exp\left(\Delta t[\lambda_n(\phi_0)-\lambda_n(\theta_k)-\lambda_n(\theta_l)]\right).\end{aligned} \tag{20}$$

Let us define

$$\begin{aligned}A &= \sum_{\mathbf{x}=0}^{\infty} x_p x_q\prod_{n=0}^{N-1}\left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x_n}\frac{1}{x_n!}\exp\left(\Delta t[\lambda_n(\phi_0)-\lambda_n(\theta_k)-\lambda_n(\theta_l)]\right)\\ &= A(p=q) + A(p\neq q),\end{aligned} \tag{21}$$

where

$$\begin{aligned}A(p=q) &= \sum_{\mathbf{x}=0}^{\infty} x_p^2\prod_{n=0}^{N-1}\left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x_n}\frac{1}{x_n!}\exp\left(\Delta t[\lambda_n(\phi_0)-\lambda_n(\theta_k)-\lambda_n(\theta_l)]\right)\\ &= \sum_{\mathbf{x}=0}^{\infty} x_p(\mathbf{x}_p-1)\prod_{n=0}^{N-1}\left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x_n}\frac{1}{x_n!}\exp\left(\Delta t[\lambda_n(\phi_0)-\lambda_n(\theta_k)-\lambda_n(\theta_l)]\right)\\ &\quad + \sum_{\mathbf{x}=0}^{\infty} x_p\prod_{m=0}^{N-1}\left[\frac{\lambda_m(\theta_k)\lambda_m(\theta_l)\Delta t}{\lambda_m(\phi_0)}\right]^{x_m}\frac{1}{x_m!}\exp\left(\Delta t[\lambda_m(\phi_0)-\lambda_m(\theta_k)-\lambda_m(\theta_l)]\right)\\ &= \sum_{\mathbf{x}'=0}^{\infty}\left[\frac{\lambda_p(\theta_k)\lambda_p(\theta_l)\Delta t}{\lambda_p(\phi_0)}\right]^2\prod_{n=0}^{N-1}\left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x'_n}\frac{1}{x'_n!}\exp\left(\Delta t[\lambda_n(\phi_0)-\lambda_n(\theta_k)-\lambda_n(\theta_l)]\right)\\ &\quad + \sum_{\mathbf{x}''=0}^{\infty}\left[\frac{\lambda_p(\theta_k)\lambda_p(\theta_l)\Delta t}{\lambda_p(\phi_0)}\right]\prod_{m=0}^{N-1}\left[\frac{\lambda_m(\theta_k)\lambda_m(\theta_l)\Delta t}{\lambda_m(\phi_0)}\right]^{x''_m}\frac{1}{x''_m!}\exp\left(\Delta t[\lambda_m(\phi_0)-\lambda_m(\theta_k)-\lambda_m(\theta_l)]\right)\\ &= \left[\frac{\lambda_p(\theta_k)\lambda_p(\theta_l)\Delta t}{\lambda_p(\phi_0)}\right]^2\mathbf{M}_{MS}(k,l) + \frac{\lambda_p(\theta_k)\lambda_p(\theta_l)\Delta t}{\lambda_p(\phi_0)}\mathbf{M}_{MS}(k,l),\end{aligned} \tag{22}$$

where $\mathbf{x}' = [x_1, \ldots, x_p - 2, \ldots, x_{N-1}]$, and $\mathbf{x}'' = [x_1, \ldots, x_p - 1, \ldots, x_{N-1}]$, and where

$$\begin{aligned} A(p \neq q) &= \sum_{\mathbf{x}=0}^{\infty} x_p x_q \prod_{n=0}^{N-1} \left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x_n} \frac{1}{x_n!} \exp\left(\Delta t[\lambda_n(\phi_0) - \lambda_n(\theta_k) - \lambda_n(\theta_l)]\right) \\ &= \frac{\lambda_p(\theta_k)\lambda_p(\theta_l)\Delta t}{\lambda_p(\phi_0)} \frac{\lambda_q(\theta_k)\lambda_q(\theta_l)\Delta t}{\lambda_q(\phi_0)} \sum_{\mathbf{x}=0}^{\infty} \prod_{n=0}^{N-1} \left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x_n'''} \frac{1}{x_n'''!} \exp\left(\Delta t[\lambda_n(\phi_0) - \lambda_n(\theta_k) - \lambda_n(\theta_l)]\right) \\ &= \frac{\lambda_p(\theta_k)\lambda_p(\theta_l)\Delta t}{\lambda_p(\phi_0)} \frac{\lambda_q(\theta_k)\lambda_q(\theta_l)\Delta t}{\lambda_q(\phi_0)} \mathbf{M}_{MS}(k,l), \end{aligned} \tag{23}$$

where $\mathbf{x}''' = [x_1, \ldots, x_p - 1, \ldots, x_q - 1, \ldots, x_{N-1}]$. Note that in the above derivation, we assumed , without loss of generality, that $p < q$. Consequently, we get

$$\begin{aligned} C_1 &= \mathbf{M}_{MS}(k,l) \Bigg\{ \sum_{p=0}^{N-1} \sum_{q=0}^{N-1} \frac{\partial \ln[\lambda_p(\theta_k)\Delta t]}{\partial \theta_k} \frac{\partial \ln[\lambda_q(\theta_l)\Delta t]}{\partial \theta_l} \frac{\lambda_p(\theta_k)\lambda_p(\theta_l)\Delta t}{\lambda_p(\phi_0)} \frac{\lambda_q(\theta_k)\lambda_q(\theta_l)\Delta t}{\lambda_q(\phi_0)} \\ &\qquad + \sum_{p=0}^{N-1} \frac{\partial \ln[\lambda_p(\theta_k)\Delta t]}{\partial \theta_k} \frac{\partial \ln[\lambda_p(\theta_l)\Delta t]}{\partial \theta_l} \frac{\lambda_p(\theta_k)\lambda_p(\theta_l)\Delta t}{\lambda_p(\phi_0)} \Bigg\} \\ &= \mathbf{M}_{MS}(k,l) \Bigg\{ \sum_{p=0}^{N-1} \frac{\partial \lambda_p(\theta_k)}{\partial \theta_k} \frac{\lambda_p(\theta_l)}{\lambda_p(\phi_0)} \Delta t \sum_{q=0}^{N-1} \frac{\partial \lambda_q(\theta_l)}{\partial \theta_l} \frac{\lambda_q(\theta_k)}{\lambda_q(\phi_0)} \Delta t + \sum_{p=0}^{N-1} \frac{\partial \lambda_p(\theta_k)}{\partial \theta_k} \frac{\partial \lambda_p(\theta_l)}{\partial \theta_l} \frac{1}{\lambda_p(\phi_0)} \Delta t \Bigg\}. \end{aligned} \tag{24}$$

Using again the same calculating technique as in the derivation of the MS matrix, we get

$$C_1 = \mathbf{M}_{MS}(k,l) \left[T_{obs}^2 \int_0^1 \frac{\partial \lambda(\phi + \xi_k)}{\partial \xi_k} \frac{\lambda(\phi + \xi_l)}{\lambda(\phi)} d\phi \int_0^1 \frac{\partial \lambda(\phi + \xi_l)}{\partial \xi_l} \frac{\lambda(\phi + \xi_k)}{\lambda(\phi)} d\phi + T_{obs} \int_0^1 \frac{\partial \lambda(\phi + \xi_k)}{\partial \xi_k} \frac{\partial \lambda(\phi + \xi_l)}{\partial \xi_l} \frac{1}{\lambda(\phi)} d\phi \right]. \tag{25}$$

Similarly, we derive the other components

$$\begin{aligned} C_2 &= \sum_{\mathbf{x}=0}^{\infty} \sum_{p=0}^{N-1} \sum_{q=0}^{N-1} \frac{\partial \lambda_p(\theta_k)\Delta t}{\partial \theta_k} \frac{\partial \lambda_q(\theta_l)\Delta t}{\partial \theta_l} \prod_{n=0}^{N-1} \left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x_n} \frac{1}{x_n!} \exp\left(\Delta t[\lambda_n(\phi_0) - \lambda_n(\theta_k) - \lambda_n(\theta_l)]\right) \\ &= \mathbf{M}_{MS}(k,l) \sum_{p=0}^{N-1} \sum_{q=0}^{N-1} \frac{\partial \lambda_p(\theta_k)\Delta t}{\partial \theta_k} \frac{\partial \lambda_q(\theta_l)\Delta t}{\partial \theta_l} \\ &= \mathbf{M}_{MS}(k,l) T_{obs}^2 \int_0^1 \frac{\partial \lambda(\phi + \xi_k)}{\partial \xi_k} d\phi \int_0^1 \frac{\partial \lambda(\phi + \xi_l)}{\partial \xi_l} d\phi, \end{aligned} \tag{26}$$

$$\begin{aligned} C_3 &= \sum_{\mathbf{x}=0}^{\infty} x_p \sum_{p=0}^{N-1} \sum_{q=0}^{N-1} \frac{\partial \ln[\lambda_p(\theta_k)\Delta t]}{\partial \theta_k} \frac{\partial \lambda_q(\theta_l)\Delta t}{\partial \theta_l} \prod_{n=0}^{N-1} \left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x_n} \frac{1}{x_n!} \exp\left(\Delta t[\lambda_n(\phi_0) - \lambda_n(\theta_k) - \lambda_n(\theta_l)]\right) \\ &= \sum_{q=0}^{N-1} \frac{\partial \lambda_q(\theta_l)\Delta t}{\partial \theta_l} \sum_{p=0}^{N-1} \frac{\partial \ln[\lambda_p(\theta_k)\Delta t]}{\partial \theta_k} \frac{\lambda_p(\theta_k)\lambda_p(\theta_l)\Delta t}{\lambda_p(\phi_0)} \mathbf{M}_{MS}(k,l) \\ &= \mathbf{M}_{MS}(k,l) T_{obs}^2 \int_0^1 \frac{\partial \lambda(\phi + \xi_l)}{\partial \xi_l} d\phi \int_0^1 \frac{\partial \lambda(\phi + \xi_k)}{\partial \xi_k} \frac{\lambda(\phi + \xi_l)}{\lambda(\phi)} d\phi, \end{aligned} \tag{27}$$

$$C_4 = \sum_{\mathbf{x}=0}^{\infty} x_q \sum_{p=0}^{N-1}\sum_{q=0}^{N-1} \frac{\partial\lambda_p(\theta_k)\Delta t}{\partial\theta_k}\frac{\partial \ln[\lambda_q(\theta_l)\Delta t]}{\partial\theta_l} \prod_{n=0}^{N-1}\left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x_n}\frac{1}{x_n!}\exp\left(\Delta t[\lambda_n(\phi_0)-\lambda_n(\theta_k)-\lambda_n(\theta_l)]\right)$$
$$= \sum_{p=0}^{N-1}\frac{\partial\lambda_p(\theta_k)\Delta t}{\partial\theta_k}\sum_{q=0}^{N-1}\frac{\partial \ln[\lambda_q(\theta_l)\Delta t]}{\partial\theta_l}\frac{\lambda_q(\theta_k)\lambda_q(\theta_l)\Delta t}{\lambda_q(\phi_0)}\mathbf{M}_{MS}(k,l)$$
$$= \mathbf{M}_{MS}(k,l)T_{obs}^2\int_0^1\frac{\partial\lambda(\phi+\xi_k)}{\partial\xi_k}d\phi\int_0^1\frac{\partial\lambda(\phi+\xi_l)}{\partial\xi_l}\frac{\lambda(\phi+\xi_k)}{\lambda(\phi)}d\phi. \quad (28)$$

Finally, plugging (25), (26), (27), and (28) into (19) we get (9).

C. Derivation of $\mathbf{H}(k,l)$

From (7), the element (k,l) of the matrix $\mathbf{H}$ is given by

$$\mathbf{H}(k,l) = E_{\mathbf{x};\phi_0}\left[\frac{\partial \ln p(\mathbf{x};\theta_k)}{\partial\theta_k}\frac{p(\mathbf{x};\theta_k)}{p(\mathbf{x};\phi_0)}\frac{p(\mathbf{x};\theta_l)}{p(\mathbf{x};\phi_0)}\right]$$
$$= \sum_{\mathbf{x}=0}^{\infty}\left[\sum_{p=0}^{N-1}x_p\frac{\partial \ln[\lambda_p(\theta_k)\Delta t]}{\partial\theta_k} - \frac{\partial\lambda_p(\theta_k)\Delta t}{\partial\theta_k}\right]\prod_{n=0}^{N-1}\left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x_n}\frac{1}{x_n!}\exp\left(\Delta t[\lambda_n(\phi_0)-\lambda_n(\theta_k)-\lambda_n(\theta_l)]\right)$$
$$= \sum_{p=0}^{N-1}\frac{\partial \ln[\lambda_p(\theta_k)\Delta t]}{\partial\theta_k}\sum_{\mathbf{x}=0}^{\infty}x_p\prod_{n=0}^{N-1}\left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x_n}\frac{1}{x_n!}\exp\left(\Delta t[\lambda_n(\phi_0)-\lambda_n(\theta_k)-\lambda_n(\theta_l)]\right)$$
$$+\sum_{p=0}^{N-1}\frac{\partial\lambda_p(\theta_k)\Delta t}{\partial\theta_k}\prod_{n=0}^{N-1}\left[\frac{\lambda_n(\theta_k)\lambda_n(\theta_l)\Delta t}{\lambda_n(\phi_0)}\right]^{x_n}\frac{1}{x_n!}\exp\left(\Delta t[\lambda_n(\phi_0)-\lambda_n(\theta_k)-\lambda_n(\theta_l)]\right)$$
$$= \mathbf{M}_{MS}(k,l)\left[T_{obs}\int_0^1\frac{\partial\lambda(\phi+\xi_k)}{\partial\xi_k}\frac{\lambda(\phi+\xi_l)}{\lambda(\phi)}d\phi - T_{obs}\int_0^1\frac{\partial\lambda(\phi+\xi_k)}{\partial\xi_k}d\phi\right]. \quad (29)$$

D. Derivation of $\eta'(\alpha,\beta)$

We have

$$\eta'(\alpha,\beta) = \sum_{\mathbf{x}=0}^{\infty}\left\{\prod_{n=0}^{N-1}\frac{[\lambda_n(\phi_0+\beta)\Delta t]^{x_n}}{x_n!}\,exp(-\lambda_n(\phi_0+\beta)\Delta t)\right\}^{\alpha}\left\{\prod_{m=0}^{N-1}\frac{[\lambda_m(\phi_0)\Delta t]^{x_m}}{x_m!}\,exp(-\lambda_m(\phi_0)\Delta t)\right\}^{1-\alpha}$$
$$= \sum_{\mathbf{x}=0}^{\infty}\prod_{n=0}^{N-1}\frac{[\lambda_n(\phi_0+\beta)^{\alpha}\lambda_n(\phi_0)^{1-\alpha}\Delta t]^{x_n}}{x_n!}\exp\left[-\alpha\lambda_n(\phi_0+\beta)\Delta t-(1-\alpha)\lambda_n(\phi_0)\Delta t\right]$$
$$= \exp\left\{T_{obs}\int_0^1\left(-\alpha\lambda(\phi+\beta)-(1-\alpha)\lambda(\phi)+\lambda(\phi+\beta)^{\alpha}\lambda(\phi)^{1-\alpha}\right)d\phi\right\}. \quad (30)$$

REFERENCES

[1] P.Ray, K. Wood, and B. Philips, "Spacecraft navigation using X-ray pulsars," *Naval Research lab. (NRL) Review*, pp. 95–102, 2006.

[2] A. R. Golshan and S. I. Sheikh, "On pulsar phase estimation and tracking of variable celestial X-ray sources," in *63rd Ann. Meet. Inst. Navigat. (INO)*, Cambridge, MA, 2007, pp. 413–422.

[3] A. A. Emadzadeh and J. L. Speyer, "On modeling and pulse phase estimation of X-ray pulsars," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4484–4495, Sep. 2010.

[4] D. C. Rife and R. R. Boorstyn, "Single tone parameter estimation from discrete time observations," *IEEE Transactions on Information Theory*, vol. 20, no. 5, pp. 591–598, Sep. 1974.

[5] E. Chaumette, J. Galy, A. Quinlan, and P. Larzabal, "A new Barankin bound approximation for the prediction of the threshold region performance of maximum likelihood estimators," *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5319–5333, Nov. 2008.

[6] R. Aharoni and D. Lee, "On the achievability of the Cramér–Rao bound for poisson distribution," *IEEE Transactions on Information Theory*, vol. 47, no. 5, pp. 2096–2100, 2001.

[7] T. L. Marzetta, "Computing the Barankin bound by solving an unconstrained quadratic optimization problem," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 5, Munich, DE, Apr. 1997, pp. 3829–3832.

[8] A. Quinlan, E. Chaumette, and P. Larzabal, "A direct method to generate approximations of the Barankin bound," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, Toulouse, FR, May 2006, pp. 808–811.

[9] A. J. Weiss and E. Weinstein, "A lower bound on the mean square error in random parameter estimation," *IEEE Transactions on Information Theory*, vol. 31, no. 5, pp. 680–682, Sep. 1985.

[10] "The European Pulsar Network Database Browser." [Online]. Available: http://www.jb.man.ac.uk/~pulsar/Resources/epn/browser.html

[11] E. W. Barankin, "Locally best unbiased estimates," *The Annals of Mathematical Statistics*, vol. 20, no. 4, pp. 477–501, Dec. 1949.

[12] V. Trees and K. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. New York: Wiley, 2007.

[13] I. Rapoport and Y. Oshman, "Weiss-Weinstein lower bounds for Markovian systems. part 1: Theory," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 2016–2030, May 2007.

[14] ——, "Weiss-Weinstein lower bounds for Markovian systems. part 1: Theory," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 2031–2042, May 2007.

[15] A. Renaux, P. Forster, and P. Larzabal, "A new derivation of the Bayesian bounds for parameter estimation," in *Proc. of IEEE Workshop on Statistical Signal Processing (SSP)*, Bordeaux, FR, Jul. 2005, pp. 567–572.

[16] A. Renaux, P. Forster, P. Larzabal, C. Richmond, and A. Nehorai, "A fresh look at the Bayesian bounds of the Weiss-Weinstein family," *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5334–5352, 2008.

Annexe I

Some results on the Weiss-Weinstein bound in array processing

Soumis à *IEEE* Transactions on Signal Processing

Some results on the Weiss-Weinstein bound in array processing

Dinh Thang VU, Alexandre RENAUX, *Member, IEEE,* Rémy BOYER, *Member, IEEE,* Sylvie MARCOS

Abstract

In this paper, the Weiss-Weinstein bound is analyzed in the context of sources localization with a planar array of sensors. Both conditional and unconditional source signal models are studied. First, some results are given in the multiple sources context without specifying the structure of the steering matrix and of the noise covariance matrix. Second, these results are applied to the particular case of a single source for two kinds of array geometries: a non-uniform linear array (elevation only) and an arbitrary planar (azimuth and elevation) array.

Index Terms

Weiss-Weinstein bound, DOA estimation, array processing.

I. Introduction

Sources localization problem has been widely investigated in the literature with many applications such as radar, sonar, medical imaging, etc. One of the objective is to estimate the direction-of-arrival (DOA) of the sources using an array of sensors.

In array processing, lower bounds on the mean square error (MSE) are usually used as a benchmark to evaluate the ultimate performance of an estimator. There exist several lower bounds in the literature. Depending on the assumptions about the parameters of interest, there are three main kinds of lower bounds. When the parameters are assumed to be deterministic (unknown), the main lower bounds on the (local) MSE used are the well known Cramér-Rao bound [2] and the Barankin bound [3] (more particularly their approximations [4]–[8]. When the parameters are assumed to be random with a known prior distribution, these lower bounds on the global MSE are called Bayesian bounds [9]. Some typical families of Bayesian bounds are the Ziv-Zakai family [10]–[12] and the Weiss-Weinstein family [13]–[16]. Finally, when the parameter vector is made from both deterministic and random parameters, the so-called hybrid bounds have been developed [17]–[20].

The authors are with Université Paris-Sud 11, CNRS Laboratoire des Signaux et Systèmes, Supelec, 3 rue Joliot Curie, 91192 Gif-sur-Yvette Cedex, France (e-mail: {Vu,Renaux,Remy.Boyer,Marcos}@lss.supelec.fr)

This project was funded by both région Île de France and Digiteo Research Park. Section V-B2 of this paper has been partially presented in [1].

Since the DOA estimation is a non-linear problem, the outliers effect can appear and the estimators MSE exhibits three distinct behaviors depending on the number of snapshots and/or on the signal to noise ratio (SNR) [21]. At high SNR and/or for a high number of snapshots, *i.e.*, in the asymptotic region, the outliers effect can be neglected and the ultimate performance are described by the (classical/Bayesian/hybrid) Cramér-Rao bound. However, when the SNR and/or the number of snapshots decrease, the outliers effect lead to a quick increase of the MSE: this is the so-called threshold effect. In this region, the behavior of the lower bounds are not the same. Some bounds, generally called global bounds (Barankin, Ziv-Zakai, Weiss-Weinstein) can predict the threshold while the others, called local bounds, like the Cramér-Rao bound or the Bhattacharyya bound cannot. Finally, at low SNR and/or at low number of snapshots, *i.e.*, in the no-information region, the deterministic bounds exceed the estimator MSE due to the fact that they do not take into account the parameter support. On the contrary, the Bayesian bounds exploit the parameter prior information leading to a "real" lower bound on the global MSE.

In this paper, we are interested in the Weiss-Weinstein bounds which is known to be one of the tightest Bayesian bound with the bounds of the Ziv-Zakai family. We will study the two main source models used in the literature [22]: the unconditional (or stochastic) model where the source signals are assumed to be Gaussian and the conditional (or deterministic) model where the source signals are assumed to be deterministic. Surprisingly, in the context of array processing, while closed-form expressions of the Ziv-Zakai bound (more precisizely its extension by Bell et. al. [23]) were proposed around 15 years ago for the unconditional model, the results concerning the Weiss-Weinstein bound are, most of the time, only conducted by way of simulations. Concerning the unconditional model, in [24], the Weiss-Weinstein bound has been evaluated by way of simulations and has been compared to the MSE of the MUSIC algorithm and classical Beamforming using a particular 8×8 element array antenna. In [25], the authors have introduced a numerical comparison between the Bayesian Cramér-Rao bound, the Ziv-Zakai bound and the Weiss-Weinstein bound for DOA estimation. In [26], numerical simulations of the Weiss-Weinstein bound to optimize sensor positions for non-uniform linear arrays have been presented. Again in the unconditional model context, in [27], by considering the matched-field estimation problem, the authors have derived a semi closed-form expression of a simplified version of the Weiss-Weinstein bound for the DOA estimation. Indeed, the integration over the prior probability density function was not performed. The conditional model (with known waveforms) is studied only in [28], where a closed-form expression of the WWB is given in the simple case of spectral analysis and in [1] which is a simplified version of the bound.

While the primary goal of this paper is to provide closed-form expressions of the Weiss-Weinstein bound for the DOA estimation of a single source with an arbitrary planar array of sensors, under both conditional and unconditional source signal models, we also provide partial closed-form expressions of the bound which could be useful for other problems. First, we study the general Gaussian observation model with parameterized mean or parameterized covariance matrix. Indeed, one of the success of the Cramér-Rao is that, for this observation model, a closed-form expression of the Fisher information matrix is available: this is the so-called Slepian-Bang formula [29]. Such kind of formulas have been less investigated in the context of bounds tighter than the Cramér-rao bound. Second, some results are given in the multiple sources context without specifying the structure of the steering matrix

and of the noise covariance matrix. Finally, these results are applied to the particular case of a single source for two kinds of array geometries: the non-uniform linear array (elevation only) and the planar (azimuth and elevation) array. Consequently, the aim of this paper is also to provide a textbook of formulas which could be applied in other fields. The Weiss-Weinstein bound is known to depend on parameters called test points and other parameters generally denoted s_i. One particularity of this paper in comparison with the previous works on the Weiss-Weinstein bound is that we do not use the assumption $s_i = 1/2, \forall i$.

This paper is organized as follows. Section II is devoted to the array processing observation model which will be used in the paper. In Section III, a short background on the Weiss-Weinstein bound is presented and two general closed-form expressions which will be the cornerstone for our array processing problems are derived. In Section IV we apply these general results to the array processing problem without specifying the structure of the steering matrix. In Section V, we study the particular case of the non-uniform linear array and of the planar array for which we provide both closed-form expressions of the bound. Some simulation results are proposed in Section VI. Finally, Section VII gives our conclusions.

II. Problem setup

In this section, the general observation model usually used in array signal processing is presented as well as the different assumptions used in the remain of the paper. Particularly, the so-called conditional and unconditional source models are emphasized.

A. Observations model

We consider the classical scenario of an array with M sensors which receives N complex bandpass signals $\mathbf{s}(t) = [s_1(t)\ \ s_2(t) \cdots s_N(t)]^T$. The output of the array is a $M \times 1$ complex vector $\mathbf{y}(t)$ which can be modelled as follows (see, *e.g.*, [30] or [22])

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta})\,\mathbf{s}(t) + \mathbf{n}(t), \qquad t = 1, \ldots, T, \tag{1}$$

where T is the number of snapshots, where $\boldsymbol{\theta} = [\theta_1\ \theta_2 \cdots \theta_q]^T$ is an unknown parameter vector of interest[1], where $\mathbf{A}(\boldsymbol{\theta})$ is the so-called $M \times N$ steering matrix of the array response to the sources, and where the $M \times 1$ random vector $\mathbf{n}(t)$ is an additive noise.

B. Assumptions

- The unknown parameters of interest are assumed to be random with an *a priori* probability density function $p(\theta_i)$, $i = 1, \ldots, q$. These random parameters are assumed to be statistically independent such that the *a priori* joint probability density function is $p(\boldsymbol{\theta}) = \prod_{i=1}^{q} p(\theta_i)$. We also assume that the parameter space, denoted Θ, is a connected subset of $\mathbb{R}^q$ (see [31]).

[1]Note that one source can be described by several parameters. Consequently, $q > N$ in general.

- The noise vector is assumed to be complex Gaussian, statistically independent of the parameters, i.i.d., circular, with zero mean and known covariance matrix $\mathbb{E}\left[\mathbf{n}\left(t\right)\mathbf{n}^{H}\left(t\right)\right]=\mathbf{R_n}$. This assumption will be cancelled in Section V where it will be assumed that $\mathbf{R_n}=\sigma_n^2\mathbf{I}$. In any case, $\mathbf{R_n}$ is assumed to be a full rank matrix.
- The steering matrix $\mathbf{A}\left(\boldsymbol{\theta}\right)$ is assumed such that the observation model is identifiable. From Section III to Section IV, the structure of $\mathbf{A}\left(\boldsymbol{\theta}\right)$ is not specified in order to obtain the more general results.
- Concerning the source signals, two kinds of models have been investigated in the literature (see, *e.g.*, [32] or [22]) and will be alternatively used in this paper.
  - $\mathcal{M}_1$: *Unconditional or stochastic model*: $\mathbf{s}(t)$ is assumed to be a complex circular random vector, i.i.d., statistically independent of the noise, Gaussian with zero-mean and known covariance matrix $\mathbb{E}\left[\mathbf{s}\left(t\right)\mathbf{s}^{H}\left(t\right)\right]=\mathbf{R_s}$. Note that concerning the previous results on the Cramér-Rao bound available in the literature [32], the covariance matrix $\mathbf{R_s}$ is assumed to be unknown. In this paper, we have made the simpler assumption that the covariance matrix $\mathbf{R_s}$ is known. These assumptions have already been used for the calculus of bounds more complex than the Cramér-Rao bound (see, *e.g.*, [27], [33], [34]).
  - $\mathcal{M}_2$: *Conditional or deterministic model*: $\forall t$, $\mathbf{s}(t)$ is assumed to be deterministic known. Note that, under the conditional model assumption, the signal waveforms can be assumed either unknown or known. While the conditional observation model with unknown waveforms seems more challenging, the conditional model with known waveforms signals which will be used in this paper can be found in several applications such as in mobile telecommunication and radar (see *e.g.* [35], [36], [37], [38], and [39]).

C. Likelihood of the observations

Let $\mathbf{R_y}=\mathbb{E}\left[\mathbf{y}\left(t\right)\mathbf{y}^{H}\left(t\right)\right]$ be the covariance matrix of the observation vector $\mathbf{y}\left(t\right)$. According to the aforementioned assumptions, it is easy to see that under $\mathcal{M}_1$, the observations $\mathbf{y}\left(t\right)$ are distributed as a complex circular Gaussian random vector with zero mean and covariance matrix $\mathbf{R_y}(\boldsymbol{\theta})=\mathbf{A}(\boldsymbol{\theta})\mathbf{R_s}\mathbf{A}^{H}(\boldsymbol{\theta})+\mathbf{R_n}$ while under $\mathcal{M}_2$, the observations $\mathbf{y}\left(t\right)$ are distributed as a complex circular Gaussian random vector with mean $\mathbf{A}(\boldsymbol{\theta})\mathbf{s}\left(t\right)$ and covariance matrix $\mathbf{R_y}=\mathbf{R_n}$. Moreover, in both case the observations are i.i.d..

Therefore, the likelihood, $p\left(\mathbf{Y};\boldsymbol{\theta}\right)$, of the full observations matrix $\mathbf{Y}=\left[\mathbf{y}\left(1\right)\ \mathbf{y}\left(2\right)\ \ldots\ \mathbf{y}\left(T\right)\right]$ under $\mathcal{M}_1$ is given by

$$p\left(\mathbf{Y};\boldsymbol{\theta}\right)=\frac{1}{\pi^{MT}\left|\mathbf{R_y}(\boldsymbol{\theta})\right|^{T}}\exp\left(-\sum_{t=1}^{T}\mathbf{y}\left(t\right)^{H}\mathbf{R_y}^{-1}(\boldsymbol{\theta})\mathbf{y}\left(t\right)\right), \tag{2}$$

where $\mathbf{R_y}(\boldsymbol{\theta})=\mathbf{A}(\boldsymbol{\theta})\mathbf{R_s}\mathbf{A}^{H}(\boldsymbol{\theta})+\mathbf{R_n}$ and the likelihood under $\mathcal{M}_2$ is given by

$$p\left(\mathbf{Y};\boldsymbol{\theta}\right)=\frac{1}{\pi^{MT}\left|\mathbf{R_n}\right|^{T}}\exp\left(-\sum_{t=1}^{T}\left(\mathbf{y}\left(t\right)-\mathbf{A}\left(\boldsymbol{\theta}\right)\mathbf{s}\left(t\right)\right)^{H}\mathbf{R_n}^{-1}\left(\mathbf{y}\left(t\right)-\mathbf{A}\left(\boldsymbol{\theta}\right)\mathbf{s}\left(t\right)\right)\right). \tag{3}$$

III. Weiss-Weinstein bound: Generalities

In this Section, we first remind to the reader the structure of the Weiss-Weinstein bound on the MSE and the assumptions used to compute this bound. Second, a general result about the Gaussian observation model with

parameterized mean or parameterized covariance matrix, which, to the best of our knowledge, does not appear in the literature is presented. This result will be useful to study both the unconditional model $\mathcal{M}_1$ and the conditional model $\mathcal{M}_2$ in the next Section.

A. Background

The Weiss-Weinstein bound for a $q \times 1$ real parameter vector $\boldsymbol{\theta}$ is a $q \times q$ matrix denoted $\mathbf{WWB}$ and is given as follows [40]

$$\mathbf{WWB} = \mathbf{H}\mathbf{G}^{-1}\mathbf{H}^T, \tag{4}$$

where the $q \times q$ matrix $\mathbf{H} = [\mathbf{h}_1 \ \mathbf{h}_2 \dots \mathbf{h}_q]$ contains the so-called test-points $\mathbf{h}_i$, $i = 1, \dots, q$ such that $\boldsymbol{\theta} + \mathbf{h}_i \in \Theta$ $\forall \mathbf{h}_i$. The $k,l-$element of the $q \times q$ matrix $\mathbf{G}$ is given by

$$\{\mathbf{G}\}_{k,l} = \frac{\mathbb{E}\left[\left(L^{s_k}\left(\mathbf{Y};\boldsymbol{\theta}+\mathbf{h}_k,\boldsymbol{\theta}\right) - L^{1-s_k}\left(\mathbf{Y};\boldsymbol{\theta}-\mathbf{h}_k,\boldsymbol{\theta}\right)\right)\left(L^{s_l}\left(\mathbf{Y};\boldsymbol{\theta}+\mathbf{h}_l,\boldsymbol{\theta}\right) - L^{1-s_l}\left(\mathbf{Y};\boldsymbol{\theta}-\mathbf{h}_l,\boldsymbol{\theta}\right)\right)\right]}{\mathbb{E}\left[L^{s_k}\left(\mathbf{Y};\boldsymbol{\theta}+\mathbf{h}_k,\boldsymbol{\theta}\right)\right]\mathbb{E}\left[L^{s_l}\left(\mathbf{Y};\boldsymbol{\theta}+\mathbf{h}_l,\boldsymbol{\theta}\right)\right]}, \tag{5}$$

where the expectations are taken over the joint probability density function $p\left(\mathbf{Y},\boldsymbol{\theta}\right)$ and where the function $L\left(\mathbf{Y};\boldsymbol{\theta}+\mathbf{h}_i,\boldsymbol{\theta}\right)$ is defined by $L\left(\mathbf{Y};\boldsymbol{\theta}+\mathbf{h}_i,\boldsymbol{\theta}\right) = \frac{p\left(\mathbf{Y},\boldsymbol{\theta}+\mathbf{h}_i\right)}{p\left(\mathbf{Y},\boldsymbol{\theta}\right)}$. The elements s_i are such that $s_i \in [0,1]$, $i = 1, \dots, q$.

Note that we have the following order relation [40]

$$\mathbf{Cov}\left(\hat{\boldsymbol{\theta}}\right) = \mathbb{E}\left[\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)^T\right] \succeq \mathbf{WWB}, \tag{6}$$

where $\mathbf{A} \succeq \mathbf{B}$ means that the matrix $\mathbf{A} - \mathbf{B}$ is a semi-positive definite matrix and where $\mathbf{Cov}\left(\hat{\boldsymbol{\theta}}\right)$ is the global (the expectation is taken over the joint pdf $p\left(\mathbf{Y},\boldsymbol{\theta}\right)$) MSE of any estimator $\hat{\boldsymbol{\theta}}$ of the parameter vector $\boldsymbol{\theta}$. Finally, in order to obtain a tight bound, one has to maximize $\mathbf{WWB}$ over the test-points $\mathbf{h}_i$ and s_i $i = 1, \dots, q$. Note that this maximization can be done by using the trace of $\mathbf{H}\mathbf{G}^{-1}\mathbf{H}^T$ or with respect to the Loewner partial ordering [41]. In this paper, we will use the trace of $\mathbf{H}\mathbf{G}^{-1}\mathbf{H}^T$ which is enough to obtain tight results.

B. A general result on the Weiss-Weinstein bound and its application to the Gaussian observation models

An analytical result on the Weiss-Weinstein bound which will be useful in the following derivations and which could be useful for other problems is derived in this part. Note that this result is independent of the parameter vector size q and of the considered observation model.

Let us denote Ω the observation space. By rewriting the elements of matrix $\mathbf{G}$ (see Eqn. (5)) involved in the Weiss-Weinstein bound, one obtains for the numerator denoted $N_{\{\mathbf{G}\}_{k,l}}$,

$$\begin{aligned} N_{\{\mathbf{G}\}_{k,l}} &= \mathbb{E}\left[\left(L^{s_k}\left(\mathbf{Y};\boldsymbol{\theta}+\mathbf{h}_k,\boldsymbol{\theta}\right) - L^{1-s_k}\left(\mathbf{Y};\boldsymbol{\theta}-\mathbf{h}_k,\boldsymbol{\theta}\right)\right)\left(L^{s_l}\left(\mathbf{Y};\boldsymbol{\theta}+\mathbf{h}_l,\boldsymbol{\theta}\right) - L^{1-s_l}\left(\mathbf{Y};\boldsymbol{\theta}-\mathbf{h}_l,\boldsymbol{\theta}\right)\right)\right] \\ &= \int_\Theta\int_\Omega \frac{p^{s_k}\left(\mathbf{Y},\boldsymbol{\theta}+\mathbf{h}_k\right)p^{s_l}\left(\mathbf{Y},\boldsymbol{\theta}+\mathbf{h}_l\right)}{p^{s_k+s_l-1}\left(\mathbf{Y},\boldsymbol{\theta}\right)}d\mathbf{Y}d\boldsymbol{\theta} + \int_\Theta\int_\Omega \frac{p^{1-s_k}\left(\mathbf{Y},\boldsymbol{\theta}-\mathbf{h}_k\right)p^{1-s_l}\left(\mathbf{Y},\boldsymbol{\theta}-\mathbf{h}_l\right)}{p^{1-s_k-s_l}\left(\mathbf{Y},\boldsymbol{\theta}\right)}d\mathbf{Y}d\boldsymbol{\theta} \\ &- \int_\Theta\int_\Omega \frac{p^{s_k}\left(\mathbf{Y},\boldsymbol{\theta}+\mathbf{h}_k\right)p^{1-s_l}\left(\mathbf{Y},\boldsymbol{\theta}-\mathbf{h}_l\right)}{p^{s_k-s_l}\left(\mathbf{Y},\boldsymbol{\theta}\right)}d\mathbf{Y}d\boldsymbol{\theta} - \int_\Theta\int_\Omega \frac{p^{1-s_k}\left(\mathbf{Y},\boldsymbol{\theta}-\mathbf{h}_k\right)p^{s_l}\left(\mathbf{Y},\boldsymbol{\theta}+\mathbf{h}_l\right)}{p^{s_l-s_k}\left(\mathbf{Y},\boldsymbol{\theta}\right)}d\mathbf{Y}d\boldsymbol{\theta}, \end{aligned} \tag{7}$$

and for the denominator denoted $D_{\{\mathbf{G}\}_{k,l}}$,

$$D_{\{\mathbf{G}\}_{k,l}} = \mathbb{E}\left[L^{s_k}\left(\mathbf{Y};\boldsymbol{\theta}+\mathbf{h}_k,\boldsymbol{\theta}\right)\right]\mathbb{E}\left[L^{s_l}\left(\mathbf{Y};\boldsymbol{\theta}+\mathbf{h}_l,\boldsymbol{\theta}\right)\right]$$
$$= \int_\Theta\int_\Omega \frac{p^{s_k}\left(\mathbf{Y},\boldsymbol{\theta}+\mathbf{h}_k\right)}{p^{s_k-1}\left(\mathbf{Y},\boldsymbol{\theta}\right)}d\mathbf{Y}d\boldsymbol{\theta}\int_\Theta\int_\Omega \frac{p^{s_l}\left(\mathbf{Y},\boldsymbol{\theta}+\mathbf{h}_l\right)}{p^{s_l-1}\left(\mathbf{Y},\boldsymbol{\theta}\right)}d\mathbf{Y}d\boldsymbol{\theta}. \quad (8)$$

Let us now define a function $\eta\left(\alpha,\beta,\mathbf{u},\mathbf{v}\right)$ as

$$\eta\left(\alpha,\beta,\mathbf{u},\mathbf{v}\right) = \int_\Theta\int_\Omega \frac{p^{\alpha}\left(\mathbf{Y},\boldsymbol{\theta}+\mathbf{u}\right)p^{\beta}\left(\mathbf{Y},\boldsymbol{\theta}+\mathbf{v}\right)}{p^{\alpha+\beta-1}\left(\mathbf{Y},\boldsymbol{\theta}\right)}d\mathbf{Y}d\boldsymbol{\theta}, \quad (9)$$

where $(\alpha,\beta)\in[0,1]^2$ and where $(\mathbf{u},\mathbf{v})$ are two $q\times 1$ vectors such that $\boldsymbol{\theta}+\mathbf{u}\in\Theta$ and $\boldsymbol{\theta}+\mathbf{v}\in\Theta$. By identification, it is easy to see that

$$\{\mathbf{G}\}_{k,l} = \frac{\eta\left(s_k,s_l,\mathbf{h}_k,\mathbf{h}_l\right)+\eta\left(1-s_k,1-s_l,-\mathbf{h}_k,-\mathbf{h}_l\right)-\eta\left(s_k,1-s_l,\mathbf{h}_k,-\mathbf{h}_l\right)-\eta\left(1-s_k,s_l,-\mathbf{h}_k,\mathbf{h}_l\right)}{\eta\left(s_k,0,\mathbf{h}_k,\mathbf{0}\right)\eta\left(0,s_l,\mathbf{0},\mathbf{h}_l\right)}. \quad (10)$$

Note that we choose the arbitrary notation $D_{\{\mathbf{G}\}_{k,l}} = \eta\left(s_k,0,\mathbf{h}_k,\mathbf{0}\right)\eta\left(0,s_l,\mathbf{0},\mathbf{h}_l\right)$ for the denominator. The notation $D_{\{\mathbf{G}\}_{k,l}} = \eta\left(s_k,1,\mathbf{h}_i,\mathbf{0}\right)\eta\left(1,s_l,\mathbf{0},\mathbf{h}_l\right)$ or, even, $D_{\{\mathbf{G}\}_{k,l}} = \eta\left(s_k,0,\mathbf{h}_k,\mathbf{v}\right)\eta\left(0,s_l,\mathbf{u},\mathbf{h}_l\right)$ will lead to the same result.

With Eqn. (10), it is clear that the knowledge of $\eta\left(\alpha,\beta,\mathbf{u},\mathbf{v}\right)$ for a particular problem leads to the Weiss-Weinstein bound (without the maximization procedure over the test-points and over the parameters s_i). Surprisingly, this simple expression is given in [40] only for a particular case $s_i=\frac{1}{2}$, $\forall i$ and not for the general case.

Let us now detail this function $\eta\left(\alpha,\beta,\mathbf{u},\mathbf{v}\right)$. Thanks to the Bayes rule, the function $\eta\left(\alpha,\beta,\mathbf{u},\mathbf{v}\right)$ can be rewritten as

$$\eta\left(\alpha,\beta,\mathbf{u},\mathbf{v}\right) = \int_\Theta \frac{p^{\alpha}\left(\boldsymbol{\theta}+\mathbf{u}\right)p^{\beta}\left(\boldsymbol{\theta}+\mathbf{v}\right)}{p^{\alpha+\beta-1}\left(\boldsymbol{\theta}\right)}\int_\Omega \frac{p^{\alpha}\left(\mathbf{Y};\boldsymbol{\theta}+\mathbf{u}\right)p^{\beta}\left(\mathbf{Y};\boldsymbol{\theta}+\mathbf{v}\right)}{p^{\alpha+\beta-1}\left(\mathbf{Y};\boldsymbol{\theta}\right)}d\mathbf{Y}d\boldsymbol{\theta}$$
$$= \int_\Theta \acute{\eta}_{\boldsymbol{\theta}}\left(\alpha,\beta,\mathbf{u},\mathbf{v}\right)\frac{p^{\alpha}\left(\boldsymbol{\theta}+\mathbf{u}\right)p^{\beta}\left(\boldsymbol{\theta}+\mathbf{v}\right)}{p^{\alpha+\beta-1}\left(\boldsymbol{\theta}\right)}d\boldsymbol{\theta}, \quad (11)$$

where we define

$$\acute{\eta}_{\boldsymbol{\theta}}\left(\alpha,\beta,\mathbf{u},\mathbf{v},\boldsymbol{\theta}\right) = \int_\Omega \frac{p^{\alpha}\left(\mathbf{Y};\boldsymbol{\theta}+\mathbf{u}\right)p^{\beta}\left(\mathbf{Y};\boldsymbol{\theta}+\mathbf{v}\right)}{p^{\alpha+\beta-1}\left(\mathbf{Y};\boldsymbol{\theta}\right)}d\mathbf{Y}. \quad (12)$$

Our aim is to give the most general result. Consequently, we will focus only on $\acute{\eta}_{\boldsymbol{\theta}}\left(\alpha,\beta,\mathbf{u},\mathbf{v}\right)$ since the *a priori* probability density function depends on the considered problem.

1) Gaussian observation model with parameterized covariance matrix: One calls (circular, i.i.d.) Gaussian observation model with parameterized covariance matrix, a model such that the observations $\mathbf{y}\left(t\right)\sim\mathcal{CN}\left(\mathbf{0},\mathbf{R}_{\mathbf{y}}\left(\boldsymbol{\theta}\right)\right)$ where $\boldsymbol{\theta}$ are the parameters of interest. Note that $\mathcal{M}_1$ is a special case of this model since the parameters of interest appear only in the covariance matrix of the observations which has the following particular structure $\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}) = \mathbf{A}(\boldsymbol{\theta})\mathbf{R}_{\mathbf{s}}\mathbf{A}^H(\boldsymbol{\theta})+\mathbf{R}_{\mathbf{n}}$. The closed-form expression of $\acute{\eta}_{\boldsymbol{\theta}}\left(\alpha,\beta,\mathbf{u},\mathbf{v}\right)$ is given by:

$$\acute{\eta}_{\boldsymbol{\theta}}\left(\alpha,\beta,\mathbf{u},\mathbf{v}\right) = \frac{\left|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta})\right|^{T(\alpha+\beta-1)}}{\left|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}+\mathbf{u})\right|^{T\alpha}\left|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}+\mathbf{v})\right|^{T\beta}\left|\alpha\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}+\mathbf{u})+\beta\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}+\mathbf{v})-(\alpha+\beta-1)\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta})\right|^{T}}. \quad (13)$$

The proof is given in Appendix A.

2) Gaussian observation model with parameterized mean: One calls (circular, i.i.d.) Gaussian observation model with parameterized mean, a model such that the observations $\mathbf{y}(t) \sim \mathcal{CN}(\mathbf{f}(\boldsymbol{\theta}), \mathbf{R}_{\mathbf{y}})$ where $\boldsymbol{\theta}$ are the parameters of interest. Note that $\mathcal{M}_2$ is a special case of this model since the parameters of interest appear only in the mean of the observations which has the following particular structure $\mathbf{f}(\boldsymbol{\theta}) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t)$ (and $\mathbf{R}_{\mathbf{y}} = \mathbf{R}_{\mathbf{n}}$). For notational convenience, we do not emphasize the dependence of $\mathbf{f}(\boldsymbol{\theta})$ on t. The closed-form expression of $\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v})$ is given in this case by

$$\begin{aligned}\ln \acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v}) = -\sum_{t=1}^{T} & \alpha(1-\alpha)\mathbf{f}^H(\boldsymbol{\theta}+\mathbf{u})\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{f}(\boldsymbol{\theta}+\mathbf{u}) + \beta(1-\beta)\mathbf{f}^H(\boldsymbol{\theta}+\mathbf{v})\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{f}(\boldsymbol{\theta}+\mathbf{v}) \\ & + (1-\alpha-\beta)(\alpha+\beta)\mathbf{f}^H(\boldsymbol{\theta})\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{f}(\boldsymbol{\theta}) - 2\,\mathrm{Re}\left\{\alpha\beta\mathbf{f}^H(\boldsymbol{\theta}+\mathbf{u})\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{f}(\boldsymbol{\theta}+\mathbf{v})\right. \\ & \left. + \alpha(1-\alpha-\beta)\mathbf{f}^H(\boldsymbol{\theta}+\mathbf{u})\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{f}(\boldsymbol{\theta}) + \beta(1-\alpha-\beta)\mathbf{f}^H(\boldsymbol{\theta}+\mathbf{v})\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{f}(\boldsymbol{\theta})\right\}, \end{aligned} \tag{14}$$

or equivalently by

$$\begin{aligned}\ln \acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v}) = -\sum_{t=1}^{T} & \alpha(1-\alpha-\beta)\left\|\mathbf{R}_{\mathbf{y}}^{-1/2}(\mathbf{f}(\boldsymbol{\theta}+\mathbf{u}) - \mathbf{f}(\boldsymbol{\theta}))\right\|^2 + \alpha\beta\left\|\mathbf{R}_{\mathbf{y}}^{-1/2}(\mathbf{f}(\boldsymbol{\theta}+\mathbf{u}) - \mathbf{f}(\boldsymbol{\theta}+\mathbf{v}))\right\|^2 \\ & + \beta(1-\alpha-\beta)\left\|\mathbf{R}_{\mathbf{y}}^{-1/2}(\mathbf{f}(\boldsymbol{\theta}+\mathbf{v}) - \mathbf{f}(\boldsymbol{\theta}))\right\|^2. \end{aligned} \tag{15}$$

The details are given in Appendix B.

IV. General application to array processing

In the previous Section, it has been shown that the Weiss-Weinstein bound computation (or, at least, the computation of matrix $\mathbf{G}$) is reduced to the knowledge of the function $\eta(\alpha, \beta, \mathbf{u}, \mathbf{v})$ given by Eqn. (9). As one can see in Eqn. (10), the elements of the matrix $\mathbf{G}$ depend on $\eta(\alpha, \beta, \mathbf{u}, \mathbf{v})$ for particular values of α, β, $\mathbf{u}$, and $\mathbf{v}$. Consequently, the goal of this Section is to detail these particular functions for our model given by Eqn. (1). Since Eqn. (9) can be decomposed into a *deterministic part* (in the sense where $\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v})$ (see Eqn. (12)) only depends on the likelihood function) and a *Bayesian part* (when we have to integrate $\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v})$ over the *a priori* probability density function of the parameters), we will first focus on the particular functions $\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v})$ by using the results of the previous Section on the Gaussian observation model with parameterized mean or covariance matrix. Second, we will detail the passage from $\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v})$ to $\eta(\alpha, \beta, \mathbf{u}, \mathbf{v})$ in the particular case where $p(\theta_i)$ is a uniform probability density function $\forall i$. Another result will also be given in the case of a Gaussian prior.

A. Analysis of $\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v})$

We will now detail the particular functions $\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v})$ involved in the different elements of $\{\mathbf{G}\}_{k,l}$, $k, l \in \{1, q\}^2$ for both models $\mathcal{M}_1$ and $\mathcal{M}_2$.

1) Unconditional observation model $\mathcal{M}_1$: Under the unconditional model $\mathcal{M}_1$, by using Eqn. (13), one obtains straightforwardly the functions $\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v})$ involved in the elements $\{\mathbf{G}\}_{k,l} = \{\mathbf{G}\}_{l,k}$

$$\begin{cases} \acute{\eta}_{\boldsymbol{\theta}}(s_k, s_l, \mathbf{h}_k, \mathbf{h}_l) = \frac{|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta})|^{T(s_k+s_l-1)}}{|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}+\mathbf{h}_k)|^{Ts_k}|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}+\mathbf{h}_l)|^{Ts_l}\left|s_k\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}+\mathbf{h}_k)+s_l\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}+\mathbf{h}_l)-(s_k+s_l-1)\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta})\right|^T}, \\ \acute{\eta}_{\boldsymbol{\theta}}(1-s_k, 1-s_l, -\mathbf{h}_k, -\mathbf{h}_l) = \frac{|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta})|^{T(1-s_k-s_l)}|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}-\mathbf{h}_k)|^{T(s_k-1)}|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}-\mathbf{h}_l)|^{T(s_l-1)}}{\left|(1-s_k)\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}-\mathbf{h}_k)+(1-s_l)\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}-\mathbf{h}_l)-(1-s_k-s_l)\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta})\right|^T}, \\ \acute{\eta}_{\boldsymbol{\theta}}(s_k, 1-s_l, \mathbf{h}_k, -\mathbf{h}_l) = \frac{|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta})|^{T(s_k-s_l)}|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}-\mathbf{h}_l)|^{T(s_l-1)}}{|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}+\mathbf{h}_k)|^{Ts_k}\left|s_k\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}+\mathbf{h}_k)+(1-s_l)\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}-\mathbf{h}_l)-(s_k-s_l)\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta})\right|^T}, \\ \acute{\eta}_{\boldsymbol{\theta}}(1-s_k, s_l, -\mathbf{h}_k, \mathbf{h}_l) = \frac{|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta})|^{T(s_l-s_k)}|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}-\mathbf{h}_k)|^{T(s_k-1)}}{|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}+\mathbf{h}_l)|^{Ts_l}\left|(1-s_k)\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}-\mathbf{h}_k)+s_l\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}+\mathbf{h}_l)-(s_l-s_k)\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta})\right|^T}, \\ \acute{\eta}_{\boldsymbol{\theta}}(s_k, 0, \mathbf{h}_k, \mathbf{0}) = \frac{|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta})|^{T(s_k-1)}}{|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}+\mathbf{h}_k)|^{Ts_k}\left|s_k\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}+\mathbf{h}_k)-(s_k-1)\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta})\right|^T}, \\ \acute{\eta}_{\boldsymbol{\theta}}(0, s_l, \mathbf{0}, \mathbf{h}_l) = \frac{|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta})|^{T(s_l-1)}}{|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}+\mathbf{h}_l)|^{Ts_l}\left|s_l\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}+\mathbf{h}_l)-(s_l-1)\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta})\right|^T}. \end{cases} \tag{16}$$

The diagonal elements of $\mathbf{G}$ are obtained by letting $k = l$ in the above equations.

2) Conditional observation model $\mathcal{M}_2$: Under the conditional model $\mathcal{M}_2$, by using Eqn. (15) with $\mathbf{f}(\boldsymbol{\theta}) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t)$ and $\mathbf{R}_{\mathbf{y}} = \mathbf{R}_{\mathbf{n}}$ one obtains straightforwardly the functions $\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v})$ involved in the elements $\{\mathbf{G}\}_{k,l} = \{\mathbf{G}\}_{l,k}$

$$\begin{cases} \ln \acute{\eta}_{\boldsymbol{\theta}}(s_k, s_l, \mathbf{h}_k, \mathbf{h}_l) = s_k(s_k+s_l-1)\zeta_{\boldsymbol{\theta}}(\mathbf{h}_k, \mathbf{0}) + s_l(s_k+s_l-1)\zeta_{\boldsymbol{\theta}}(\mathbf{h}_l, \mathbf{0}) - s_k s_l \zeta_{\boldsymbol{\theta}}(\mathbf{h}_k, \mathbf{h}_l), \\ \ln \acute{\eta}_{\boldsymbol{\theta}}(1-s_k, 1-s_l, -\mathbf{h}_k, -\mathbf{h}_l) = (s_k-1)(s_k+s_l-1)\zeta_{\boldsymbol{\theta}}(-\mathbf{h}_k, \mathbf{0}) + (s_l-1)(s_k+s_l-1)\zeta_{\boldsymbol{\theta}}(-\mathbf{h}_l, \mathbf{0}) \\ \qquad\qquad - (1-s_k)(1-s_l)\zeta_{\boldsymbol{\theta}}(-\mathbf{h}_k, -\mathbf{h}_l), \\ \ln \acute{\eta}_{\boldsymbol{\theta}}(s_k, 1-s_l, \mathbf{h}_k, -\mathbf{h}_l) = s_k(s_k-s_l)\zeta_{\boldsymbol{\theta}}(\mathbf{h}_k, \mathbf{0}) + (1-s_l)(s_k-s_l)\zeta_{\boldsymbol{\theta}}(-\mathbf{h}_l, \mathbf{0}) + s_k(s_l-1)\zeta_{\boldsymbol{\theta}}(\mathbf{h}_k, -\mathbf{h}_l), \\ \ln \acute{\eta}_{\boldsymbol{\theta}}(1-s_k, s_l, -\mathbf{h}_k, \mathbf{h}_l) = (s_k-1)(s_k-s_l)\zeta_{\boldsymbol{\theta}}(-\mathbf{h}_k, \mathbf{0}) + s_l(s_l-s_k)\zeta_{\boldsymbol{\theta}}(\mathbf{h}_l, \mathbf{0}) + (s_k-1)s_l\zeta_{\boldsymbol{\theta}}(-\mathbf{h}_k, \mathbf{h}_l), \\ \ln \acute{\eta}_{\boldsymbol{\theta}}(s_k, 0, \mathbf{h}_k, \mathbf{0}) = s_k(s_k-1)\zeta_{\boldsymbol{\theta}}(\mathbf{h}_k, \mathbf{0}), \\ \ln \acute{\eta}_{\boldsymbol{\theta}}(0, s_l, \mathbf{0}, \mathbf{h}_l) = s_l(s_l-1)\zeta_{\boldsymbol{\theta}}(\mathbf{h}_l, \mathbf{0}), \end{cases} \tag{17}$$

where we define

$$\zeta_{\boldsymbol{\theta}}(\boldsymbol{\mu}, \boldsymbol{\rho}) = \sum_{t=1}^{T} \left\| \mathbf{R}_{\mathbf{n}}^{-1/2} \left(\mathbf{A}(\boldsymbol{\theta}+\boldsymbol{\mu}) - \mathbf{A}(\boldsymbol{\theta}+\boldsymbol{\rho}) \right) \mathbf{s}(t) \right\|^2. \tag{18}$$

The diagonal elements of $\mathbf{G}$ are obtained by letting $k = l$ in the above equations. Note that, since we are working on matrix $\mathbf{G}$, all the previously proposed results are made whatever the number of test-points.

B. Analysis of $\eta(\alpha, \beta, \mathbf{u}, \mathbf{v})$ with a uniform prior

Of course, the analysis of $\eta(\alpha, \beta, \mathbf{u}, \mathbf{v})$ given by Eqn. (11) can only be conducted by specifying the *a priori* probability density functions of the parameters. Consequently, the results provided here are very specific. However, note that, in general, this aspect is less emphasized in the literature where most of the authors give results without specifying the prior probability density functions and compute the "remain" of the bound numerically (see e.g., [27] [23]).

We assume that all the parameters θ_i have a uniform prior distribution over the interval $[a_i, b_i]$ and are statistically independent. We will also assume one test-point per parameter, otherwise there is no possibility to obtain (pseudo) closed-form expressions. Consequently, the matrix $\mathbf{H}$ is such that

$$\mathbf{H} = Diag\left([h_1 \ h_2 \cdots h_q]\right), \tag{19}$$

and the vector $\mathbf{h}_i$, $i = 1, \ldots, q$, takes the value h_i at the i^{th} row and zero elsewhere. So, in this analysis, the vector $\mathbf{u}$ takes the value u_i at the i^{th} row and zero elsewhere and the vector $\mathbf{v}$ takes the value v_j at the j^{th} row and zero elsewhere (of course, we can have $i = j$). Under these assumptions, $\eta(\alpha, \beta, \mathbf{u}, \mathbf{v})$ can be rewritten for $i \neq j$

$$\begin{aligned}\eta(\alpha,\beta,\mathbf{u},\mathbf{v}) &= \int_{\Theta} \acute{\eta}_{\boldsymbol{\theta}}(\alpha,\beta,\mathbf{u},\mathbf{v}) \frac{p^{\alpha}(\theta_i+u_i)\,p^{\beta}(\theta_j+v_j)\,p^{\beta}(\theta_i)\,p^{\alpha}(\theta_j)}{p^{\alpha+\beta-1}(\theta_i)\,p^{\alpha+\beta-1}(\theta_j)} \prod_{\substack{k=1\\k\neq i,k\neq j}}^{q} p(\theta_k)\,d\boldsymbol{\theta} \\ &= \frac{1}{\prod_{k=1}^{q}(b_k-a_k)} \int_{\Theta^{q-2}} \int_{\Theta_j} \int_{\Theta_i} \acute{\eta}_{\boldsymbol{\theta}}(\alpha,\beta,\mathbf{u},\mathbf{v})\,d\theta_i d\theta_j d(\boldsymbol{\theta}/\{\theta_i,\theta_j\}), \end{aligned} \tag{20}$$

where $\Theta_i = \begin{cases} [a_i, b_i - u_i] & \text{if } u_i > 0, \\ [a_i - u_i, b_i] & \text{if } u_i < 0, \end{cases}$ and $\Theta_j = \begin{cases} [a_j, b_j - v_j] & \text{if } v_j > 0, \\ [a_j - v_j, b_j] & \text{if } v_j < 0, \end{cases}$. For $i = j$, one can have $\mathbf{v} = \pm\mathbf{u}$, then one obtains

$$\begin{aligned}\eta(\alpha,\beta,\mathbf{u},\mathbf{v}=\pm\mathbf{u}) &= \int_{\Theta} \acute{\eta}_{\boldsymbol{\theta}}(\alpha,\beta,\mathbf{u},\mathbf{v}) \frac{p^{\alpha}(\theta_i+u_i)\,p^{\beta}(\theta_i\pm u_i)}{p^{\alpha+\beta-1}(\theta_i)} \prod_{\substack{k=1\\k\neq i}}^{q} p(\theta_k)\,d\boldsymbol{\theta} \\ &= \frac{1}{\prod_{k=1}^{q}(b_k-a_k)} \int_{\Theta^{q-1}} \int_{\Theta_i} \acute{\eta}_{\boldsymbol{\theta}}(\alpha,\beta,\mathbf{u},\mathbf{v}=\pm\mathbf{u})\,d\theta_i d(\boldsymbol{\theta}/\{\theta_i\}). \end{aligned} \tag{21}$$

In the last equation, if $\mathbf{v} = -\mathbf{u}$, then $\Theta_i = \begin{cases} [a_i + u_i, b_i - u_i] & \text{if } u_i > 0, \\ [a_i - u_i, b_i + u_i] & \text{if } u_i < 0, \end{cases}$, while, if $\mathbf{v} = \mathbf{u}$, then $\Theta_i = \begin{cases} [a_i, b_i - u_i] & \text{if } u_i > 0, \\ [a_i - u_i, b_i] & \text{if } u_i < 0, \end{cases}$.

Depending on the structure of $\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v})$, $\eta(\alpha, \beta, \mathbf{u}, \mathbf{v})$ has to be computed numerically or a closed-form expression can be found.

Another particular case which appears sometimes is when the function $\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v})$ does not depend on $\boldsymbol{\theta}$ (see, e.g., [28] and Section V of this paper). In this case, $\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v})$ is denoted $\acute{\eta}(\alpha, \beta, \mathbf{u}, \mathbf{v})$ and one obtains from Eqn. (20)

$$\begin{aligned}\eta(\alpha,\beta,\mathbf{u},\mathbf{v}) &= \frac{\acute{\eta}(\alpha,\beta,\mathbf{u},\mathbf{v})}{\prod_{k=1}^{q}(b_k-a_k)} \left(\prod_{\substack{k=1\\k\neq i,k\neq j}}^{q} \int_{a_k}^{b_k} d\theta_k \right) \int_{\Theta_i} d\theta_i \int_{\Theta_j} d\theta_j \\ &= \frac{(b_i - a_i - |u_i|)(b_j - a_j - |v_j|)}{(b_i - a_i)(b_j - a_j)} \acute{\eta}(\alpha,\beta,\mathbf{u},\mathbf{v}), \end{aligned} \tag{22}$$

and from Eqn. (21)

$$\eta(\alpha,\beta,\mathbf{u},\mathbf{v}=\mathbf{u}) = \frac{(b_i - a_i - |u_i|)}{(b_i - a_i)} \acute{\eta}(\alpha,\beta,\mathbf{u},\mathbf{v}), \tag{23}$$

and

$$\eta(\alpha,\beta,\mathbf{u},\mathbf{v}=-\mathbf{u}) = \frac{(b_i - a_i - 2|u_i|)}{(b_i - a_i)} \acute{\eta}(\alpha,\beta,\mathbf{u},\mathbf{v}). \tag{24}$$

C. Analysis of $\eta(\alpha,\beta,\mathbf{u},\mathbf{v})$ with a Gaussian prior

Finally, one can mention that if the prior is now assumed to be Gaussian, *i.e.*, $\theta_i \sim \mathcal{N}\left(\mu_i,\sigma_i^2\right)\forall i$ and $\acute{\eta}_{\boldsymbol{\theta}}(\alpha,\beta,\mathbf{u},\mathbf{v})$ does not depend on $\boldsymbol{\theta}$ one obtains after a straightforward calculus

$$\begin{aligned}\eta(\alpha,\beta,\mathbf{u},\mathbf{v}) &= \acute{\eta}(\alpha,\beta,\mathbf{u},\mathbf{v})\int_{\mathbb{R}}\frac{p^{\alpha}(\theta_i+u_i)}{p^{\alpha-1}(\theta_i)}d\theta_i\int_{\mathbb{R}}\frac{p^{\beta}(\theta_j+v_j)}{p^{\beta-1}(\theta_j)}d\theta_j \\ &= \acute{\eta}(\alpha,\beta,\mathbf{u},\mathbf{v})\exp\left(-\frac{1}{2}\left(\frac{\alpha(1-\alpha)u_i^2}{\sigma_i^2}+\frac{\beta(1-\beta)v_j^2}{\sigma_j^2}\right)\right), \end{aligned} \tag{25}$$

$$\begin{aligned}\eta(\alpha,\beta,\mathbf{u},\mathbf{v}=\mathbf{u}) &= \acute{\eta}(\alpha,\beta,\mathbf{u},\mathbf{v})\int_{\mathbb{R}}\frac{p^{\alpha+\beta}(\theta_i+u_i)}{p^{\alpha+\beta-1}(\theta_i)}d\theta_i \\ &= \acute{\eta}(\alpha,\beta,\mathbf{u},\mathbf{v})\exp\left(-\frac{(\alpha+\beta)(1-\alpha-\beta)u_i^2}{2\sigma_i^2}\right), \end{aligned} \tag{26}$$

and

$$\begin{aligned}\eta(\alpha,\beta,\mathbf{u},\mathbf{v}=-\mathbf{u}) &= \acute{\eta}(\alpha,\beta,\mathbf{u},\mathbf{v})\int_{\mathbb{R}}\frac{p^{\alpha}(\theta_i+u_i)\,p^{\beta}(\theta_i-u_i)}{p^{\alpha+\beta-1}(\theta_i)}d\theta_i \\ &= \acute{\eta}(\alpha,\beta,\mathbf{u},\mathbf{v})\exp\left(-\frac{\left(\alpha+\beta-\alpha^2-\beta^2+2\alpha\beta\right)u_i^2}{2\sigma_i^2}\right). \end{aligned} \tag{27}$$

V. Specific applications to array processing: DOA estimation

We now consider the application of the Weiss-Weinstein bound in the particular context of source localization. Indeed, until now, the structure of the steering matrix $\mathbf{A}(\boldsymbol{\theta})$ for a particular problem has not been used in the proposed (semi) closed-form expressions. Consequently, these previous results can be applied to a large class of estimation problems such as far-field and near-field sources localization, passive localization with polarized array of sensors, or radar (known waveforms).

Here, we want to focus on the direction-of-arrival estimation of a single source in the far-field area with narrow-band signal. In this case, the steering matrix $\mathbf{A}(\boldsymbol{\theta})$ becomes a steering vector denoted $\mathbf{a}(\boldsymbol{\theta})$ (except for one preliminary result concerning the conditional model which will be given whatever the number of sources in Section V-A2). The structure of this vector will be specified by the analysis of two kinds of array geometry: the non-uniform linear array from which only one angle-of-arrival can be estimated ($\boldsymbol{\theta}$ becomes a scalar) and the arbitrary planar array from which both azimuth and elevation can be estimated ($\boldsymbol{\theta}$ becomes a 2×1 vector). In any cases, the array always consists of M identical, omnidirectional sensors. Both model $\mathcal{M}_1$ and $\mathcal{M}_2$ will be considered and the noise will be assumed spatially uncorrelated: $\mathbf{R}_{\mathbf{n}}=\sigma_n^2\mathbf{I}$. Since we focus on the single source scenario, the variance of the source signal $s(t)$ is denoted σ_s^2 for the model $\mathcal{M}_1$.

The general structure of the i^{th} element of the steering vector is as follows

$$\{\mathbf{a}(\boldsymbol{\theta})\}_i = \exp\left(j\frac{2\pi}{\lambda}\mathbf{r}_i^T\boldsymbol{\theta}\right),\ i=1,\ldots,M \tag{28}$$

where $\boldsymbol{\theta}$ represents the parameter vector, where λ denotes the wavelength, and where $\mathbf{r}_i$ denotes the coordinate of the i^{th} sensor position with respect to a given referential. In the following, $\mathbf{r}_i$ will be a scalar or a 2×1 vector depending on the context (linear array or planar array).

A. Preliminary results

Since our analysis is now reduced to the single source case, we give here some other closed-form expressions which will be useful when we will detail the specific linear and planar arrays.

1) Unconditional observation model $\mathcal{M}_1$: In order to detail the set of functions $\acute{\eta}_\theta$ given by Eqn. (16), one has to find closed-form expressions of the determinant $|\mathbf{R}_\mathbf{y}(\boldsymbol{\theta}+\mathbf{u})|$ and of determinants having the following structure: $\left|m_1\mathbf{R}_\mathbf{y}^{-1}(\boldsymbol{\theta}_1)+m_2\mathbf{R}_\mathbf{y}^{-1}(\boldsymbol{\theta}_2)\right|$ with $m_1+m_2=1$ or $\left|m_1\mathbf{R}_\mathbf{y}^{-1}(\boldsymbol{\theta}_1)+m_2\mathbf{R}_\mathbf{y}^{-1}(\boldsymbol{\theta}_2)+m_3\mathbf{R}_\mathbf{y}^{-1}(\boldsymbol{\theta}_3)\right|$ with $m_1+m_2+m_3=1$. Under $\mathcal{M}_1$, the observation covariance matrix is now given by

$$\mathbf{R}_\mathbf{y}(\boldsymbol{\theta})=\sigma_s^2\mathbf{a}(\boldsymbol{\theta})\mathbf{a}^H(\boldsymbol{\theta})+\sigma_n^2\mathbf{I}_M. \tag{29}$$

Concerning the calculus of $|\mathbf{R}_\mathbf{y}(\boldsymbol{\theta}+\mathbf{u})|$, it is easy to find

$$|\mathbf{R}_\mathbf{y}(\boldsymbol{\theta}+\mathbf{u})|=\sigma_n^{2M}\left(1+\frac{\sigma_s^2}{\sigma_n^2}\left\|\mathbf{a}(\boldsymbol{\theta}+\mathbf{u})\right\|^2\right). \tag{30}$$

Moreover, after calculus detailed in Appendix C, one obtains for the other determinants

$$\begin{aligned}\left|m_1\mathbf{R}_\mathbf{y}^{-1}(\boldsymbol{\theta}_1)+m_2\mathbf{R}_\mathbf{y}^{-1}(\boldsymbol{\theta}_2)\right|=\frac{1}{\left(\sigma_n^2\right)^M}\Big(&1-\varphi_1 m_1\left\|\mathbf{a}(\boldsymbol{\theta}_1)\right\|^2+m_2\varphi_2\left\|\mathbf{a}(\boldsymbol{\theta}_2)\right\|^2\\&-\varphi_1 m_1\varphi_2 m_2\left(\left\|\mathbf{a}^H(\boldsymbol{\theta}_1)\mathbf{a}(\boldsymbol{\theta}_2)\right\|^2-\left\|\mathbf{a}(\boldsymbol{\theta}_1)\right\|^2\left\|\mathbf{a}(\boldsymbol{\theta}_2)\right\|^2\right)\Big)\end{aligned} \tag{31}$$

and

$$\begin{aligned}&\left|m_1\mathbf{R}_\mathbf{y}^{-1}(\boldsymbol{\theta}_1)+m_2\mathbf{R}_\mathbf{y}^{-1}(\boldsymbol{\theta}_2)+m_3\mathbf{R}_\mathbf{y}^{-1}(\boldsymbol{\theta}_3)\right|=\\&\frac{1}{\left(\sigma_n^2\right)^M}\left(1-\sum_{k=1}^{3}m_k\varphi_k\left\|\mathbf{a}(\boldsymbol{\theta}_k)\right\|^2-\frac{1}{2}\sum_{k=1}^{3}\sum_{\substack{k'=1\\k'\neq k}}^{3}m_k\varphi_k m_{k'}\varphi_{k'}\left(\left\|\mathbf{a}^H(\boldsymbol{\theta}_k)\mathbf{a}(\boldsymbol{\theta}_{k'})\right\|^2-\left\|\mathbf{a}(\boldsymbol{\theta}_k)\right\|^2\left\|\mathbf{a}(\boldsymbol{\theta}_{k'})\right\|^2\right)\right.\\&-\left(\prod_{k=1}^{3}m_k\varphi_k\right)\left(\prod_{k=1}^{3}\left\|\mathbf{a}(\boldsymbol{\theta}_k)\right\|^2-\frac{1}{2}\sum_{k=1}^{3}\sum_{\substack{k'=1\\k'\neq k}}^{3}\sum_{\substack{k''=1\\k''\neq k'\neq k}}^{3}\left\|\mathbf{a}^H(\boldsymbol{\theta}_k)\mathbf{a}(\boldsymbol{\theta}_{k'})\right\|^2\left\|\mathbf{a}(\boldsymbol{\theta}_{k''})\right\|^2\right.\\&\left.\left.+\mathbf{a}^H(\boldsymbol{\theta}_3)\mathbf{a}(\boldsymbol{\theta}_2)\mathbf{a}^H(\boldsymbol{\theta}_1)\mathbf{a}(\boldsymbol{\theta}_3)\mathbf{a}^H(\boldsymbol{\theta}_2)\mathbf{a}(\boldsymbol{\theta}_1)+\mathbf{a}^H(\boldsymbol{\theta}_3)\mathbf{a}(\boldsymbol{\theta}_1)\mathbf{a}^H(\boldsymbol{\theta}_1)\mathbf{a}(\boldsymbol{\theta}_2)\mathbf{a}^H(\boldsymbol{\theta}_2)\mathbf{a}(\boldsymbol{\theta}_3)\right)\right),\end{aligned} \tag{32}$$

where

$$\varphi_k=\frac{\sigma_s^2}{\sigma_s^2\left\|\mathbf{a}(\boldsymbol{\theta}_k)\right\|^2+\sigma_n^2},\quad k=1,2,3. \tag{33}$$

2) Conditional observation model $\mathcal{M}_2$: Note that the results proposed here are in the context of any number of sources. Under the conditional model, the set of functions $\acute{\eta}_\theta$ given by Eqn. (17) is linked to the function $\zeta_{\boldsymbol{\theta}}(\boldsymbol{\mu},\boldsymbol{\rho})$ given by Eqn. (18). In this analysis, the vector $\boldsymbol{\mu}$ takes the value μ_i at the i^{th} row and zero elsewhere and the vector $\boldsymbol{\rho}$ takes the value ρ_j at the j^{th} row and zero elsewhere (of course, one can has $i=j$). In Appendix D, the calculus of the following closed-form expressions for $\zeta_{\boldsymbol{\theta}}(\boldsymbol{\mu},\boldsymbol{\rho})$ are detailed.

- If $(m-1)\,p+1 \leq i,j \leq mp$, where p denotes the number of parameters per source, then, we have

$$\zeta_{\boldsymbol{\theta}}(\boldsymbol{\mu},\boldsymbol{\rho}) = \sum_{t=1}^{T} \|\{\mathbf{s}(t)\}_m\|^2 \sum_{i=1}^{M}\sum_{j=1}^{M} \left\{\mathbf{R}_{\mathbf{n}}^{-1}\right\}_{i,j} \exp\left(j\frac{2\pi}{\lambda}\left(\mathbf{r}_j^T - \mathbf{r}_i^T\right)\boldsymbol{\theta}_m\right) \times \left(\exp\left(-j\frac{2\pi}{\lambda}\mathbf{r}_i^T\boldsymbol{\mu}_m\right) - \exp\left(-j\frac{2\pi}{\lambda}\mathbf{r}_i^T\boldsymbol{\rho}_m\right)\right)\left(\exp\left(j\frac{2\pi}{\lambda}\mathbf{r}_j^T\boldsymbol{\mu}_m\right) - \exp\left(j\frac{2\pi}{\lambda}\mathbf{r}_j^T\boldsymbol{\rho}_m\right)\right). \tag{34}$$

- Otherwise, if $(m-1)\,p+1 \leq i \leq mp$ and $(n-1)\,p+1 \leq j \leq np$, then we have

$$\begin{aligned}
&\zeta_{\boldsymbol{\theta}}(\boldsymbol{\mu},\boldsymbol{\rho}) = \\
&\sum_{t=1}^{T} \|\{\mathbf{s}(t)\}_m\|^2 \sum_{i=1}^{M}\sum_{j=1}^{M} \left\{\mathbf{R}_{\mathbf{n}}^{-1}\right\}_{i,j} \exp\left(j\frac{2\pi}{\lambda}\left(\mathbf{r}_j^T - \mathbf{r}_i^T\right)\boldsymbol{\theta}_m\right) \exp\left(-j\frac{2\pi}{\lambda}\mathbf{r}_i^T\boldsymbol{\mu}_m\right) \exp\left(j\frac{2\pi}{\lambda}\mathbf{r}_j^T\boldsymbol{\mu}_m\right) \\
&+\sum_{t=1}^{T} \|\{\mathbf{s}(t)\}_n\|^2 \sum_{i=1}^{M}\sum_{j=1}^{M} \left\{\mathbf{R}_{\mathbf{n}}^{-1}\right\}_{i,j} \exp\left(j\frac{2\pi}{\lambda}\left(\mathbf{r}_j^T - \mathbf{r}_i^T\right)\boldsymbol{\theta}_n\right) \exp\left(-j\frac{2\pi}{\lambda}\mathbf{r}_i^T\boldsymbol{\rho}_n\right) \exp\left(j\frac{2\pi}{\lambda}\mathbf{r}_j^T\boldsymbol{\rho}_n\right) \\
&-2\,\mathrm{Re}\left(\sum_{t=1}^{T} \{\mathbf{s}(t)\}_m^* \{\mathbf{s}(t)\}_n \right. \\
&\qquad \left. \times \sum_{i=1}^{M}\sum_{j=1}^{M} \left\{\mathbf{R}_{\mathbf{n}}^{-1}\right\}_{i,j} \exp\left(j\frac{2\pi}{\lambda}\left(\mathbf{r}_j^T\boldsymbol{\theta}_n - \mathbf{r}_i^T\boldsymbol{\theta}_m\right)\right) \exp\left(-j\frac{2\pi}{\lambda}\mathbf{r}_i^T\boldsymbol{\mu}_m\right) \exp\left(j\frac{2\pi}{\lambda}\mathbf{r}_j^T\boldsymbol{\rho}_n\right)\right).
\end{aligned} \tag{35}$$

In particular, if one assumes $\mathbf{R}_n = \sigma_n^2\mathbf{I}$, then, several simplifications can be done:

- If $(m-1)\,p+1 \leq i,j \leq mp$, then

$$\zeta_{\boldsymbol{\theta}}(\boldsymbol{\mu},\boldsymbol{\rho}) = \frac{1}{\sigma_n^2}\sum_{i=1}^{M}\left\|\exp\left(-j\frac{2\pi}{\lambda}\mathbf{r}_i^T\boldsymbol{\mu}_m\right) - \exp\left(-j\frac{2\pi}{\lambda}\mathbf{r}_i^T\boldsymbol{\rho}_m\right)\right\|^2 \sum_{t=1}^{T}\|\{\mathbf{s}(t)\}_m\|^2, \tag{36}$$

  where we note that the function $\zeta_{\boldsymbol{\theta}}(\boldsymbol{\mu},\boldsymbol{\rho})$ does not depend on the parameter $\boldsymbol{\theta}$.

- Otherwise, if $(m-1)\,p+1 \leq i \leq mp$ and $(n-1)\,p+1 \leq j \leq np$, then

$$\begin{aligned}
\zeta_{\boldsymbol{\theta}}(\boldsymbol{\mu},\boldsymbol{\rho}) &= \frac{1}{\sigma_n^2}\sum_{i=1}^{M}\left\|\exp\left(-j\frac{2\pi}{\lambda}\mathbf{r}_i^T\boldsymbol{\mu}_m\right)\right\|^2 \sum_{t=1}^{T}\|\{\mathbf{s}(t)\}_m\|^2 + \frac{1}{\sigma_n^2}\sum_{i=1}^{M}\left\|\exp\left(-j\frac{2\pi}{\lambda}\mathbf{r}_i^T\boldsymbol{\rho}_n\right)\right\|^2 \sum_{t=1}^{T}\|\{\mathbf{s}(t)\}_n\|^2 \\
&-2\,\mathrm{Re}\left(\frac{1}{\sigma_n^2}\sum_{i=1}^{M}\exp\left(j\frac{2\pi}{\lambda}\mathbf{r}_i^T\left(\boldsymbol{\theta}_n - \boldsymbol{\theta}_m\right)\right)\exp\left(-j\frac{2\pi}{\lambda}\mathbf{r}_i^T\boldsymbol{\mu}_m\right)\exp\left(j\frac{2\pi}{\lambda}\mathbf{r}_i^T\boldsymbol{\rho}_n\right)\sum_{t=1}^{T}\{\mathbf{s}(t)\}_m^*\{\mathbf{s}(t)\}_n\right)
\end{aligned} \tag{37}$$

It is clear that the proposed above formulas for both the unconditional and the conditional models can be applied to any kind of array geometry and whatever the number of sources. However, they generally depend on the parameter vector $\boldsymbol{\theta}$. This means that, in general, the calculus of the set of functions η will have to be performed numerically (except if one is able to find a closed-form expression of Eqn. (11)). In the following we present a kind of array geometry where, fortunately, the set of functions $\acute{\eta}_\theta$ will not depend on $\boldsymbol{\theta}$ leading to a straightforward calculus of the bound.

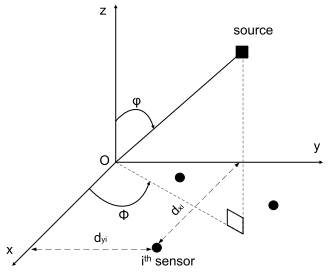


Fig. 1. 3D source localization using a planar array antenna.

B. 3D Source localization with a planar array

We first consider the problem of DOA estimation of a single narrow band source in the far field area by using an arbitrary planar array. In fact, we start by this general setting because the non-uniform linear array is clearly a particular case of this array. Without loss of generality, we assume that the sensors of this array lay on the xOy plan with Cartesian coordinates (see Fig. 1). Therefore, the vector $\mathbf{r}_i$ contains the coordinate of the i^{th} sensor position with respect to this referential, *i.e.*, $\mathbf{r}_i = [d_{x_i} \;\; d_{y_i}]^T$, $i = 1, \ldots, M$. From (28), the steering vector is given by

$$\mathbf{a}(\boldsymbol{\theta}) = \left[\exp\left(j\frac{2\pi}{\lambda}(d_{x_1}u + d_{y_1}v)\right) \ldots \exp\left(j\frac{2\pi}{\lambda}(d_{x_M}u + d_{y_M}v)\right)\right]^T, \tag{38}$$

where, as in [23], the parameter vector of interest is $\boldsymbol{\theta} = [u \quad v]^T$ where

$$\begin{cases} u = \sin\varphi\cos\phi, \\ v = \sin\varphi\sin\phi, \end{cases} \tag{39}$$

and where φ and ϕ represent the elevation and azimuth angles of the source, respectively. The parameters space is such that $u \in [-1, 1]$ and $v \in [-1, 1]$. Therefore, we assume that they both follow a uniform distribution over $[-1, 1]$. Note that from a physical point of view, it should be more tempting to choose a uniform prior for φ and ϕ. This will lead to a probability density functions for u and v not uniform. To the best of our knowledge, this assumption has only been used in the context of lower bounds in [25]. Unfortunately, such prior leads to an untractable expression of the bound (see Eqn. (21) of [25]). Consequently, other authors have generally not specified the prior leading to semi closed-form expressions of bounds (*i.e.* that it remains a numerical integration to perform over the parameters) [23] [27]. On the other hand, in order to obtain a closed-form expression, authors have generally used a simplified assumption, *i.e.* a uniform prior directly on u and v (see, for example, [26] [42]). In this paper, we have followed the same way by expecting a slight modification of performance with respect to a more physical model and in order to be able to get closed-form expressions of the bound.

We choose the matrix of test points such that

$$\mathbf{H} = [\mathbf{h}_u \quad \mathbf{h}_v] = \begin{bmatrix} h_u & 0 \\ 0 & h_v \end{bmatrix}. \tag{40}$$

Then, we have: $\boldsymbol{\theta}+\mathbf{h_u} = [u+h_u \quad v]^T$ and $\boldsymbol{\theta}+\mathbf{h_v} = [u \quad v+h_v]^T$. Moreover, we now have two elements $s_i \in [0,1]\,,\ i=1,2$ for which we will prefer the notation s_u and s_v, respectively.

1) Unconditional observation model $\mathcal{M}_1$: Under $\mathcal{M}_1$, let us set $U_{SNR} = \frac{\sigma_s^4}{\sigma_n^2(M\sigma_s^2+\sigma_n^2)}$. The closed-form expressions of the elements of matrix $\mathbf{G} = \begin{bmatrix} \{\mathbf{G}\}_{uu} & \{\mathbf{G}\}_{uv} \\ \{\mathbf{G}\}_{vu} & \{\mathbf{G}\}_{vv} \end{bmatrix}$ are given by (see Appendix E for the proof):

$$\{\mathbf{G}\}_{uu} = \frac{\left(\begin{array}{l} \left(1-\frac{|h_u|}{2}\right)\left(1+2s_u(1-2s_u)U_{SNR}\left(M^2-\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right\|^2\right)\right)^{-T} \\ +\left(1-\frac{|h_u|}{2}\right)\left(1+2(1-s_u)(2s_u-1)U_{SNR}\left(M^2-\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right\|^2\right)\right)^{-T} \\ -2\left(1-|h_u|\right)\left(1+s_u(1-s_u)U_{SNR}\left(M^2-\left\|\sum_{k=1}^{M}\exp\left(-j\frac{4\pi}{\lambda}d_{x_k}h_u\right)\right\|^2\right)\right)^{-T} \end{array} \right)}{\left(1-\frac{|h_u|}{2}\right)^2\left(1+s_u(1-s_u)U_{SNR}\left(M^2-\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right\|^2\right)\right)^{-2T}}, \tag{41}$$

$$\{\mathbf{G}\}_{vv} = \frac{\left(\begin{array}{l} \left(1-\frac{|h_v|}{2}\right)\left(1+2s_v(1-2s_v)U_{SNR}\left(M^2-\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right\|^2\right)\right)^{-T} \\ +\left(1-\frac{|h_v|}{2}\right)\left(1+2(1-s_v)(2s_v-1)U_{SNR}\left(M^2-\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right\|^2\right)\right)^{-T} \\ -2\left(1-|h_v|\right)\left(1+s_v(1-s_v)U_{SNR}\left(M^2-\left\|\sum_{k=1}^{M}\exp\left(-j\frac{4\pi}{\lambda}d_{y_k}h_v\right)\right\|^2\right)\right)^{-T} \end{array} \right)}{\left(1-\frac{|h_v|}{2}\right)^2\left(1+s_v(1-s_v)U_{SNR}\left(M^2-\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right\|^2\right)\right)^{-2T}}, \tag{42}$$

$$\{\mathbf{G}\}_{uv} = \frac{\left(\begin{array}{l} \left(\begin{array}{l} 1 - U_{SNR} \left(\begin{array}{l} s_u s_v \left(\left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda}(d_{x_k} h_u - d_{y_k} h_v)\right) \right\|^2 - M^2 \right) \\ +s_u(1-s_u-s_v) \left(\left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda} d_{x_k} h_u\right) \right\|^2 - M^2 \right) \\ +s_v(1-s_u-s_v) \left(\left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda} d_{y_k} h_v\right) \right\|^2 - M^2 \right) \end{array} \right) \\ -s_u s_v (1-s_u-s_v) \frac{U_{SNR}^2 \sigma_n^2}{\sigma_s^2} \times \\ \times \left(\begin{array}{l} \sum_{k=1}^{M} \exp\left(j\frac{2\pi d_{y_k} h_v}{\lambda}\right) \sum_{k=1}^{M} \exp\left(-j\frac{2\pi d_{x_k} h_u}{\lambda}\right) \sum_{k=1}^{M} \exp\left(j\frac{2\pi (d_{x_k} h_u - d_{y_k} h_v)}{\lambda}\right) \\ +\sum_{k=1}^{M} \exp\left(-j\frac{2\pi d_{y_k} h_v}{\lambda}\right) \sum_{k=1}^{M} \exp\left(j\frac{2\pi d_{x_k} h_u}{\lambda}\right) \sum_{k=1}^{M} \exp\left(-j\frac{2\pi (d_{x_k} h_u - d_{y_k} h_v)}{\lambda}\right) \\ -M \left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda} d_{y_k} h_v\right) \right\|^2 - M \left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda} d_{x_k} h_u\right) \right\|^2 \\ -M \left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda}(d_{x_k} h_u - d_{y_k} h_v)\right) \right\|^2 + M^3 \end{array} \right) \end{array} \right)^{-T} \\ + \left(\begin{array}{l} 1 - U_{SNR} \left(\begin{array}{l} (1-s_u)(1-s_v) \left(\left\| \sum_{k=1}^{M} \exp\left(j\frac{2\pi}{\lambda}(d_{x_k} h_u - d_{y_k} h_v)\right) \right\|^2 - M^2 \right) \\ +(1-s_u)(s_u+s_v-1) \left(\left\| \sum_{k=1}^{M} \exp\left(j\frac{2\pi}{\lambda} d_{x_k} h_u\right) \right\|^2 - M^2 \right) \\ +(1-s_v)(s_u+s_v-1) \left(\left\| \sum_{k=1}^{M} \exp\left(j\frac{2\pi}{\lambda} d_{y_k} h_v\right) \right\|^2 - M^2 \right) \end{array} \right) \\ -(1-s_u)(1-s_v)(s_u+s_v-1) \frac{U_{SNR}^2 \sigma_n^2}{\sigma_s^2} \times \\ \times \left(\begin{array}{l} \sum_{k=1}^{M} \exp\left(j\frac{2\pi d_{y_k} h_v}{\lambda}\right) \sum_{k=1}^{M} \exp\left(-j\frac{2\pi d_{x_k} h_u}{\lambda}\right) \sum_{k=1}^{M} \exp\left(j\frac{2\pi (d_{x_k} h_u - d_{y_k} h_v)}{\lambda}\right) \\ +\sum_{k=1}^{M} \exp\left(-j\frac{2\pi d_{y_k} h_v}{\lambda}\right) \sum_{k=1}^{M} \exp\left(j\frac{2\pi d_{x_k} h_u}{\lambda}\right) \sum_{k=1}^{M} \exp\left(-j\frac{2\pi (d_{x_k} h_u - d_{y_k} h_v)}{\lambda}\right) \\ -M \left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda} d_{y_k} h_v\right) \right\|^2 - M \left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda} d_{x_k} h_u\right) \right\|^2 \\ -M \left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda}(d_{x_k} h_u - d_{y_k} h_v)\right) \right\|^2 + M^3 \end{array} \right) \end{array} \right)^{-T} \\ - \left(\begin{array}{l} 1 - U_{SNR} \left(\begin{array}{l} s_u(1-s_v) \left(\left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda}(d_{x_k} h_u + d_{y_k} h_v)\right) \right\|^2 - M^2 \right) \\ +s_u(s_v-s_u) \left(\left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda} d_{x_k} h_u\right) \right\|^2 - M^2 \right) \\ +(1-s_v)(s_v-s_u) \left(\left\| \sum_{k=1}^{M} \exp\left(j\frac{2\pi}{\lambda} d_{y_k} h_v\right) \right\|^2 - M^2 \right) \end{array} \right) \\ -s_u(1-s_v)(s_v-s_u) \frac{U_{SNR}^2 \sigma_n^2}{\sigma_s^2} \times \\ \times \left(\begin{array}{l} \sum_{k=1}^{M} \exp\left(j\frac{2\pi d_{y_k} h_v}{\lambda}\right) \sum_{k=1}^{M} \exp\left(j\frac{2\pi d_{x_k} h_u}{\lambda}\right) \sum_{k=1}^{M} \exp\left(-j\frac{2\pi (d_{x_k} h_u + d_{y_k} h_v)}{\lambda}\right) \\ +\sum_{k=1}^{M} \exp\left(-j\frac{2\pi d_{y_k} h_v}{\lambda}\right) \sum_{k=1}^{M} \exp\left(-j\frac{2\pi d_{x_k} h_u}{\lambda}\right) \sum_{k=1}^{M} \exp\left(j\frac{2\pi (d_{x_k} h_u + d_{y_k} h_v)}{\lambda}\right) \\ -M \left\| \sum_{k=1}^{M} \exp\left(j\frac{2\pi}{\lambda} d_{y_k} h_v\right) \right\|^2 - M \left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda} d_{x_k} h_u\right) \right\|^2 \\ -M \left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda}(d_{x_k} h_u + d_{y_k} h_v)\right) \right\|^2 + M^3 \end{array} \right) \end{array} \right)^{-T} \\ - \left(\begin{array}{l} 1 - U_{SNR} \left(\begin{array}{l} s_v(1-s_u) \left(\left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda}(d_{x_k} h_u + d_{y_k} h_v)\right) \right\|^2 - M^2 \right) \\ +s_v(s_u-s_v) \left(\left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda} d_{x_k} h_u\right) \right\|^2 - M^2 \right) \\ +(1-s_u)(s_u-s_v) \left(\left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda} d_{y_k} h_v\right) \right\|^2 - M^2 \right) \end{array} \right) \\ -s_v(1-s_u)(s_u-s_v) \frac{U_{SNR}^2 \sigma_n^2}{\sigma_s^2} \times \\ \times \left(\begin{array}{l} \sum_{k=1}^{M} \exp\left(j\frac{2\pi d_{y_k} h_v}{\lambda}\right) \sum_{k=1}^{M} \exp\left(j\frac{2\pi d_{x_k} h_u}{\lambda}\right) \sum_{k=1}^{M} \exp\left(-j\frac{2\pi (d_{x_k} h_u + d_{y_k} h_v)}{\lambda}\right) \\ +\sum_{k=1}^{M} \exp\left(-j\frac{2\pi d_{y_k} h_v}{\lambda}\right) \sum_{k=1}^{M} \exp\left(-j\frac{2\pi d_{x_k} h_u}{\lambda}\right) \sum_{k=1}^{M} \exp\left(j\frac{2\pi (d_{x_k} h_u + d_{y_k} h_v)}{\lambda}\right) \\ -M \left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda} d_{y_k} h_v\right) \right\|^2 - M \left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda} d_{x_k} h_u\right) \right\|^2 \\ -M \left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda}(d_{x_k} h_u + d_{y_k} h_v)\right) \right\|^2 + M^3 \end{array} \right) \end{array} \right)^{-T} \end{array} \right)}{\left(1 + s_u(1-s_u) U_{SNR} \left(M^2 - \left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda} d_{x_k} h_u\right) \right\|^2 \right) \right)^{-T} \left(1 + s_v(1-s_v) U_{SNR} \left(M^2 - \left\| \sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda} d_{y_k} h_v\right) \right\|^2 \right) \right)^{-T}}, \tag{43}$$

and, of course, $\{\mathbf{G}\}_{uv} = \{\mathbf{G}\}_{vu}$. Consequently, the unconditional Weiss-Weinstein bound is 2×2 matrix given by:

$$\begin{aligned}\mathbf{UWWB} &= \mathbf{H}\mathbf{G}^{-1}\mathbf{H}^T \\ &= \frac{1}{\{\mathbf{G}\}_{uu}\{\mathbf{G}\}_{vv} - \{\mathbf{G}\}_{uv}^2} \begin{bmatrix} h_u^2\{\mathbf{G}\}_{vv} & -h_u h_v \{\mathbf{G}\}_{uv} \\ -h_u h_v \{\mathbf{G}\}_{uv} & h_v^2 \{\mathbf{G}\}_{uu} \end{bmatrix},\end{aligned} \tag{44}$$

which has to be optimized over s_u, s_v, h_u, and h_v. Concerning the optimization over s_u and s_v, several other works in the literature have suggested to simply use $s_u = s_v = 1/2$. Most of the time, numerical simulations of this simplified bound compared with the bound obtained after optimization over s_u and s_v leads to the same results while their is no formal proof of this fact (see [9] page 41 footnote 17). Note that, thanks to the expressions obtained in the next Section concerning the linear array, we will be able to prove that $s = 1/2$ is a (maybe not unique) correct choice for any linear array. In the case of the planar array treated in this Section, we will only check this property by simulation.

In the particular case where $s_u = s_v = 1/2$ one obtains the following simplified expressions

$$\{\mathbf{G}\}_{uu} = \frac{2\left(1-\frac{|h_u|}{2}\right) - 2\left(1-|h_u|\right)\left(1+\frac{U_{SNR}}{4}\left(M^2 - \left\|\sum_{k=1}^{M}\exp\left(-j\frac{4\pi}{\lambda}d_{x_k}h_u\right)\right\|^2\right)\right)^{-T}}{\left(1-\frac{|h_u|}{2}\right)^2\left(1+\frac{U_{SNR}}{4}\left(M^2 - \left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right\|^2\right)\right)^{-2T}}, \tag{45}$$

$$\{\mathbf{G}\}_{vv} = \frac{2\left(1-\frac{|h_v|}{2}\right) - 2\left(1-|h_v|\right)\left(1+\frac{U_{SNR}}{4}\left(M^2 - \left\|\sum_{k=1}^{M}\exp\left(-j\frac{4\pi}{\lambda}d_{y_k}h_v\right)\right\|^2\right)\right)^{-T}}{\left(1-\frac{|h_v|}{2}\right)^2\left(1+\frac{U_{SNR}}{4}\left(M^2 - \left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right\|^2\right)\right)^{-2T}}, \tag{46}$$

and

$$\{\mathbf{G}\}_{uv} = \frac{\begin{pmatrix} 2\left(1+\frac{U_{SNR}}{4}\left(M^2 - \left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}(d_{x_k}h_u - d_{y_k}h_v)\right)\right\|^2\right)\right)^{-T} \\ -2\left(1+\frac{U_{SNR}}{4}\left(M^2 - \left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}(d_{x_k}h_u + d_{y_k}h_v)\right)\right\|^2\right)\right)^{-T} \end{pmatrix}}{\left(1+\frac{U_{SNR}}{4}\left(M^2 - \left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right\|^2\right)\right)^{-T}\left(1+\frac{U_{SNR}}{4}\left(M^2 - \left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right\|^2\right)\right)^{-T}}. \tag{47}$$

Again, the Weiss-Weinstein bound is obtained by using the above expressions in Eqn. (44) and after an optimization over the test points. The optimization over the test points can be done over a search grid or by using the ambiguity diagram of the array in order to reduce significantly the computational cost (see [18], [27], [34], [43], [44]).

2) Conditional observation model $\mathcal{M}_2$: Under $\mathcal{M}_2$, let us set $C_{SNR} = \frac{1}{\sigma_n^2}\sum_{t=1}^{T}\|s(t)\|^2$. The closed-form expressions of the elements of matrix $\mathbf{G}$ are given by (see Appendix F for the proof):

$$\{\mathbf{G}\}_{uu} = \frac{\left(\begin{array}{l}\left(1-\frac{|h_u|}{2}\right)\exp\left(4s_u(2s_u-1)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right)\right)\\ +\left(1-\frac{|h_u|}{2}\right)\exp\left(4(2s_u-1)(s_u-1)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right)\right)\\ -2(1-|h_u|)\exp\left(2s_u(s_u-1)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{4\pi}{\lambda}d_{x_k}h_u\right)\right)\right)\end{array}\right)}{\left(1-\frac{|h_u|}{2}\right)^2\exp\left(4s_u(s_u-1)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right)\right)}, \tag{48}$$

$$\{\mathbf{G}\}_{vv} = \frac{\left(\begin{array}{l}\left(1-\frac{|h_v|}{2}\right)\exp\left(4s_v(2s_v-1)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right)\right)\\ +\left(1-\frac{|h_v|}{2}\right)\exp\left(4(2s_v-1)(s_v-1)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right)\right)\\ -2(1-|h_v|)\exp\left(2s_v(s_v-1)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{4\pi}{\lambda}d_{y_k}h_v\right)\right)\right)\end{array}\right)}{\left(1-\frac{|h_v|}{2}\right)^2\exp\left(4s_v(s_v-1)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right)\right)}, \tag{49}$$

$$\{\mathbf{G}\}_{uv} = \frac{\left(\begin{array}{l}\exp\left(\begin{array}{l}2s_u(s_u+s_v-1)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right)\\ +2s_v(s_u+s_v-1)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right)\\ -2s_us_vC_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}(d_{x_k}h_u-d_{y_k}h_v)\right)\right)\end{array}\right)\\ +\exp\left(\begin{array}{l}2(s_u-1)(s_u+s_v-1)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right)\\ +2(s_v-1)(s_u+s_v-1)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right)\\ -2(1-s_u)(1-s_v)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}(d_{x_k}h_u-d_{y_k}h_v)\right)\right)\end{array}\right)\\ -\exp\left(\begin{array}{l}2s_u(s_u-s_v)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right)\\ +2(1-s_v)(s_u-s_v)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right)\\ +2s_u(s_v-1)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}(d_{x_k}h_u+d_{y_k}h_v)\right)\right)\end{array}\right)\\ -\exp\left(\begin{array}{l}2(s_u-1)(s_u-s_v)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right)\\ +2s_v(s_v-s_u)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right)\\ +2(s_u-1)s_vC_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}(d_{x_k}h_u+d_{y_k}h_v)\right)\right)\end{array}\right)\end{array}\right)}{\exp\left(2s_u(s_u-1)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right)\right)\exp\left(2s_v(s_v-1)C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right)\right)}, \tag{50}$$

and $\{\mathbf{G}\}_{uv} = \{\mathbf{G}\}_{vu}$. Consequently, the conditional Weiss-Weinstein bound is 2×2 matrix given by using the above equations in Eqn. (44). As for the unconditional case, if we set $s_u = s_v = 1/2$, one obtains the following simplified expressions

$$\{\mathbf{G}\}_{uu} = \frac{2\left(1-\frac{|h_u|}{2}\right) - 2(1-|h_u|)\exp\left(-\frac{C_{SNR}}{2}\left(M - \sum_{k=1}^{M}\cos\left(\frac{4\pi}{\lambda}d_{x_k}h_u\right)\right)\right)}{\left(1-\frac{|h_u|}{2}\right)^2\exp\left(-C_{SNR}\left(M - \sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right)\right)}, \tag{51}$$

$$\{\mathbf{G}\}_{vv} = \frac{2\left(1-\frac{|h_v|}{2}\right) - 2(1-|h_v|)\exp\left(-\frac{C_{SNR}}{2}\left(M - \sum_{k=1}^{M}\cos\left(\frac{4\pi}{\lambda}d_{y_k}h_v\right)\right)\right)}{\left(1-\frac{|h_v|}{2}\right)^2\exp\left(-C_{SNR}\left(M - \sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right)\right)}, \tag{52}$$

$$\{\mathbf{G}\}_{uv} = \frac{\left(\begin{array}{l} 2\exp\left(-\frac{C_{SNR}}{2}\left(M - \sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}(d_{x_k}h_u - d_{y_k}h_v)\right)\right)\right) \\ -2\exp\left(-\frac{C_{SNR}}{2}\left(M - \sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}(d_{x_k}h_u + d_{y_k}h_v)\right)\right)\right) \end{array}\right)}{\exp\left(-\frac{C_{SNR}}{2}\left(2M - \sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{x_k}h_u\right) - \sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right)\right)}. \tag{53}$$

By using the above expressions in Eqn. (44) and after an optimization over the test points, one obtains the Weiss-Weinstein bound.

C. Source localization with a non-uniform linear array

We now briefly consider the DOA estimation of a single narrow band source in the far area by using a non-uniform linear array antenna. Without loss of generality, let us assume that the linear array antenna lays on the Ox axis of the coordinate system (see Fig. 1), consequently, $d_{y_i} = 0,\ \forall i$. The sensor positions vector is denoted $[d_{x_1} \dots d_{x_M}]$. By letting $\theta = \sin\varphi$, where φ denotes the elevation angle of the source, the steering vector is then given by

$$\mathbf{a}(\theta) = \left[\exp\left(j\frac{2\pi}{\lambda}d_{x_1}\theta\right) \dots \exp\left(j\frac{2\pi}{\lambda}d_{x_M}\theta\right)\right]^T. \tag{54}$$

We assume that the parameter θ follows a uniform distribution over $[0,1]$. As in Section IV-B and since the parameter of interest is a scalar, matrix $\mathbf{H}$ of the test points becomes a scalar denoted h_θ. In the same way, there is only one element $s_i \in [0,1]$ which will be simply denoted s. The closed-form expressions given here are straightforwardly obtained from the aforementioned results on the planar array about the element denoted $\{\mathbf{G}\}_{uu}$. We will continue to use the previously introduced notations $U_{SNR} = \frac{\sigma_s^4}{\sigma_n^2(M\sigma_s^2+\sigma_n^2)}$ and $C_{SNR} = \frac{1}{\sigma_n^2}\sum_{t=1}^{T}\|s(t)\|^2$.

1) Unconditional observation model $\mathcal{M}_1$: The closed-form expression of the unconditional Weiss-Weinstein bound, denoted $UWWB$, is given by

$$UWWB = \frac{h_\theta^2\left(1-\frac{|h_\theta|}{2}\right)^2\left(1+s(1-s)U_{SNR}\left(M^2 - \left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_\theta\right)\right\|^2\right)\right)^{-2T}}{\left(\begin{array}{l} \left(1-\frac{|h_\theta|}{2}\right)\left(\begin{array}{l}\left(1+2s(1-2s)U_{SNR}\left(M^2 - \left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_\theta\right)\right\|^2\right)\right)^{-T} \\ +\left(1+2(1-s)(2s-1)U_{SNR}\left(M^2 - \left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_\theta\right)\right\|^2\right)\right)^{-T}\end{array}\right) \\ -2\left(1-|h_\theta|\right)\left(1+s(1-s)U_{SNR}\left(M^2 - \left\|\sum_{k=1}^{M}\exp\left(-j\frac{4\pi}{\lambda}d_{x_k}h_\theta\right)\right\|^2\right)\right)^{-T} \end{array}\right)}. \tag{55}$$

In order to find one optimal value of s that maximizes $\mathbf{H}\mathbf{G}^{-1}\mathbf{H}^T$, $\forall h_\theta$ we have considered the derivative of $\mathbf{H}\mathbf{G}^{-1}\mathbf{H}^T$ w.r.t. s. The calculus (not reported here) is straightforward and it is easy to see that $\frac{\partial \mathbf{H}\mathbf{G}^{-1}\mathbf{H}^T}{\partial s}\Big|_{s=\frac{1}{2}} = 0$. Consequently, the Weiss-Weinstein bound has just to be optimized over h_θ and is simplified leading to

$$UWWB = \sup_{h_\theta} \frac{h_\theta^2 \left(1-\frac{|h_\theta|}{2}\right)^2 \left(1+\frac{U_{SNR}}{4}\left(M^2 - \left\|\sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_\theta\right)\right\|^2\right)\right)^{-2T}}{2\left(1-\frac{|h_\theta|}{2}\right) - 2\left(1-|h_\theta|\right)\left(1+\frac{U_{SNR}}{4}\left(M^2 - \left\|\sum_{k=1}^{M} \exp\left(-j\frac{4\pi}{\lambda}d_{x_k}h_\theta\right)\right\|^2\right)\right)^{-T}}. \quad (56)$$

In the classical case of a uniform linear array (*i.e.*, $d_{x_k} = d$), this expression can be still simplified by noticing that $\sum_{k=1}^{M} \exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_\theta\right) = M\exp\left(-j\frac{2\pi d}{\lambda}h_\theta\right)$.

2) Conditional observation model $\mathcal{M}_2$: The closed-form expression of the conditional Weiss-Weinstein bound $CWWB$ is given by

$$CWWB = \frac{h_\theta^2 \left(1-\frac{|h_\theta|}{2}\right)^2 \exp\left(4s(s-1)C_{SNR}\left(M - \sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{x_k}h_\theta\right)\right)\right)}{\left(\begin{array}{l}\left(1-\frac{|h_\theta|}{2}\right)\left(\begin{array}{l}\exp\left(4s(2s-1)C_{SNR}\left(M - \sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{x_k}h_\theta\right)\right)\right)\\ +\exp\left(4(2s-1)(s-1)C_{SNR}\left(M - \sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{x_k}h_\theta\right)\right)\right)\end{array}\right)\\ -2\left(1-|h_\theta|\right)\exp\left(2s(s-1)C_{SNR}\left(M - \sum_{k=1}^{M}\cos\left(\frac{4\pi}{\lambda}d_{x_k}h_\theta\right)\right)\right)\end{array}\right)}. \quad (57)$$

Again, it is easy to check that $\frac{\partial \mathbf{H}\mathbf{G}^{-1}\mathbf{H}^T}{\partial s}\Big|_{s=\frac{1}{2}} = 0$. Consequently, one optimal value of s that maximizes $\mathbf{H}\mathbf{G}^{-1}\mathbf{H}^T$, $\forall h_\theta$ is $s = \frac{1}{2}$. The Weiss-Weinstein bound is then simplified as follows

$$CWWB = \sup_{h_\theta} \frac{h_\theta^2 \left(1-\frac{|h_\theta|}{2}\right)^2 \exp\left(-C_{SNR}\left(M - \sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{x_k}h_\theta\right)\right)\right)}{2\left(1-\frac{|h_\theta|}{2}\right) - 2\left(1-|h_\theta|\right)\exp\left(-\frac{1}{2}C_{SNR}\left(M - \sum_{k=1}^{M}\cos\left(\frac{4\pi}{\lambda}d_{x_k}h_\theta\right)\right)\right)}. \quad (58)$$

In the classical case of a uniform linear array (*i.e.*, $d_{x_k} = d$), this expression can be still simplified by noticing that $\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{x_k}h_\theta\right) = M\cos\left(\frac{2\pi d}{\lambda}h_\theta\right)$.

VI. Simulation results and analysis

As an illustration of the previously derived results, we first consider the scenario proposed in [23] Fig. 5, *i.e.*, the DOA estimation under the unconditional model using an uniform circular array consisting of $M = 16$ sensors with a half-wavelength inter-sensors spacing. The numbers of snapshots is $T = 100$. Since the array is symmetric, the performance estimation concerning the parameters u and v are the same, this is why only the performance with respect to the parameters u is given in Fig. 2. The Weiss-Weinstein bound is computed using Eqn. (45), (46) and (47). The Ziv-Zakai bound is computed using Eqn. (24) in [23]. The empirical global MSE (MSE) of the maximum *a posteriori* (MAP) estimator is obtained over 2000 Monte Carlo trials. As in [23] Fig. (1b), one observes that both the Weiss-Weinstein bound and the Ziv-Zakai bound are tight w.r.t. the MSE of the MAP and capture the SNR threshold. Note that, in [23] Fig. (1b), the Weiss-Weinstein bound was computed numerically only.

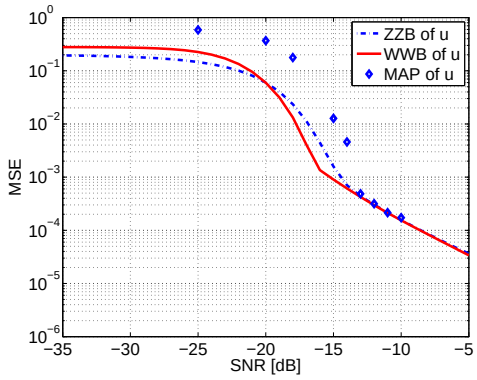


Fig. 2. Ziv-Zakai bound, Weiss-Weinstein bound and empirical MSE of the MAP estimator: unconditional case.

To the best of our knowledge, their are no closed-form expressions of the Ziv-Zakai bound for the conditional model available in the literature. In this case, we consider 3D source localization using a V-shaped array. Indeed, it has been shown that this kind of array is able to outperform other classical planar arrays, more particularly the uniform circular array [45]. This array is made from two branches of uniform linear arrays with 6 sensors located on each branches and one sensor located at the origin. We denote Δ the angle between these two branches. The sensors are equally spaced with a half-wavelength. The number of snapshots is $T = 20$. Fig. 3 shows the behavior of the Weiss-Weinstein bound with respect to the opening angle Δ. One can observe that when Δ varies, the estimation performance concerning the estimation of parameter u varies slightly. On the contrary, the estimation performance concerning the estimation of parameter v is strongly dependent on Δ. When Δ increases from 0° to 90°, the Weiss-Weinstein bound of v decreases, as well as the SNR threshold. Fig. 3 also shows that $\Delta = 90^\circ$ is the optimal value, which is different with the optimal value $\Delta = 53.13^\circ$ in [45] since the assumptions concerning the source signal are not the same.

VII. Conclusion

In this paper, the Weiss-Weinstein bound on the MSE has been studied in the array processing context. In order to analyze the unconditional and conditional signal source models, the structure of the bound has been detailed for both Gaussian observation models with parameterized mean or parameterized covariance matrix.

Appendix

A. Closed-form expression of $\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v})$ under the Gaussian observation model with parameterized covariance

Since $\mathbf{y}(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}))$, one has,

$$\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v}) = \frac{|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta})|^{T(\alpha+\beta-1)}}{\pi^{MT} |\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}+\mathbf{u})|^{T\alpha} |\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}+\mathbf{v})|^{T\beta}} \int_{\Omega} \exp\left(-\sum_{t=1}^{T} \mathbf{y}^H(t) \boldsymbol{\Gamma}^{-1} \mathbf{y}(t)\right) d\mathbf{Y}, \tag{59}$$

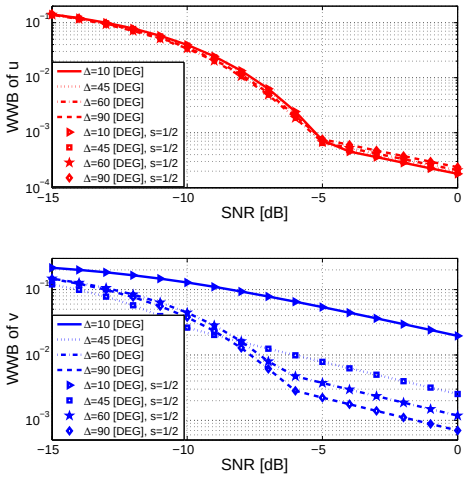


Fig. 3. Weiss-Weinstein bounds of the V-shaped array w.r.t. the opening angle Δ.

where $\mathbf{\Gamma}^{-1} = \alpha \mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}+\mathbf{u}) + \beta \mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}+\mathbf{v}) - (\alpha+\beta-1)\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta})$. Then, since

$$\int_{\Omega} \exp\left\{-\sum_{t=1}^{T} \mathbf{y}^H(t)\mathbf{\Gamma}^{-1}\mathbf{y}(t)\right\} d\mathbf{Y} = \pi^{MT} \left|\mathbf{\Gamma}\right|^T, \tag{60}$$

one has

$$\acute{\eta}_{\boldsymbol{\theta}}(\alpha,\beta,\mathbf{u},\mathbf{v}) = \frac{\left|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta})\right|^{T(\alpha+\beta-1)} \left|\mathbf{\Gamma}\right|^T}{\left|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}+\mathbf{u})\right|^{T\alpha} \left|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}+\mathbf{v})\right|^{T\beta}} = \frac{\left|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta})\right|^{T(\alpha+\beta-1)}}{\left|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}+\mathbf{u})\right|^{T\alpha} \left|\mathbf{R}_{\mathbf{y}}(\boldsymbol{\theta}+\mathbf{v})\right|^{T\beta} \left|\mathbf{\Gamma}^{-1}\right|^T} \tag{61}$$

B. Closed-form expression of $\acute{\eta}_{\boldsymbol{\theta}}(\alpha,\beta,\mathbf{u},\mathbf{v})$ under the Gaussian observation model with parameterized mean

Since $\mathbf{y}(t) \sim \mathcal{CN}(\mathbf{f}(\boldsymbol{\theta}), \mathbf{R}_{\mathbf{y}})$, one has

$$\acute{\eta}_{\boldsymbol{\theta}}(\alpha,\beta,\mathbf{u},\mathbf{v}) = \frac{1}{\pi^{MT}\left|\mathbf{R}_{\mathbf{y}}\right|^T} \int_{\Omega} \exp\left(-\sum_{t=1}^{T} \xi(t)\right) d\mathbf{Y}, \tag{62}$$

with[2]

$$\begin{aligned}
\xi(t) &= \alpha\left(\mathbf{y}(t) - \mathbf{f}(\boldsymbol{\theta}+\mathbf{u})\right)^H \mathbf{R}_{\mathbf{y}}^{-1}\left(\mathbf{y}(t) - \mathbf{f}(\boldsymbol{\theta}+\mathbf{u})\right) + \beta\left(\mathbf{y}(t) - \mathbf{f}(\boldsymbol{\theta}+\mathbf{v})\right)^H \mathbf{R}_{\mathbf{y}}^{-1}\left(\mathbf{y}(t) - \mathbf{f}(\boldsymbol{\theta}+\mathbf{v})\right) \\
&\quad + (1-\alpha-\beta)\left(\mathbf{y}(t) - \mathbf{f}(\boldsymbol{\theta})\right)^H \mathbf{R}_{\mathbf{y}}^{-1}\left(\mathbf{y}(t) - \mathbf{f}(\boldsymbol{\theta})\right) \\
&= \mathbf{y}(t)^H \mathbf{R}_{\mathbf{y}}^{-1}\mathbf{y}(t) + \alpha \mathbf{f}^H(\boldsymbol{\theta}+\mathbf{u})\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{f}(\boldsymbol{\theta}+\mathbf{u}) + \beta \mathbf{f}^H(\boldsymbol{\theta}+\mathbf{v})\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{f}(\boldsymbol{\theta}+\mathbf{v}) + (1-\alpha-\beta)\mathbf{f}^H(\boldsymbol{\theta})\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{f}(\boldsymbol{\theta}) \\
&\quad - 2\,\mathrm{Re}\left\{\mathbf{y}(t)^H \mathbf{R}_{\mathbf{y}}^{-1}\left(\alpha \mathbf{f}(\boldsymbol{\theta}+\mathbf{u}) + \beta \mathbf{f}(\boldsymbol{\theta}+\mathbf{v}) + (1-\alpha-\beta)\mathbf{f}(\boldsymbol{\theta})\right)\right\}.
\end{aligned} \tag{63}$$

[2]For simplicity, the dependance on t of $\mathbf{f}$ and $\mathbf{y}$ is not emphasized.

Let us set $\mathbf{x}(t) = \mathbf{y}(t) - (\alpha \mathbf{f}(\boldsymbol{\theta}+\mathbf{u}) + \beta \mathbf{f}(\boldsymbol{\theta}+\mathbf{v}) + (1-\alpha-\beta)\mathbf{f}(\boldsymbol{\theta}))$. Consequently,

$$\begin{aligned}\mathbf{x}(t)^H \mathbf{R}_{\mathbf{y}}^{-1} \mathbf{x}(t) = \mathbf{y}(t)^H \mathbf{R}_{\mathbf{y}}^{-1} \mathbf{y}(t) - 2\,\mathrm{Re}\left\{\mathbf{y}(t)^H \mathbf{R}_{\mathbf{y}}^{-1} \left(\alpha \mathbf{f}(\boldsymbol{\theta}+\mathbf{u}) + \beta \mathbf{f}(\boldsymbol{\theta}+\mathbf{v}) + (1-\alpha-\beta)\mathbf{f}(\boldsymbol{\theta})\right)\right\} \\ + \left(\alpha \mathbf{f}^H(\boldsymbol{\theta}+\mathbf{u}) + \beta \mathbf{f}^H(\boldsymbol{\theta}+\mathbf{v}) + (1-\alpha-\beta)\mathbf{f}^H(\boldsymbol{\theta})\right) \mathbf{R}_{\mathbf{y}}^{-1} \left(\alpha \mathbf{f}(\boldsymbol{\theta}+\mathbf{u}) + \beta \mathbf{f}(\boldsymbol{\theta}+\mathbf{v}) + (1-\alpha-\beta)\mathbf{f}(\boldsymbol{\theta})\right).\end{aligned} \tag{64}$$

And $\xi(t)$ can be rewritten as

$$\xi(t) = \mathbf{x}(t)^H \mathbf{R}_{\mathbf{y}}^{-1} \mathbf{x}(t) + \acute{\xi}(t), \tag{65}$$

where

$$\begin{aligned}\acute{\xi}(t) = \alpha(1-\alpha)\mathbf{f}^H(\boldsymbol{\theta}+\mathbf{u})\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{f}(\boldsymbol{\theta}+\mathbf{u}) + \beta(1-\beta)\mathbf{f}^H(\boldsymbol{\theta}+\mathbf{v})\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{f}(\boldsymbol{\theta}+\mathbf{v}) \\ + (1-\alpha-\beta)(\alpha+\beta)\mathbf{f}^H(\boldsymbol{\theta})\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{f}(\boldsymbol{\theta}) - 2\,\mathrm{Re}\left\{\alpha\beta \mathbf{f}^H(\boldsymbol{\theta}+\mathbf{u})\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{f}(\boldsymbol{\theta}+\mathbf{v}) \right. \\ \left. + \alpha(1-\alpha-\beta)\mathbf{f}^H(\boldsymbol{\theta}+\mathbf{u})\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{f}(\boldsymbol{\theta}) + \beta(1-\alpha-\beta)\mathbf{f}^H(\boldsymbol{\theta}+\mathbf{v})\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{f}(\boldsymbol{\theta})\right\}.\end{aligned} \tag{66}$$

Note that $\acute{\xi}(t)$ is independent of $\mathbf{x}(t)$. By defining $\mathbf{X} = [\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(T)]$, the function $\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v})$ becomes

$$\acute{\eta}_{\boldsymbol{\theta}}(\alpha, \beta, \mathbf{u}, \mathbf{v}) = \frac{1}{\pi^{MT}|\mathbf{R}_{\mathbf{y}}|^T} \int_{\Omega} \exp\left(-\sum_{t=1}^{T} \mathbf{x}(t)^H \mathbf{R}_{\mathbf{y}}^{-1} \mathbf{x}(t) + \acute{\xi}(t)\right) d\mathbf{X} = \exp\left(-\sum_{t=1}^{T} \acute{\xi}(t)\right), \tag{67}$$

since $\frac{1}{\pi^{MT}|\mathbf{R}_{\mathbf{y}}|^T} \int_{\Omega} \exp\left(-\sum_{t=1}^{T} \mathbf{x}(t)^H \mathbf{R}_{\mathbf{y}}^{-1} \mathbf{x}(t)\right) d\mathbf{X} = 1$.

C. Closed-form expressions of $\left|m_1 \mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}_1) + m_2 \mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}_2)\right|$ *and* $\left|m_1 \mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}_1) + m_2 \mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}_2) + m_3 \mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}_3)\right|$

Note that this calculus is actually an extension of the result obtained in [27] Appendix A in which $m_1 = m_2 = \frac{1}{2}$ and $m_3 = 0$, but follows the same method. The inverse of $\mathbf{R}_{\mathbf{y}}$ can be deduced from the Woodbury formula

$$\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}) = \frac{1}{\sigma_n^2}\left(\mathbf{I}_M - \frac{\sigma_s^2 \mathbf{a}(\boldsymbol{\theta})\mathbf{a}^H(\boldsymbol{\theta})}{\sigma_s^2 \|\mathbf{a}(\boldsymbol{\theta})\|^2 + \sigma_n^2}\right).$$

Then,

$$\sum_{k=1}^{3} m_k \mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}_k) = \frac{1}{\sigma_n^2} \sum_{k=1}^{3} m_k \left(\mathbf{I} - \frac{\sigma_s^2 \mathbf{a}(\boldsymbol{\theta}_k)\mathbf{a}^H(\boldsymbol{\theta}_k)}{\sigma_s^2 \|\mathbf{a}(\boldsymbol{\theta}_k)\|^2 + \sigma_n^2}\right). \tag{68}$$

Since the rank of $\mathbf{a}(\boldsymbol{\theta}_k)\mathbf{a}^H(\boldsymbol{\theta}_k)$ is equal to 1 and since $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2 \neq \boldsymbol{\theta}_3$ (except for $\mathbf{h}_k = \mathbf{h}_l = \mathbf{0}$), the above matrix has $M-3$ eigenvalues equal to $\frac{1}{\sigma_n^2}\sum_{k=1}^{3} m_k$ and 3 eigenvalues corresponding to the eigenvectors made from the linear combination of $\mathbf{a}(\boldsymbol{\theta}_1)$, $\mathbf{a}(\boldsymbol{\theta}_2)$, and $\mathbf{a}(\boldsymbol{\theta}_3)$: $\mathbf{a}(\boldsymbol{\theta}_1) + p\mathbf{a}(\boldsymbol{\theta}_2) + q\mathbf{a}(\boldsymbol{\theta}_3)$. The determinant will then be the product of these M eigenvalues[3]. Let us set

$$\varphi_k = \frac{\sigma_s^2}{\sigma_s^2 \|\mathbf{a}(\boldsymbol{\theta}_k)\|^2 + \sigma_n^2}, \quad k = 1, 2, 3. \tag{69}$$

[3]Note that we are only interested by the eigenvalues. Consequently, the linear combination of of $\mathbf{a}(\boldsymbol{\theta}_1)$, $\mathbf{a}(\boldsymbol{\theta}_2)$, and $\mathbf{a}(\boldsymbol{\theta}_3)$ can be written $\mathbf{a}(\boldsymbol{\theta}_1) + p\mathbf{a}(\boldsymbol{\theta}_2) + q\mathbf{a}(\boldsymbol{\theta}_3)$ instead of $r\mathbf{a}(\boldsymbol{\theta}_1) + p\mathbf{a}(\boldsymbol{\theta}_2) + q\mathbf{a}(\boldsymbol{\theta}_3)$

Then, the three aforementioned eigenvalues denoted λ must satisfy:

$$\left(\sum_{k=1}^{3} m_k \mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}_k)\right)\left(\mathbf{a}(\boldsymbol{\theta}_1)+p\mathbf{a}(\boldsymbol{\theta}_2)+q\mathbf{a}(\boldsymbol{\theta}_3)\right)=\lambda\left(\mathbf{a}(\boldsymbol{\theta}_1)+p\mathbf{a}(\boldsymbol{\theta}_2)+q\mathbf{a}(\boldsymbol{\theta}_3)\right). \tag{70}$$

By using Eqn. (68) in the above equation and after a factorization with respect to $\mathbf{a}(\boldsymbol{\theta}_1)$, $\mathbf{a}(\boldsymbol{\theta}_2)$, and $\mathbf{a}(\boldsymbol{\theta}_3)$ one obtains

$$\begin{aligned}&\left(x-m_1\varphi_1\left\|\mathbf{a}(\boldsymbol{\theta}_1)\right\|^2-pm_1\varphi_1\mathbf{a}^H(\boldsymbol{\theta}_1)\mathbf{a}(\boldsymbol{\theta}_2)-qm_1\varphi_1\mathbf{a}^H(\boldsymbol{\theta}_1)\mathbf{a}(\boldsymbol{\theta}_3)\right)\mathbf{a}(\boldsymbol{\theta}_1)\\&+\left(-m_2\varphi_2\mathbf{a}^H(\boldsymbol{\theta}_2)\mathbf{a}(\boldsymbol{\theta}_1)+p\left(x-m_2\varphi_2\left\|\mathbf{a}(\boldsymbol{\theta}_2)\right\|^2\right)-qm_2\varphi_2\mathbf{a}^H(\boldsymbol{\theta}_2)\mathbf{a}(\boldsymbol{\theta}_3)\right)\mathbf{a}(\boldsymbol{\theta}_2)\\&+\left(-m_3\varphi_3\mathbf{a}^H(\boldsymbol{\theta}_3)\mathbf{a}(\boldsymbol{\theta}_1)-m_3\varphi_3p\mathbf{a}^H(\boldsymbol{\theta}_3)\mathbf{a}(\boldsymbol{\theta}_2)+q\left(x-m_3\varphi_3\left\|\mathbf{a}(\boldsymbol{\theta}_3)\right\|^2\right)\right)\mathbf{a}(\boldsymbol{\theta}_3)=0,\end{aligned} \tag{71}$$

where[4]

$$x=1-\sigma_n^2\lambda. \tag{72}$$

Consequently, the coefficients of $\mathbf{a}(\boldsymbol{\theta}_1)$, $\mathbf{a}(\boldsymbol{\theta}_2)$, and $\mathbf{a}(\boldsymbol{\theta}_3)$ are equals to zero leading to a system of three equations with two unknown (p and q). Solving the two first equations to find[5] p and q, and applying the solution into the last equation, one obtains the following polynomial equation of x

$$\begin{aligned}&x^3-x^2\sum_{k=1}^{3}m_k\varphi_k\left\|\mathbf{a}(\boldsymbol{\theta}_k)\right\|^2-\frac{x}{2}\sum_{k=1}^{3}\sum_{\substack{k'=1\\k'\neq k}}^{3}m_k\varphi_k m_{k'}\varphi_{k'}\left(\left|\mathbf{a}^H(\boldsymbol{\theta}_k)\mathbf{a}(\boldsymbol{\theta}_{k'})\right|^2-\left\|\mathbf{a}(\boldsymbol{\theta}_k)\right\|^2\left\|\mathbf{a}(\boldsymbol{\theta}_{k'})\right\|^2\right)\\&-m_1m_2m_3\varphi_1\varphi_2\varphi_3\left(\left\|\mathbf{a}(\boldsymbol{\theta}_1)\right\|^2\left\|\mathbf{a}(\boldsymbol{\theta}_2)\right\|^2\left\|\mathbf{a}(\boldsymbol{\theta}_3)\right\|^2-\left|\mathbf{a}^H(\boldsymbol{\theta}_2)\mathbf{a}(\boldsymbol{\theta}_3)\right|^2\left\|\mathbf{a}(\boldsymbol{\theta}_1)\right\|^2\right.\\&-\left|\mathbf{a}^H(\boldsymbol{\theta}_1)\mathbf{a}(\boldsymbol{\theta}_2)\right|^2\left\|\mathbf{a}(\boldsymbol{\theta}_3)\right\|^2-\left|\mathbf{a}^H(\boldsymbol{\theta}_3)\mathbf{a}(\boldsymbol{\theta}_1)\right|^2\left|\mathbf{a}^H(\boldsymbol{\theta}_2)\right\|^2+\mathbf{a}^H(\boldsymbol{\theta}_3)\mathbf{a}(\boldsymbol{\theta}_2)\mathbf{a}^H(\boldsymbol{\theta}_1)\mathbf{a}(\boldsymbol{\theta}_3)\mathbf{a}^H(\boldsymbol{\theta}_2)\mathbf{a}(\boldsymbol{\theta}_1)\\&\left.+\mathbf{a}^H(\boldsymbol{\theta}_3)\mathbf{a}(\boldsymbol{\theta}_1)\mathbf{a}^H(\boldsymbol{\theta}_1)\mathbf{a}(\boldsymbol{\theta}_2)\mathbf{a}^H(\boldsymbol{\theta}_2)\mathbf{a}(\boldsymbol{\theta}_3)\right)=0\end{aligned}$$

Since we are only interested by the product of the three eigenvalues, we do not have to solve this polynomial in λ and only the opposite of the last term is required. This leads to Eqn. (31) with $\sum_{k=1}^{3} m_k = 1$. Of course, the closed-form expression of $\left|m_1\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}_1)+m_2\mathbf{R}_{\mathbf{y}}^{-1}(\boldsymbol{\theta}_2)\right|$ is obtained by letting $m_3=0$ and $\sum_{k=1}^{2} m_k = 1$ in Eqn. (32).

D. Closed-form expressions of $\zeta_{\boldsymbol{\theta}}(\boldsymbol{\mu},\boldsymbol{\rho})$

Remind that the function $\zeta_{\boldsymbol{\theta}}(\boldsymbol{\mu},\boldsymbol{\rho})$ is defined by Eqn. (18). Let us define p as the number of parameters per sources (assumed to be constant for each sources). Then, without loss of generality, the full parameter vector $\boldsymbol{\theta}$

[4]Note that, from Eqn. (16), $\sum_{k=1}^{3} m_k = 1$.

[5]p and q are given by

$$p=\frac{m_2\varphi_2\mathbf{a}^H(\boldsymbol{\theta}_2)\left(m_1\varphi_1\mathbf{a}(\boldsymbol{\theta}_1)\mathbf{a}^H(\boldsymbol{\theta}_1)+\left(x-m_1\varphi_1\left\|\mathbf{a}(\boldsymbol{\theta}_1)\right\|^2\right)\mathbf{I}\right)\mathbf{a}(\boldsymbol{\theta}_3)}{m_1\varphi_1\mathbf{a}^H(\boldsymbol{\theta}_1)\left(m_2\varphi_2\mathbf{a}(\boldsymbol{\theta}_2)\mathbf{a}^H(\boldsymbol{\theta}_2)+\left(x-m_2\varphi_2\left\|\mathbf{a}(\boldsymbol{\theta}_2)\right\|^2\right)\mathbf{I}\right)\mathbf{a}(\boldsymbol{\theta}_3)}, \tag{73}$$

and

$$q=\frac{\left(x-m_1\varphi_1\left\|\mathbf{a}(\boldsymbol{\theta}_1)\right\|^2\right)\left(x-m_2\varphi_2\left\|\mathbf{a}(\boldsymbol{\theta}_2)\right\|^2\right)-m_1\varphi_1m_2\varphi_2\mathbf{a}^H(\boldsymbol{\theta}_1)\mathbf{a}(\boldsymbol{\theta}_2)\mathbf{a}^H(\boldsymbol{\theta}_2)\mathbf{a}(\boldsymbol{\theta}_1)}{m_1\varphi_1\mathbf{a}^H(\boldsymbol{\theta}_1)\left(m_2\varphi_2\mathbf{a}(\boldsymbol{\theta}_2)\mathbf{a}^H(\boldsymbol{\theta}_2)+\left(x-m_2\varphi_2\left\|\mathbf{a}(\boldsymbol{\theta}_2)\right\|^2\right)\mathbf{I}\right)\mathbf{a}(\boldsymbol{\theta}_3)}. \tag{74}$$

can be decomposed as $\boldsymbol{\theta} = \left[\boldsymbol{\theta}_1^T \dots \boldsymbol{\theta}_N^T\right]^T$ where $\boldsymbol{\theta}_i = [\theta_{i,1} \dots \theta_{i,p}]^T$, $i = 1, \dots, N$ with $q = Np$. Remind that $\boldsymbol{\mu} = [0 \dots \mu_i \dots 0]^T$ and $\boldsymbol{\rho} = [0 \dots \rho_j \dots 0]^T$. It exists two distinct cases to study: when both index i and j are such that $(m-1)p+1 \leq i \leq mp$, $m = 1, \dots, N$ and $(m-1)p+1 \leq j \leq mp$ or when $(m-1)p+1 \leq i \leq mp$, $m = 1, \dots, N$ and $(n-1)p+1 \leq j \leq np$, $n = 1, \dots, N$ with $m \neq n$. Therefore let us denote:

$$\begin{cases} \boldsymbol{\mu}_m = [0 \cdots 0 \quad h_i \quad 0 \cdots 0]^T \in \mathbb{R}^p \\ \boldsymbol{\rho}_m = [0 \cdots 0 \quad h_j \quad 0 \cdots 0]^T \in \mathbb{R}^p \end{cases} \text{ if } (m-1)\,p+1 \leq i,j \leq mp \tag{75}$$

and

$$\begin{cases} \boldsymbol{\mu}_m = [0 \cdots 0 \quad h_i \quad 0 \cdots 0]^T \in \mathbb{R}^p, \\ \boldsymbol{\rho}_n = [0 \cdots 0 \quad h_j \quad 0 \cdots 0]^T \in \mathbb{R}^p, \end{cases} \text{ if } \begin{cases} (m-1)\,p+1 \leq i \leq mp \\ (n-1)\,p+1 \leq j \leq np \end{cases}, \text{ with } m \neq n. \tag{76}$$

1) The case where $(m-1)\,p+1 \leq i,j \leq mp$*:* In this case, one has:

$$\mathbf{A}(\boldsymbol{\theta}+\boldsymbol{\mu}) - \mathbf{A}(\boldsymbol{\theta}+\boldsymbol{\rho}) = [\mathbf{0} \cdots \mathbf{0} \quad \mathbf{a}(\boldsymbol{\theta}_m+\boldsymbol{\mu}_m) - \mathbf{a}(\boldsymbol{\theta}_m+\boldsymbol{\rho}_m) \quad \mathbf{0} \cdots \mathbf{0}] \in \mathbb{C}^{p \times N}, \tag{77}$$

and consequently,

$$\zeta_{\boldsymbol{\theta}}(\boldsymbol{\mu},\boldsymbol{\rho}) = \left\| \mathbf{R}_{\mathbf{n}}^{-1/2} \left(\mathbf{a}(\boldsymbol{\theta}_m+\boldsymbol{\mu}_m) - \mathbf{a}(\boldsymbol{\theta}_m+\boldsymbol{\rho}_m) \right) \right\|^2 \sum_{t=1}^{T} \left\| \{\mathbf{s}(t)\}_m \right\|^2. \tag{78}$$

Due to Eqn. (28), one has

$$\left\| \mathbf{R}_{\mathbf{n}}^{-1/2} \left(\mathbf{a}(\boldsymbol{\theta}_m+\boldsymbol{\mu}_m) - \mathbf{a}(\boldsymbol{\theta}_m+\boldsymbol{\rho}_m) \right) \right\|^2 =$$

$$\sum_{i=1}^{M}\sum_{j=1}^{M} \left\{\mathbf{R}_{\mathbf{n}}^{-1}\right\}_{i,j} \exp\left(j\frac{2\pi}{\lambda} \left(\mathbf{r}_j^T - \mathbf{r}_i^T\right) \boldsymbol{\theta}_m \right) \left(\exp\left(-j\frac{2\pi}{\lambda} \mathbf{r}_i^T \boldsymbol{\mu}_m \right) - \exp\left(-j\frac{2\pi}{\lambda} \mathbf{r}_i^T \boldsymbol{\rho}_m \right) \right)$$
$$\times \left(\exp\left(j\frac{2\pi}{\lambda} \mathbf{r}_j^T \boldsymbol{\mu}_m \right) - \exp\left(j\frac{2\pi}{\lambda} \mathbf{r}_j^T \boldsymbol{\rho}_m \right) \right). \tag{79}$$

In particular, in the case where $\mathbf{R}_{\mathbf{n}} = \sigma_n^2 \mathbf{I}$ one obtains

$$\left\| \mathbf{R}_{\mathbf{n}}^{-1/2} \left(\mathbf{a}(\boldsymbol{\theta}_m+\boldsymbol{\mu}_m) - \mathbf{a}(\boldsymbol{\theta}_m+\boldsymbol{\rho}_m) \right) \right\|^2 = \frac{1}{\sigma_n^2} \sum_{i=1}^{M} \left\| \exp\left(-j\frac{2\pi}{\lambda} \mathbf{r}_i^T \boldsymbol{\mu}_m \right) - \exp\left(-j\frac{2\pi}{\lambda} \mathbf{r}_i^T \boldsymbol{\rho}_m \right) \right\|^2. \tag{80}$$

2) The case where $(m-1)\,p+1 \leq i \leq mp$ *and where* $(n-1)\,p+1 \leq j \leq np$*:* Without loss generality, we assume that $n > m$. Then,

$$\begin{aligned} \mathbf{A}(\boldsymbol{\theta}+\boldsymbol{\mu}) - \mathbf{A}(\boldsymbol{\theta}+\boldsymbol{\rho}) &= [\mathbf{a}(\boldsymbol{\theta}_1) - \mathbf{a}(\boldsymbol{\theta}_1) \cdots \mathbf{a}(\boldsymbol{\theta}_m+\boldsymbol{\mu}_m) - \mathbf{a}(\boldsymbol{\theta}_m) \cdots \mathbf{a}(\boldsymbol{\theta}_n) - \mathbf{a}(\boldsymbol{\theta}_n+\boldsymbol{\rho}_n) \cdots \mathbf{a}(\boldsymbol{\theta}_N) - \mathbf{a}(\boldsymbol{\theta}_N)] \\ &= [\mathbf{0} \cdots \mathbf{0} \quad \mathbf{a}(\boldsymbol{\theta}_m+\boldsymbol{\mu}_m) - \mathbf{a}(\boldsymbol{\theta}_m) \quad \mathbf{0} \cdots \mathbf{0} \quad \mathbf{a}(\boldsymbol{\theta}_m) - \mathbf{a}(\boldsymbol{\theta}_n+\boldsymbol{\rho}_n) \quad \mathbf{0} \cdots \mathbf{0}], \end{aligned} \tag{81}$$

and consequently,

$$\zeta_{\boldsymbol{\theta}}(\boldsymbol{\mu},\boldsymbol{\rho}) = \sum_{t=1}^{T} \left\| \mathbf{R}_{\mathbf{n}}^{-1/2} \left(\mathbf{a}(\boldsymbol{\theta}_m+\boldsymbol{\mu}_m) - \mathbf{a}(\boldsymbol{\theta}_m) \right) \{\mathbf{s}(t)\}_m + \left(\mathbf{a}(\boldsymbol{\theta}_n) - \mathbf{a}(\boldsymbol{\theta}_n+\boldsymbol{\rho}_n) \right) \{\mathbf{s}(t)\}_n \right\|^2. \tag{82}$$

Let us set $\varkappa = \mathbf{R}_{\mathbf{n}}^{-1/2}\left(\mathbf{a}\left(\boldsymbol{\theta}_m+\boldsymbol{\mu}_m\right)-\mathbf{a}\left(\boldsymbol{\theta}_m\right)\right)$and $\boldsymbol{\varrho} = \mathbf{R}_{\mathbf{n}}^{-1/2}\left(\mathbf{a}\left(\boldsymbol{\theta}_n\right)-\mathbf{a}\left(\boldsymbol{\theta}_n+\boldsymbol{\rho}_n\right)\right)$. Then, $\zeta_{\boldsymbol{\theta}}\left(\boldsymbol{\mu},\boldsymbol{\rho}\right)$ can be rewritten

$$\begin{aligned}\zeta_{\boldsymbol{\theta}}\left(\boldsymbol{\mu},\boldsymbol{\rho}\right) &= \sum_{t=1}^{T}\left\|\varkappa\left\{\mathbf{s}\left(t\right)\right\}_m+\boldsymbol{\varrho}\left\{\mathbf{s}\left(t\right)\right\}_n\right\|^2 \\ &= \sum_{t=1}^{T}\left(\varkappa^H\varkappa\left\|\left\{\mathbf{s}\left(t\right)\right\}_m\right\|^2+\varkappa^H\boldsymbol{\varrho}\left\{\mathbf{s}\left(t\right)\right\}_m^*\left\{\mathbf{s}\left(t\right)\right\}_n+\boldsymbol{\varrho}^H\varkappa\left\{\mathbf{s}\left(t\right)\right\}_m\left\{\mathbf{s}\left(t\right)\right\}_n^*+\boldsymbol{\varrho}^H\boldsymbol{\varrho}\left\|\left\{\mathbf{s}\left(t\right)\right\}_n\right\|^2\right) \\ &= \varkappa^H\varkappa\sum_{t=1}^{T}\left\|\left\{\mathbf{s}\left(t\right)\right\}_m\right\|^2+\boldsymbol{\varrho}^H\boldsymbol{\varrho}\sum_{t=1}^{T}\left\|\left\{\mathbf{s}\left(t\right)\right\}_n\right\|^2+2\,\mathrm{Re}\left(\varkappa^H\boldsymbol{\varrho}\sum_{t=1}^{T}\left\{\mathbf{s}\left(t\right)\right\}_m^*\left\{\mathbf{s}\left(t\right)\right\}_n\right). \end{aligned} \tag{83}$$

By using the structure of the steering matrix $\mathbf{A}$, it leads to

$$\left\{\begin{array}{l} \varkappa^H\varkappa = \sum_{i=1}^{M}\sum_{j=1}^{M}\left\{\mathbf{R}_{\mathbf{n}}^{-1}\right\}_{i,j}\exp\left(j\frac{2\pi}{\lambda}\left(\mathbf{r}_j^T-\mathbf{r}_i^T\right)\boldsymbol{\theta}_m\right)\exp\left(-j\frac{2\pi}{\lambda}\mathbf{r}_i^T\boldsymbol{\mu}_m\right)\exp\left(j\frac{2\pi}{\lambda}\mathbf{r}_j^T\boldsymbol{\mu}_m\right), \\ \boldsymbol{\varrho}^H\boldsymbol{\varrho} = \sum_{i=1}^{M}\sum_{j=1}^{M}\left\{\mathbf{R}_{\mathbf{n}}^{-1}\right\}_{i,j}\exp\left(j\frac{2\pi}{\lambda}\left(\mathbf{r}_j^T-\mathbf{r}_i^T\right)\boldsymbol{\theta}_n\right)\exp\left(-j\frac{2\pi}{\lambda}\mathbf{r}_i^T\boldsymbol{\rho}_n\right)\exp\left(j\frac{2\pi}{\lambda}\mathbf{r}_j^T\boldsymbol{\rho}_n\right), \\ \varkappa^H\boldsymbol{\varrho} = -\sum_{i=1}^{M}\sum_{j=1}^{M}\left\{\mathbf{R}_{\mathbf{n}}^{-1}\right\}_{i,j}\exp\left(j\frac{2\pi}{\lambda}\left(\mathbf{r}_j^T\boldsymbol{\theta}_n-\mathbf{r}_i^T\boldsymbol{\theta}_m\right)\right)\exp\left(-j\frac{2\pi}{\lambda}\mathbf{r}_i^T\boldsymbol{\mu}_m\right)\exp\left(j\frac{2\pi}{\lambda}\mathbf{r}_j^T\boldsymbol{\rho}_n\right). \end{array}\right. \tag{84}$$

E. Proof of Eqn. (41), (42) and (43)

In fact, one only has to prove Eqn. (43) since Eqn. (41) and (42) can be obtained by letting $h_u = h_v$ and $s_u = s_v$ in Eqn. (43) and by using (h_u, s_u) for Eqn. (41) and (h_v, s_v) for Eqn. (42). By plugging Eqn. (30) and (32) into Eqn. (16), and by considering the following expressions

$$\mathbf{a}^H(\boldsymbol{\theta}+\mathbf{h}_u)\mathbf{a}(\boldsymbol{\theta}+\mathbf{h}_v) = \sum_{i=1}^{M}\exp\left(j\tfrac{2\pi}{\lambda}(d_{y_i}h_v-d_{x_i}h_u)\right) = \left(\mathbf{a}^H(\boldsymbol{\theta}+\mathbf{h}_v)\mathbf{a}(\boldsymbol{\theta}+\mathbf{h}_u)\right)^H,$$

$$\mathbf{a}^H(\boldsymbol{\theta}\pm\mathbf{h}_u)\mathbf{a}(\boldsymbol{\theta}) = \sum_{i=1}^{M}\exp\left(\mp j\tfrac{2\pi}{\lambda}d_{x_i}h_u\right), \text{ and } \mathbf{a}^H(\boldsymbol{\theta}+\mathbf{h}_u)\mathbf{a}(\boldsymbol{\theta}-\mathbf{h}_u) = \sum_{i=1}^{M}\exp\left(-j\tfrac{4\pi}{\lambda}d_{x_i}h_u\right),$$

one obtains the closed-form expressions for the set of functions $\acute{\eta}_{\boldsymbol{\theta}}\left(\alpha,\beta,\mathbf{u},\mathbf{v}\right)$

$$\acute{\eta}_{\boldsymbol{\theta}}(s_u,s_v,\mathbf{h}_u,\mathbf{h}_v) = \left(\begin{array}{l} 1-U_{SNR}\left(\begin{array}{l} s_us_v\left(\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}(d_{x_k}h_u-d_{y_k}h_v)\right)\right\|^2-M^2\right) \\ +s_u(1-s_u-s_v)\left(\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right\|^2-M^2\right) \\ +s_v(1-s_u-s_v)\left(\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right\|^2-M^2\right)\end{array}\right) \\ -s_us_v(1-s_u-s_v)\frac{U_{SNR}^2\sigma_n^2}{\sigma_s^2}\times \\ \times\left(\begin{array}{l} \sum_{k=1}^{M}\exp\left(j\frac{2\pi d_{y_k}h_v}{\lambda}\right)\sum_{k=1}^{M}\exp\left(-j\frac{2\pi d_{x_k}h_u}{\lambda}\right)\sum_{k=1}^{M}\exp\left(j\frac{2\pi(d_{x_k}h_u-d_{y_k}h_v)}{\lambda}\right) \\ +\sum_{k=1}^{M}\exp\left(-j\frac{2\pi d_{y_k}h_v}{\lambda}\right)\sum_{k=1}^{M}\exp\left(j\frac{2\pi d_{x_k}h_u}{\lambda}\right)\sum_{k=1}^{M}\exp\left(-j\frac{2\pi(d_{x_k}h_u-d_{y_k}h_v)}{\lambda}\right) \\ -M\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right\|^2-M\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right\|^2 \\ -M\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}(d_{x_k}h_u-d_{y_k}h_v)\right)\right\|^2+M^3\end{array}\right)\end{array}\right)^{-T} \tag{85}$$

26

$$\acute{\eta}_{\boldsymbol{\theta}}(1-s_u,1-s_v,-\mathbf{h}_u,-\mathbf{h}_v)=$$
$$\left(\begin{array}{l}1-U_{SNR}\left(\begin{array}{l}(1-s_u)(1-s_v)\left(\left\|\sum_{k=1}^{M}\exp\left(j\frac{2\pi}{\lambda}(d_{x_k}h_u-d_{y_k}h_v)\right)\right\|^2-M^2\right)\\+(1-s_u)(s_u+s_v-1)\left(\left\|\sum_{k=1}^{M}\exp\left(j\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right\|^2-M^2\right)\\+(1-s_v)(s_u+s_v-1)\left(\left\|\sum_{k=1}^{M}\exp\left(j\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right\|^2-M^2\right)\end{array}\right)\\-(1-s_u)(1-s_v)(s_u+s_v-1)\frac{U_{SNR}^2\sigma_n^2}{\sigma_s^2}\times\\\times\left(\begin{array}{l}\sum_{k=1}^{M}\exp\left(j\frac{2\pi d_{y_k}h_v}{\lambda}\right)\sum_{k=1}^{M}\exp\left(-j\frac{2\pi d_{x_k}h_u}{\lambda}\right)\sum_{k=1}^{M}\exp\left(j\frac{2\pi(d_{x_k}h_u-d_{y_k}h_v)}{\lambda}\right)\\+\sum_{k=1}^{M}\exp\left(-j\frac{2\pi d_{y_k}h_v}{\lambda}\right)\sum_{k=1}^{M}\exp\left(j\frac{2\pi d_{x_k}h_u}{\lambda}\right)\sum_{k=1}^{M}\exp\left(-j\frac{2\pi(d_{x_k}h_u-d_{y_k}h_v)}{\lambda}\right)\\-M\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right\|^2-M\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right\|^2\\-M\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}(d_{x_k}h_u-d_{y_k}h_v)\right)\right\|^2+M^3\end{array}\right)\end{array}\right)^{-T}\quad(86)$$

$$\acute{\eta}_{\boldsymbol{\theta}}(s_u,1-s_v,\mathbf{h}_u,-\mathbf{h}_v)=$$
$$\left(\begin{array}{l}1-U_{SNR}\left(\begin{array}{l}s_u(1-s_v)\left(\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}(d_{x_k}h_u+d_{y_k}h_v)\right)\right\|^2-M^2\right)\\+s_u(s_v-s_u)\left(\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right\|^2-M^2\right)\\+(1-s_v)(s_v-s_u)\left(\left\|\sum_{k=1}^{M}\exp\left(j\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right\|^2-M^2\right)\end{array}\right)\\-s_u(1-s_v)(s_v-s_u)\frac{U_{SNR}^2\sigma_n^2}{\sigma_s^2}\times\\\times\left(\begin{array}{l}\sum_{k=1}^{M}\exp\left(j\frac{2\pi d_{y_k}h_v}{\lambda}\right)\sum_{k=1}^{M}\exp\left(j\frac{2\pi d_{x_k}h_u}{\lambda}\right)\sum_{k=1}^{M}\exp\left(-j\frac{2\pi(d_{x_k}h_u+d_{y_k}h_v)}{\lambda}\right)\\+\sum_{k=1}^{M}\exp\left(-j\frac{2\pi d_{y_k}h_v}{\lambda}\right)\sum_{k=1}^{M}\exp\left(-j\frac{2\pi d_{x_k}h_u}{\lambda}\right)\sum_{k=1}^{M}\exp\left(j\frac{2\pi(d_{x_k}h_u+d_{y_k}h_v)}{\lambda}\right)\\-M\left\|\sum_{k=1}^{M}\exp\left(j\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right\|^2-M\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right\|^2\\-M\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}(d_{x_k}h_u+d_{y_k}h_v)\right)\right\|^2+M^3\end{array}\right)\end{array}\right)^{-T}\quad(87)$$

$$\acute{\eta}_{\boldsymbol{\theta}}(1-s_u,s_v,-\mathbf{h}_u,\mathbf{h}_v)= \left(\begin{array}{l} 1-U_{SNR}\left(\begin{array}{l} s_v(1-s_u)\left(\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}(d_{x_k}h_u+d_{y_k}h_v)\right)\right\|^2-M^2\right)\\ +s_v(s_u-s_v)\left(\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right\|^2-M^2\right)\\ +(1-s_u)(s_u-s_v)\left(\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right\|^2-M^2\right)\end{array}\right)\\ -s_v(1-s_u)(s_u-s_v)\frac{U_{SNR}^2\sigma_n^2}{\sigma_s^2}\times\\ \times\left(\begin{array}{l}\sum_{k=1}^{M}\exp\left(j\frac{2\pi d_{y_k}h_v}{\lambda}\right)\sum_{k=1}^{M}\exp\left(j\frac{2\pi d_{x_k}h_u}{\lambda}\right)\sum_{k=1}^{M}\exp\left(-j\frac{2\pi(d_{x_k}h_u+d_{y_k}h_v)}{\lambda}\right)\\ +\sum_{k=1}^{M}\exp\left(-j\frac{2\pi d_{y_k}h_v}{\lambda}\right)\sum_{k=1}^{M}\exp\left(-j\frac{2\pi d_{x_k}h_u}{\lambda}\right)\sum_{k=1}^{M}\exp\left(j\frac{2\pi(d_{x_k}h_u+d_{y_k}h_v)}{\lambda}\right)\\ -M\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right\|^2-M\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right\|^2\\ -M\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}(d_{x_k}h_u+d_{y_k}h_v)\right)\right\|^2+M^3\end{array}\right)\end{array}\right)^{-T} \tag{88}$$

$$\acute{\eta}_{\boldsymbol{\theta}}(s_u,0,\mathbf{h}_u,\mathbf{0})=\left(1+s_u(1-s_u)U_{SNR}\left(M^2-\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{x_k}h_u\right)\right\|^2\right)\right)^{-T}, \tag{89}$$

and

$$\acute{\eta}_{\boldsymbol{\theta}}(0,s_v,\mathbf{0},\mathbf{h}_v)=\left(1+s_v(1-s_v)U_{SNR}\left(M^2-\left\|\sum_{k=1}^{M}\exp\left(-j\frac{2\pi}{\lambda}d_{y_k}h_v\right)\right\|^2\right)\right)^{-T}. \tag{90}$$

One notices that the set of functions $\acute{\eta}_{\boldsymbol{\theta}}(\alpha,\beta,\mathbf{u},\mathbf{v})$ does not depend on $\boldsymbol{\theta}$. Consequently, it is also easy to obtain the Weiss-Weinstein bound (throughout the set of functions $\eta(\alpha,\beta,\mathbf{u},\mathbf{v})$) by using the results of Section IV-B whatever the considered prior on θ (only the integral $\int_\Theta \frac{p^{\alpha+\beta}(\theta+u)}{p^{\alpha+\beta-1}(\theta)}d\theta$ has to be calculated or computed numerically). In our case of a uniform prior, the results are straightforward and leads to Eqn. (41), (42) and (43).

F. Proof of Eqn. (48), (49) and (50)

The set of functions $\acute{\eta}_{\boldsymbol{\theta}}(\alpha,\beta,\mathbf{u},\mathbf{v})$ is given by Eqn. (17). So, it only remains the calculus of functions $\zeta_{\boldsymbol{\theta}}(\boldsymbol{\mu},\boldsymbol{\rho})$ from Eqn. (18). Since $\mathbf{R_n}=\sigma_n^2\mathbf{I}$, one obtains

28

$$\begin{cases}
\zeta_{\boldsymbol{\theta}}(\mathbf{h}_u,\mathbf{0}) = \zeta_{\boldsymbol{\theta}}(-\mathbf{h}_u,\mathbf{0}) = 2C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{xk}h_u\right)\right), \\
\zeta_{\boldsymbol{\theta}}(\mathbf{h}_v,\mathbf{0}) = \zeta_{\boldsymbol{\theta}}(-\mathbf{h}_v,\mathbf{0}) = 2C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}d_{yk}h_v\right)\right), \\
\zeta_{\boldsymbol{\theta}}(\mathbf{h}_u,-\mathbf{h}_u) = \zeta_{\boldsymbol{\theta}}(-\mathbf{h}_u,\mathbf{h}_u) = 2C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{4\pi}{\lambda}d_{xk}h_u\right)\right), \\
\zeta_{\boldsymbol{\theta}}(\mathbf{h}_v,-\mathbf{h}_v) = \zeta_{\boldsymbol{\theta}}(-\mathbf{h}_v,\mathbf{h}_v) = 2C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{4\pi}{\lambda}d_{yk}h_v\right)\right), \\
\zeta_{\boldsymbol{\theta}}(\mathbf{h}_u,\mathbf{h}_v) = \zeta_{\boldsymbol{\theta}}(\mathbf{h}_v,\mathbf{h}_u) = \zeta_{\boldsymbol{\theta}}(-\mathbf{h}_u,-\mathbf{h}_v) \\
\qquad = \zeta_{\boldsymbol{\theta}}(-\mathbf{h}_v,-\mathbf{h}_u) = 2C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}(d_{xk}h_u-d_{yk}h_v)\right)\right), \\
\zeta_{\boldsymbol{\theta}}(-\mathbf{h}_u,\mathbf{h}_v) = \zeta_{\boldsymbol{\theta}}(\mathbf{h}_u,-\mathbf{h}_v) = \zeta_{\boldsymbol{\theta}}(\mathbf{h}_v,-\mathbf{h}_u) \\
\qquad = \zeta_{\boldsymbol{\theta}}(-\mathbf{h}_v,\mathbf{h}_u) = 2C_{SNR}\left(M-\sum_{k=1}^{M}\cos\left(\frac{2\pi}{\lambda}(d_{xk}h_u+d_{yk}h_v)\right)\right), \\
\zeta_{\boldsymbol{\theta}}(\mathbf{h}_u,\mathbf{h}_u) = \zeta_{\boldsymbol{\theta}}(\mathbf{h}_v,\mathbf{h}_v) = \zeta_{\boldsymbol{\theta}}(-\mathbf{h}_u,-\mathbf{h}_u) = \zeta_{\boldsymbol{\theta}}(-\mathbf{h}_v,-\mathbf{h}_v) = 0.
\end{cases} \tag{91}$$

Again, since the set of functions $\zeta_{\boldsymbol{\theta}}(\boldsymbol{\mu},\boldsymbol{\rho})$ does not depend on $\boldsymbol{\theta}$, the set of functions $\acute{\eta}_{\boldsymbol{\theta}}(\alpha,\beta,\mathbf{u},\mathbf{v})$ is given by plugging the above equations into Eqn. (17) and does not depend on $\boldsymbol{\theta}$. Consequently, as in unconditional case, the set of functions $\eta(\alpha,\beta,\mathbf{u},\mathbf{v})$ is obtained by using the results of Section IV-B whatever the considered prior on θ. In our case of a uniform prior, the results are straightforward and leads to Eqn. (48), (49) and (50).

REFERENCES

[1] D. T. Vu, A. Renaux, R. Boyer, and S. Marcos, "Closed-form expression of the Weiss-Weinstein bound for 3D source localization: the conditional case," in *Proc. of IEEE Workshop on Sensor Array and Multi-channel Processing (SAM)*, vol. 1, Kibutz Ma'ale Hahamisha, Israel, Oct. 2010, pp. 125–128.

[2] H. Cramér, *Mathematical Methods of Statistics*, ser. Princeton Mathematics. New-York: Princeton University Press, Sep. 1946, vol. 9.

[3] E. W. Barankin, "Locally best unbiased estimates," *The Annals of Mathematical Statistics*, vol. 20, no. 4, pp. 477–501, Dec. 1949.

[4] R. J. McAulay and L. P. Seidman, "A useful form of the Barankin lower bound and its application to PPM threshold analysis," *IEEE Transactions on Information Theory*, vol. 15, no. 2, pp. 273–279, Mar. 1969.

[5] R. J. McAulay and E. M. Hofstetter, "Barankin bounds on parameter estimation," *IEEE Transactions on Information Theory*, vol. 17, no. 6, pp. 669–676, Nov. 1971.

[6] J. S. Abel, "A bound on mean square estimate error," *IEEE Transactions on Information Theory*, vol. 39, no. 5, pp. 1675–1680, Sep. 1993.

[7] E. Chaumette, J. Galy, A. Quinlan, and P. Larzabal, "A new Barankin bound approximation for the prediction of the threshold region performance of maximum likelihood estimators," *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5319–5333, Nov. 2008.

[8] K. Todros and J. Tabrikian, "General classes of performance lower bounds for parameter estimation - part I: non-Bayesian bounds for unbiased estimators," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 5045–5063, Oct. 2010.

[9] H. L. Van Trees and K. L. Bell, Eds., *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. New-York, NY, USA: Wiley/IEEE Press, Sep. 2007.

[10] J. Ziv and M. Zakai, "Some lower bounds on signal parameter estimation," *IEEE Transactions on Information Theory*, vol. 15, no. 3, pp. 386–391, May 1969.

[11] S. Bellini and G. Tartara, "Bounds on error in signal parameter estimation," *IEEE Transactions on Communications*, vol. 22, no. 3, pp. 340–342, Mar. 1974.

[12] K. L. Bell, Y. Steinberg, Y. Ephraim, and H. L. Van Trees, "Extended Ziv-Zakaï lower bound for vector parameter estimation," *IEEE Transactions on Information Theory*, vol. 43, no. 2, pp. 624–637, Mar. 1997.

[13] A. J. Weiss and E. Weinstein, "A lower bound on the mean square error in random parameter estimation," *IEEE Transactions on Information Theory*, vol. 31, no. 5, pp. 680–682, Sep. 1985.

[14] I. Rapoport and Y. Oshman, "WeissWeinstein lower bounds for markovian systems. part I: Theory," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 2016–2030, May 2007.

[15] A. Renaux, P. Forster, P. Larzabal, C. D. Richmond, and A. Nehorai, "A fresh look at the Bayesian bounds of the Weiss-Weinstein family," *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5334–5352, Nov. 2008.

[16] K. Todros and J. Tabrikian, "General classes of performance lower bounds for parameter estimation - part II: Bayesian bounds," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 5064–5082, Oct. 2010.

[17] Y. Rockah and P. Schultheiss, "Array shape calibration using sources in unknown locations–part I: Far-field sources," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 3, pp. 286–299, Mar. 1987.

[18] I. Reuven and H. Messer, "A Barankin-type lower bound on the estimation error of a hybrid parameter vector," *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 1084–1093, May 1997.

[19] S. Bay, B. Geller, A. Renaux, J.-P. Barbot, and J.-M. Brossier, "On the hybrid Cramér-Rao bound and its application to dynamical phase estimation," *IEEE Signal Processing Letters*, vol. 15, pp. 453–456, 2008.

[20] Y. Noam and H. Messer, "Notes on the tightness of the hybrid Cramér-Rao lower bound," *IEEE Transactions on Signal Processing*, vol. 57, no. 6, pp. 2074–2084, 2009.

[21] H. L. Van Trees, *Detection, Estimation and Modulation Theory*. New-York, NY, USA: John Wiley & Sons, 1968, vol. 1.

[22] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, "Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing," in *Radar Array Processing*, S. S. Haykin, J. Litva, and T. J. Shepherd, Eds. Berlin: Springer-Verlag, 1993, ch. 4, pp. 99–151.

[23] K. L. Bell, Y. Ephraim, and H. L. Van Trees, "Explicit Ziv-Zakaï lower bound for bearing estimation," *IEEE Transactions on Signal Processing*, vol. 44, no. 11, pp. 2810–2824, Nov. 1996.

[24] T. J. Nohara and S. Haykin, "Application of the Weiss-Weinstein bound to a two dimensional antenna array," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 9, pp. 1533–1534, Sep. 1988.

[25] H. Nguyen and H. L. Van Trees, "Comparison of performance bounds for DOA estimation," in *Proc. of IEEE Workshop on Statistical Signal and Array Processing (SSAP)*, vol. 1, Jun. 1994, pp. 313–316.

[26] F. Athley, "Optimization of element positions for direction finding with sparse arrays," in *Proc. of IEEE Workshop on Statistical Signal Processing (SSP)*, vol. 1, 2001, pp. 516–519.

[27] W. Xu, A. B. Baggeroer, and C. D. Richmond, "Bayesian bounds for matched-field parameter estimation," *IEEE Transactions on Signal Processing*, vol. 52, no. 12, pp. 3293–3305, Dec. 2004.

[28] A. Renaux, "Weiss-Weinstein bound for data aided carrier estimation," *IEEE Signal Processing Letters*, vol. 14, no. 4, pp. 283–286, Apr. 2007.

[29] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., Mar. 1993, vol. 1.

[30] H. L. Van Trees, *Detection, Estimation and Modulation theory: Optimum Array Processing*. New-York, NY, USA: John Wiley & Sons, Mar. 2002, vol. 4.

[31] Z. Ben Haim and Y. Eldar, "A comment on the Weiss-Weinstein bound for constrained parameter sets," *IEEE Transactions on Information Theory*, vol. 54, no. 10, pp. 4682–4684, Oct. 2008.

[32] P. Stoica and A. Nehorai, "Performances study of conditional and unconditional direction of arrival estimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1783–1795, Oct. 1990.

[33] K. L. Bell, Y. Ephraim, and H. L. Van Trees, "Explicit Ziv-Zakaï lower bounds for bearing estimation using planar arrays," in *Proc. of Workshop on Adaptive Sensor Array Processing (ASAP)*. Lexington, MA, USA: MIT Lincoln Laboratory, Mar. 1996.

[34] I. Reuven and H. Messer, "The use of the Barankin bound for determining the threshold SNR in estimating the bearing of a source in the presence of another," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, Detroit, MI, USA, May 1995, pp. 1645–1648.

[35] J. Li and R. T. Compton, "Maximum likelihood angle estimation for signals with known waveforms," *IEEE Transactions on Signal Processing*, vol. 41, no. 9, pp. 2850–2862, Sep. 93.

[36] M. Cedervall and R. L. Moses, "Efficient maximum likelihood DOA estimation for signals with known waveforms in presence of multipath," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 808–811, Mar. 1997.

30

[37] J. Li, B. Halder, P. Stoica, and M. Viberg, "Computationally efficient angle estimation for signals with known waveforms," *IEEE Transactions on Signal Processing*, vol. 43, no. 9, pp. 2154–2163, Sep. 1995.

[38] A. Leshem and A.-J. Van der Veen, "Direction-of-arrival estimation for constant modulus signals," *IEEE Transactions on Signal Processing*, vol. 47, no. 11, pp. 3125–3129, Nov. 1999.

[39] Y. H. Choi, "Unified approach to Cramér-Rao bounds in direction estimation with known signal structures," *ELSEVIER Signal Processing*, vol. 84, no. 10, pp. 1875–1882, Oct. 2004.

[40] E. Weinstein and A. J. Weiss, "A general class of lower bounds in parameter estimation," *IEEE Transactions on Information Theory*, vol. 34, no. 2, pp. 338–342, Mar. 1988.

[41] P. S. La Rosa, A. Renaux, A. Nehorai, and C. H. Muravchik, "Barankin-type lower bound on multiple change-point estimation," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5534–5549, Nov. 2010.

[42] W. Xu, A. B. Baggeroer, and K. L. Bell, "A bound on mean-square estimation error with background parameter mismatch," *IEEE Transactions on Information Theory*, vol. 50, no. 4, pp. 621–632, Apr. 2004.

[43] J. Tabrikian and J. L. Krolik, "Barankin bounds for source localization in an uncertain ocean environment," *IEEE Transactions on Signal Processing*, vol. 47, no. 11, pp. 2917–2927, Nov. 1999.

[44] A. Renaux, L. N. Atallah, P. Forster, and P. Larzabal, "A useful form of the Abel bound and its application to estimator threshold prediction," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, Part 2, pp. 2365–2369, May 2007.

[45] H. Gazzah and S. Marcos, "Cramér-Rao bounds for antenna array design," *IEEE Transactions on Signal Processing*, vol. 54, no. 1, pp. 336–345, Jan. 2006.

Annexe J

Lower bounds on the mean square error derived from a mixture of linear and non-linear transformations of the unbiasness definition

Proc. of *IEEE* International Conference on Acoustics, Speech, and Signal Processing, ICASSP-09, Taipei, Taiwan

LOWER BOUNDS ON THE MEAN SQUARE ERROR DERIVED FROM MIXTURE OF LINEAR AND NON-LINEAR TRANSFORMATIONS OF THE UNBIASNESS DEFINITION

Eric Chaumette [(1)], *Alexandre Renaux* [(2)] *and Pascal Larzabal*[(3)]

[(1)] ONERA - DEMR/TSI, The French Aerospace Lab, Chemin de la Hunière, F-91120 Palaiseau, France
[(2)] Université Paris-Sud 11, L2S, Supelec, 3 rue Joliot Curie, F-91190 Gif-Sur-Yvette, France
[(3)] SATIE, ENS Cachan, CNRS, UniverSud, 61 av President Wilson, F-94230 Cachan, France

ABSTRACT

It is well known that in non-linear estimation problems the ML estimator exhibits a threshold effect, i.e. a rapid deterioration of estimation accuracy below a certain SNR or number of snapshots. This effect is caused by outliers and is not captured by standard tools such as the Cramér-Rao bound (CRB). The search of the SNR threshold value can be achieved with the help of approximations of the Barankin bound (BB) proposed by many authors. These approximations result from a linear transformation (discrete or integral) of the uniform unbiasness constraint introduced by Barankin. Nevertheless, non-linear transformations can be used as well for some class of p.d.f. including the Gaussian case. The benefit is their combination with existing linear transformation to get tighter lower bounds improving the SNR threshold prediction.

Index Terms— Parameter estimation, mean-square-error bounds, SNR threshold

1. INTRODUCTION

Minimal performance bounds allow for calculation of the best performance that may be achieved, in the Mean Square Error (MSE) sense, when estimating a set of model parameters from noisy observations. Historically the first MSE lower bound for deterministic parameters to be derived was the Cramér-Rao Bound (CRB) [8], which has been the most widely used since. Its popularity is largely due to its simplicity of calculation leading to closed form expressions useful for system analysis and design. Additionally, the CRB can be achieved asymptotically (high SNR [6] and/or large number of snapshots [8]) by Maximum Likelihood Estimators (MLE), and last but not least, it is the lowest bound on the MSE of unbiased estimators, since it derives from a local formulation of unbiasedness in the vicinity of the true parameters [3]. This initial characterization of locally unbiased estimators has been improved first by Bhattacharyya's works [8] which refined the characterization of local unbiasedness, and significantly generalized by Barankin works [1], who established the general form of the greatest lower bound on MSE (BB) taking into account a uniform unbiasedness definition (eq. (1)). Unfortunately the BB is the solution of an integral equation with a generally incomputable analytic solution (eq. (8)). Therefore, since then, numerous works detailed in [3][7] have been devoted to deriving computable approximations of the BB and have shown that the CRB and the BB can be regarded as key representatives of two general classes of bounds, respectively the Small-Error bounds and the Large-Error bounds. These works have also shown that in non-linear estimation problems three distinct regions of operation can be observed. In the asymptotic region, *i.e.* at a high number of independent snapshots and/or at high SNR, the MSE is small and, in many cases, close to the Small-Error bounds. In the *a priori* performance region where the number of independent snapshots and/or the SNR are very low, the observations provide little information and the MSE is close to that obtained from the prior knowledge about the problem. Between these two extremes, there is a transition region where MSE of MLEs usually deteriorates rapidly with respect to existing MSE lower bounds (Large or Small) and exhibits a threshold behaviour, which corresponds to a "performance breakdown" of the estimators due to the appearance of outliers. Small-Error bounds are not able to handle the threshold phenomena, whereas it is revealed by Large-Error bounds that can be used to predict the threshold value. On the other hand, large-Error bounds suffer from their computational cost. Indeed, each BB approximation request the search of an optimum over a set of test points and their tightness depends on the chosen set of test points.

And tightness is the matter, since a more accurate knowledge of the BB allows a better prediction of the SNR threshold value.

Therefore, at least two strategies can be adopted. The first one consists in the search for the tightest and computable approximation of the BB, for a given set of test points [3]. The second one consists in exploiting a particular property of a given class of p.d.f. in order to design a lower bound able to reveal the set of test points optimizing its tightness [7]. All these approximations of the BB result from a linear transformation (discrete or integral) of the uniform unbiasness constraint introduced by Barankin (eq. (1)). Indeed, they are different solutions of the same norm minimization problem under sets of appropriate linear constraints (eq. (4)) and derive from a simple formula (eq. (5)).

Another possible strategy to derive new BB approximations, possibly tighter, is to resort to non-linear transformations of the uniform unbiasness constraint (eq. (1)). It seems that, since the original idea came from non regular-estimation (class of p.d.f. for which the CRB is the trivial bound 0) [2], its application to regular-estimation (class of p.d.f. for which the CRB is not the trivial bound 0) has been completely overlooked, and has sunk into oblivion.

Therefore the aim of this paper is to bring non-linear transformations out of oblivion. We show that the rationale introduced in [2]

This work has been partly funded by the European Network of excellence NEWCOM++ under the number 216715

This project is partialy funded by both the Région Ile-de-France and the Digiteo Research Park

is applicable to Gaussian p.d.f., which is one of the most important regular p.d.f. in signal processing, with application the to single tone threshold analysis. Moreover, the generalization of this rationale allows to introduce a more general class of possible transformations (eq. (12)) of the uniform unbiasness constraint (eq. (1)), i.e. the mixture of integral linear and non-linear transformations, opening new directions in the search of computable tighter BB approximations improving the SNR threshold prediction.
Last, but not least, the next Section provides an unified simple framework for the derivation of any Barankin bound approximations based on linear transformations, including the last ones introduced in [7].

2. BARANKIN BOUND APPROXIMATIONS BY LINEAR TRANSFORMATIONS

For the sake of simplicity we will focus on the estimation of a single real function $g(\theta)$ of a single unknown real deterministic parameter θ. In the following, unless otherwise stated, $\mathbf{x}$ denotes the random observation vector of dimension M, Ω the observations space, and $p(\mathbf{x};\theta)$ the probability density function (p.d.f.) of $\mathbf{x}$ depending on $\theta \in \Theta$, where Θ denotes the parameter space. Let $L^2(\Omega)$ be the real Hilbert space of square integrable functions over Ω.
In the search for a lower bound on the MSE of unbiased estimators, two fundamental properties of the problem at hand, introduced by Barankin [1], must be noticed. The first property is that the MSE of a particular estimator $\widehat{g(\theta^0)}(\mathbf{x}) \in L^2(\Omega)$ of $g(\theta^0)$, where θ^0 is a selected value of the parameter θ, is a norm associated with a particular scalar product $\langle \mid \rangle_\theta$:

$$\begin{aligned} MSE_{\theta^0}\left[\widehat{g(\theta^0)}\right] &= \left\| \widehat{g(\theta^0)}(\mathbf{x}) - g(\theta^0) \right\|_{\theta^0}^2, \\ \langle g(\mathbf{x}) \mid h(\mathbf{x}) \rangle_{\theta^0} &= E_{\theta^0}\left[g(\mathbf{x})^* h(\mathbf{x}) \right]. \end{aligned}$$

The second property is that an unbiased estimator $\widehat{g(\theta^0)}(\mathbf{x})$ of $g(\theta)$ should be uniformly unbiased, i.e. for all possible values of the unknown parameter $\theta \in \Theta$ it must verify:

$$E_\theta\left[\widehat{g(\theta^0)}(\mathbf{x})\right] = g(\theta) = E_{\theta^0}\left[\widehat{g(\theta^0)}(\mathbf{x})\,\nu(\mathbf{x};\theta)\right], \quad (1)$$

where $\nu(\mathbf{x};\theta) = \frac{p(\mathbf{x};\theta)}{p(\mathbf{x};\theta^0)}$ denotes the Likelihood Ratio (LR). As a consequence, the locally-best (at θ^0) unbiased estimator is the solution of a norm minimization under linear constraints

$$\min\left\{ MSE_{\theta^0}\left[\widehat{g(\theta^0)}\right] \right\} \text{ under } E_{\theta^0}\left[\widehat{g(\theta^0)}(\mathbf{x})\,\nu(\mathbf{x};\theta)\right] = g(\theta),$$

solution that can be obtained by using the norm minimization lemma

$$\begin{aligned} &\min\left\{ \mathbf{u}^H\mathbf{u} \text{ under } \mathbf{c}_k^H\mathbf{u} = v_k, 1 \le k \le K \right\} = \mathbf{v}^H\mathbf{G}^{-1}\mathbf{v} \\ &\mathbf{u}_{opt} = \sum_{k=1}^{K} \alpha_k \mathbf{c}_k, \quad \boldsymbol{\alpha} = \mathbf{G}^{-1}\mathbf{v}, \quad \mathbf{G}_{n,k} = \mathbf{c}_n^H\mathbf{c}_k \end{aligned}. \quad (2)$$

Unfortunately, as shown hereinafter, if Θ contains a continuous subset of $\mathbb{R}$, then the norm minimization under a set of an infinite number of linear constraints (1) leads to an integral equation (8) with no analytical solution in general. Therefore, since the original work of Barankin [1], many studies [3, and references therein][7] have been dedicated to the derivation of "computable" lower bounds approximating the MSE of the locally-best unbiased estimator (BB). All these approximations derive from sets of discrete or integral linear transform of the "Barankin" constraint (1), and accordingly of the LR, and can be obtained using the following simple rationale.

Let $\boldsymbol{\theta}^N = \left(\theta^1, \ldots, \theta^N\right)^T \in \Theta^N$ be a vector of N test points, $\boldsymbol{\nu}\left(\mathbf{x};\boldsymbol{\theta}^N\right) = \left(\nu\left(\mathbf{x};\theta^1\right), \ldots, \nu\left(\mathbf{x};\theta^N\right)\right)^T$ be the vector of LR associated to $\boldsymbol{\theta}^N$, $\xi(\theta) = g(\theta) - g\left(\theta^0\right)$ and $\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right) = \left(\xi\left(\theta^1\right), \ldots, \xi\left(\theta^N\right)\right)^T$. Any unbiased estimator $\widehat{g\left(\theta^0\right)}(\mathbf{x})$ verifying (1) must comply with

$$E_{\theta^0}\left[\left(\widehat{g\left(\theta^0\right)}(\mathbf{x}) - g\left(\theta^0\right)\right)\boldsymbol{\nu}\left(\mathbf{x};\boldsymbol{\theta}^N\right)\right] = \boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right), \quad (3)$$

and with any subsequent linear transformation of (3). Therefore, any given set of K ($K \le N$) independent linear transformations of (3):

$$E_{\theta^0}\left[\left(\widehat{g\left(\theta^0\right)}(\mathbf{x}) - g\left(\theta^0\right)\right)\mathbf{h}_k^T\boldsymbol{\nu}\left(\mathbf{x};\boldsymbol{\theta}^N\right)\right] = \mathbf{h}_k^T\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right), \quad (4)$$

$\mathbf{h}_k \in \mathbb{R}^N, k \in [1,K]$, provides with a lower bound on the MSE (2):

$$MSE_{\theta^0}\left[\widehat{g\left(\theta^0\right)}\right] \ge \boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T \widetilde{\mathbf{G}}_{\mathbf{H}_K}\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right), \quad (5)$$

where $\widetilde{\mathbf{G}}_{\mathbf{H}_K} = \mathbf{H}_K\left(\mathbf{H}_K^T\mathbf{R}_{\boldsymbol{\nu}}\mathbf{H}_K\right)^{-1}\mathbf{H}_K^T$, $\mathbf{H}_K = [\mathbf{h}_1 \ldots \mathbf{h}_K]$ and $\left(\mathbf{R}_{\boldsymbol{\nu}}\right)_{n,m} = E_{\theta^0}\left[\nu(\mathbf{x};\theta^n)\,\nu(\mathbf{x};\theta^m)\right]$. The BB is obtained by taking the supremum of (5) over all the existing degrees of freedom $\left(N, \boldsymbol{\theta}^N, K, \mathbf{H}_K\right)$. Moreover, for a given vector of test points $\boldsymbol{\theta}^N$, the lower bound (5) reaches its maximum iff the matrix $\mathbf{H}_K$ is invertible ($K = N$) [5], which represents a bijective transformation of the set of the N initial constraints (3):

$$MSE_{\theta^0}\left[\widehat{g\left(\theta^0\right)}\right] \ge \boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T \widetilde{\mathbf{G}}_{\mathbf{I}_N}\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right) \ge \boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T \widetilde{\mathbf{G}}_{\mathbf{H}_K}\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right),$$

where $\mathbf{I}_N$ is the identity matrix with dimension N. All known bounds on the MSE deriving from the Barankin Bound is a particular implementation of (5), including the most general formalism introduced lately in [7]. Indeed, the limit of (4) where $N \to \infty$ and $\boldsymbol{\theta}^N$ uniformly samples Θ leads to the linear integral constraint:

$$E_{\theta^0}\left[\left(\widehat{g\left(\theta^0\right)}(\mathbf{x}) - g\left(\theta^0\right)\right)\eta(\mathbf{x},\tau)\right] = \Gamma_h(\tau), \quad (6)$$

$$\eta(\mathbf{x},\tau) = \int_\Theta h(\tau,\theta)\,\nu(\mathbf{x};\theta)\,d\theta, \quad \Gamma_h(\tau) = \int_\Theta h(\tau,\theta)\,\xi(\theta)\,d\theta,$$

where each $\mathbf{h}_k = \left(h\left(\tau_k,\theta^1\right), \ldots, h\left(\tau_k,\theta^N\right)\right)^T$ is the vector of samples of a parametric function $h(\tau,\theta), \tau \in \Lambda \subset \mathbb{R}$, integrable over $\Theta, \forall \tau \in \Lambda$. Then, for any subset of K values of τ, $\{\tau_k\}_{k\in[1,K]}$, the subset of the associated K linear integral constraints (6) leads to the following lower bound (2):

$$\lim_{K\to\infty} \left| \begin{array}{c} MSE_{\theta^0}\left[\widehat{g\left(\theta^0\right)}(\mathbf{x})\right] \ge MSE_{\theta^0}\left[\widehat{g\left(\theta^0\right)}_{lmvu}(\mathbf{x})\right] \\ MSE_{\theta^0}\left[\widehat{g\left(\theta^0\right)}_{lmvu}(\mathbf{x})\right] = \boldsymbol{\Gamma}_h^T\mathbf{R}_\eta^{-1}\boldsymbol{\Gamma}_h = \boldsymbol{\Gamma}_h^T\left(\frac{\mathbf{a}}{\lambda}\right) \\ \widehat{g\left(\theta^0\right)}_{lmvu}(\mathbf{x}) - g\left(\theta^0\right) = \sum_{k=1}^{K}\frac{a_k}{\lambda}\eta(\mathbf{x},\tau_k) \\ \mathbf{R}_\eta\left(\frac{\mathbf{a}}{\lambda}\right) = \boldsymbol{\Gamma}_h \end{array} \right. \quad (7)$$

where $\left(\mathbf{R}_\eta\right)_{k,k'} = E_{\theta^0}\left[\eta(\mathbf{x},\tau_k)\,\eta(\mathbf{x},\tau_{k'})\right]$ and $\left(\boldsymbol{\Gamma}_h\right)_k = \Gamma(\tau_k)$. Therefore, when $K \to \infty$ and the set $\{\tau_k\}_{k\in[1,K]}$ uniformly samples Λ, by setting $\frac{1}{\lambda} = d\tau = \tau_{k+1} - \tau_k$, $\boldsymbol{\beta} = \frac{\mathbf{a}}{\lambda}$, the integral form of the above lower bound appears straightforwardly:

$$\left\{ \begin{array}{l} MSE_{\theta^0}\left[\widehat{g\left(\theta^0\right)}_{lmvu}(\mathbf{x})\right] = \int_\Lambda \Gamma_h(\tau)\,\beta(\tau)\,d\tau \\ \widehat{g\left(\theta^0\right)}_{lmvu}(\mathbf{x}) - g\left(\theta^0\right) = \int_\Lambda \eta(\mathbf{x},\tau)\,\beta(\tau)\,d\tau \\ \int_\Lambda K_h(\tau',\tau)\,\beta(\tau)\,d\tau = \Gamma_h(\tau') \end{array} \right., \quad (8)$$

$$K_h(\tau,\tau') = E_{\theta^0}[\eta(\mathbf{x},\tau)\,\eta(\mathbf{x},\tau')] = \iint_{\Theta} h(\tau,\theta)\,R_{\boldsymbol{\nu}}(\theta,\theta')\,h(\tau',\theta')\,d\theta d\theta',$$

$$R_{\boldsymbol{\nu}}(\theta,\theta') = E_{\theta^0}\left[\frac{p(\mathbf{x};\theta)}{p(\mathbf{x};\theta^0)}\frac{p(\mathbf{x};\theta')}{p(\mathbf{x};\theta^0)}\right] = \int_{\Omega}\frac{p(\mathbf{x};\theta)p(\mathbf{x};\theta')}{p(\mathbf{x};\theta^0)}d\mathbf{x},$$

which is exactly the main result introduced in [7] and is a generalization of the Kiefer Bound [4] ($K=2$). Note that if $h(\tau,\theta)=\delta(\tau-\theta)$ (limit case of $\mathbf{H}_N=\mathbf{I}_N$ where $N=K\to\infty$) then $K_h(\tau,\tau')=R_{\boldsymbol{\nu}}(\tau,\tau')$ and (8) becomes the simplest expression of the exact Barankin Bound [3, (10)]. As mentioned above, in most practical cases, it is impossible to find either the limit of (7) or an analytical solution of (8) to obtain an explicit form of the exact Barankin Bound on the MSE, which somewhat limits its interest.
Nevertheless this formalism allows to use discrete (4) or integral (6) linear transforms of the LR, possibly non-invertible, possibly optimized for a set of p.d.f. (such as the Fourier transform in [7]) in order to get a tight approximation of the BB.

3. BARANKIN BOUND APPROXIMATIONS BY NON-LINEAR TRANSFORMATIONS

At the opposite, the use of a non-linear transformation of the unbiasness definition (1) of type

$$E_{\theta^0}\left[\widehat{g(\theta^0)}(\mathbf{x})\,t(\nu(\mathbf{x};\theta))\right] = h(g(\theta)) \tag{9}$$

is more obscure since it seems a difficult mathematical task to compute the bias transformation function $h(\,)$ as a function of the LR transformation function $t(\,)$ and of the LR. Nevertheless there is a class of estimation problems where non-linear transformations of the LR can be used to derive new lower bounds on the MSE. It is the class of estimation problems characterized by a p.d.f. $p(\mathbf{x};\theta)$ for which there exists at least one real valued function $t(\,)$ such that, the transformation of p.d.f. $p(\mathbf{x};\theta)$ by $t(\,)$ is still - up to a normalization constant, w.r.t. $\mathbf{x}$, $k(\theta,t)$ - a p.d.f. of the form $p(\mathbf{x};\,)$ but parameterized by a modified parameter value γ, function of the initial parameter θ and of the transformation $t(\,)$:

$$t(p(\mathbf{x};\theta)) = k(\theta,t)\,p(\mathbf{x};\gamma(\theta,t)), \quad k(\theta,t)=\int_{\Omega} t(p(\mathbf{x};\theta))\,d\mathbf{x} \tag{10}$$

Then an unbiased estimator verifying (1) verifies as well,$\forall\theta\in\Theta$:

$$\int_{\Omega}\widehat{g(\theta^0)}(\mathbf{x})\,t(p(\mathbf{x};\theta))\,d\mathbf{x} = k(\theta,t)\int_{\Omega}\widehat{g(\theta^0)}(\mathbf{x})\,p(\mathbf{x};\gamma(\theta,t))\,d\mathbf{x} = k(\theta,t)\,g(\gamma(\theta,t))$$

what implies,$\forall\theta\in\Theta$

$$E_{\theta^0}\left[\left(\widehat{g(\theta^0)}(\mathbf{x})-g(\theta^0)\right)\frac{t(p(\mathbf{x};\theta))}{p(\mathbf{x};\theta^0)}\right] = k(\theta,t)\left[g(\gamma(\theta,t))-g(\theta^0)\right].$$

In the most general case, if there exists a set of functions $t_\theta(\,)$ verifying (10), then any unbiased estimator also verifies,$\forall\theta\in\Theta$:

$$E_{\theta^0}\left[\left(\widehat{g(\theta^0)}(\mathbf{x})-g(\theta^0)\right)\frac{t_\theta(p(\mathbf{x};\theta))}{p(\mathbf{x};\theta^0)}\right] = k(\theta,t_\theta)\left[g(\gamma(\theta,t_\theta))-g(\theta^0)\right].$$

Therefore, if we update the definition of $\nu(\mathbf{x};\theta)$ and $\xi(\theta)$ in (6) according to

$$\nu(\mathbf{x};\theta) = \frac{t_\theta(p(\mathbf{x};\theta))}{p(\mathbf{x};\theta^0)}, \quad \xi(\theta) = k(\theta,t_\theta)\left[g(\gamma(\theta,t_\theta))-g(\theta^0)\right], \tag{11}$$

all the results released in the previous Section still hold, the linear integral transformation becoming a mixture of linear and non-linear integral transformations:

$$\eta(\mathbf{x},\tau) = \int_{\Theta} h(\tau,\theta)\,\frac{t_\theta(p(\mathbf{x};\theta))}{p(\mathbf{x};\theta^0)}d\theta,$$
$$\Gamma_h(\tau) = \int_{\Theta} h(\tau,\theta)\,k(\theta,t_\theta)\left[g(\gamma(\theta,t_\theta))-g(\theta^0)\right]d\theta. \tag{12}$$

The proposed rationale is a generalization of [2] where the authors has extended the Hammersley-Chapman-Robbins bound (HCRB) and the Bhattacharyya bound (BaB) for a particular non-linear transformation $t_q(y)=y^q, q\in]0,1]$, to overcome a non-regular estimation problem : the estimation of the parameters of a Pearson Type III p.d.f.. Indeed for such a p.d.f., the usual bounds based on linear transformations such as the Cramér-Rao bound (CRB), the HCRB and the BaB yield the trivial bound 0.
At first sight, the proposed rationale does not seem appealing, since a non-linear transformation of type (9) or (10) is unlikely to exist whatever the form of the p.d.f., although the linear transformation of the LR (6) is always possible, however possibly yielding the trivial bound 0. It is probably the reason why the application of the proposed rationale to regular estimation problem has been completely overlooked, even by authors of [2], although it is applicable to Gaussian p.d.f..

3.1. Application to the Gaussian observation model

We focus on M-dimensional complex circular Gaussian p.d.f.:

$$p(\mathbf{x};\boldsymbol{\theta}) = p(\mathbf{x};\mathbf{m}(\boldsymbol{\theta}),\mathbf{C}(\boldsymbol{\theta})) = \frac{e^{-(\mathbf{x}-\mathbf{m}(\boldsymbol{\theta}))^H\mathbf{C}(\boldsymbol{\theta})^{-1}(\mathbf{x}-\mathbf{m}(\boldsymbol{\theta}))}}{\pi^M|\mathbf{C}(\boldsymbol{\theta})|}$$

Then it is worth noticing that the transformation $t_q(y)=y^q$ can be applied to the observation model resulting from a mixture of deterministic and stochastic signals in presence of Gaussian interference. Indeed, in this case $\mathbf{m}(\boldsymbol{\theta})=\mathbf{m}(\boldsymbol{\varepsilon})$, $\mathbf{C}(\boldsymbol{\theta})=\boldsymbol{\Psi}(\boldsymbol{\zeta})\,\mathbf{C_s}\boldsymbol{\Psi}(\boldsymbol{\zeta})^H+\mathbf{C_n}$, $\boldsymbol{\theta}=\left[\boldsymbol{\varepsilon}^T,\boldsymbol{\zeta}^T,vec(\mathbf{C_s})^T,vec(\mathbf{C_n})^T\right]^T$ and:

$$\begin{aligned} t_q(p(\mathbf{x};\boldsymbol{\theta})) &= k(\boldsymbol{\theta},q)\,p(\mathbf{x};\boldsymbol{\gamma}(\boldsymbol{\theta},q)) \\ k(\boldsymbol{\theta},q) &= \frac{\pi^{M(1-q)}}{q^{qM}}\left|\frac{\mathbf{C}(\boldsymbol{\theta})}{q}\right|^{1-q} \\ \boldsymbol{\gamma}(\boldsymbol{\theta},q) &= \left[\boldsymbol{\varepsilon}^T,\boldsymbol{\zeta}^T,\frac{vec(\mathbf{C_s})^T}{q},\frac{vec(\mathbf{C_n})^T}{q}\right]^T \end{aligned} \tag{13}$$

3.2. Single tone threshold analysis

A reference problem in threshold analysis is the estimation of a single tone $\eta\in]-0.5,0.5[$ for a deterministic observation model:

$$\begin{aligned} \mathbf{x} &= a\boldsymbol{\psi}(\eta^0)+\mathbf{n}, \quad \boldsymbol{\psi}(\eta)=\left[1,...,e^{j(M-1)2\pi\eta}\right]^T \\ p(\mathbf{x};\boldsymbol{\theta}) &= \frac{e^{-\|\mathbf{x}-a\boldsymbol{\psi}(\eta)\|^2}}{\pi^M} \end{aligned} \tag{14}$$

In this problem, $\boldsymbol{\theta} = \left[\eta, a, \sigma_{\mathbf{n}}^2 = 1\right]$, $\mathbf{m}(\boldsymbol{\theta}) = a\boldsymbol{\psi}(\eta)$, $\mathbf{C}(\boldsymbol{\theta}) = \mathbf{I}$, a^2 is the known SNR ($a > 0$) and $\mathbf{n}$ is a complex circular Gaussian noise, with zero mean and a known covariance matrix $\mathbf{C_n} = \mathbf{I}$. As a consequence: $k(\boldsymbol{\theta},q) = \frac{\pi^{M(1-q)}}{q^M}$, $\boldsymbol{\gamma}(\boldsymbol{\theta}, q) = \left(\eta, a, \frac{1}{q}\right)$ and $g(\boldsymbol{\gamma}(\boldsymbol{\theta}, q)) = g(\boldsymbol{\theta}) = \eta$.

Let us now consider the generalization of the HCRB, the simplest approximation of the BB (5) based on 2 test-points $\boldsymbol{\eta}^2 = \left(\eta^0 + h, \eta^0 - h\right)^T$ where $\mathbf{H}_2 = \mathbf{I}$, obtained by using the non-linear transformation mentioned above (13). Application of (5) and (11) where $\nu(\mathbf{x};\boldsymbol{\theta}) = \frac{p(\mathbf{x};\boldsymbol{\theta})^q}{p\left(\mathbf{x};\boldsymbol{\theta}^0\right)}$ and $\xi(\eta) = \frac{\pi^{M(1-q)}}{q^M}\left(\eta - \eta^0\right)$ leads to the following lower bound $HCRB_q$:

$$MSE_{\eta^0}\left[\widehat{\eta^0}\right] \geq HCRB_q = \boldsymbol{\xi}\left(\boldsymbol{\eta}^2\right)^T \mathbf{R}^{-1} \boldsymbol{\xi}\left(\boldsymbol{\eta}^2\right)$$

$$\boldsymbol{\xi}\left(\boldsymbol{\eta}^2\right) = \frac{\pi^{M(1-q)}}{q^M}\begin{bmatrix} h \\ -h \end{bmatrix}, \qquad \mathbf{R} = \begin{bmatrix} R_{1,1} & R_{1,2} \\ R_{1,2} & R_{2,2} \end{bmatrix}$$

$$R_{1,1} = E_{\boldsymbol{\theta}^0}\left[\frac{p\left(\mathbf{x};\boldsymbol{\theta}^0 + h\right)^{2q}}{p\left(\mathbf{x};\boldsymbol{\theta}^0\right)^2}\right] = \alpha e^{\frac{2qa^2}{2q-1}\left\|\boldsymbol{\psi}\left(\eta^0+h\right) - \boldsymbol{\psi}\left(\eta^0\right)\right\|^2}$$

$$R_{1,2} = E_{\boldsymbol{\theta}^0}\left[\frac{p\left(\mathbf{x};\boldsymbol{\theta}^0 + h\right)^q p\left(\mathbf{x};\boldsymbol{\theta}^0 - h\right)^q}{p\left(\mathbf{x};\boldsymbol{\theta}^0\right)^2}\right]$$

$$= \alpha \frac{e^{\frac{2qa^2}{2q-1}\left\|\frac{\boldsymbol{\psi}\left(\eta^0+h\right)+\boldsymbol{\psi}\left(\eta^0-h\right)}{2} - \boldsymbol{\psi}\left(\eta^0\right)\right\|^2}}{e^{2qa^2\left\|\frac{\boldsymbol{\psi}\left(\eta^0+h\right)-\boldsymbol{\psi}\left(\eta^0-h\right)}{2}\right\|^2}}$$

$$R_{2,2} = E_{\boldsymbol{\theta}^0}\left[\frac{p\left(\mathbf{x};\boldsymbol{\theta}^0 - h\right)^{2q}}{p\left(\mathbf{x};\boldsymbol{\theta}^0\right)^2}\right] = \alpha e^{\frac{2qa^2}{2q-1}\left\|\boldsymbol{\psi}\left(\eta^0-h\right) - \boldsymbol{\psi}\left(\eta^0\right)\right\|^2}$$

where $\alpha = \frac{\pi 2^{M(1-q)}}{(2q-1)^M}$. The classical $HCRB$, i.e. the $HCRB$ obtained from a linear transformation of the LR, is the particular case of its generalized form $HCRB_q$, obtained from a non-linear transformation of the LR, where $q = 1$.

The benefits on tightness of the introduction of a non-linear transformation is highlighted on figure (1) where the bounds $HCRB = \sup_h \{HCRB_{q=1}\}$ and $HCRB_{NL} = \sup_{h,q} \{HCRB_q\}$ $\left(0 \leq h \leq 0.5, \frac{1}{2} < q < 2\right)$ are displayed and compared with both the CRB and the MSE of the MLE estimator for $M = 10$ and $\eta^0 = 0$.

3.3. Results and Perspectives

First, the proof of the gain in tightness of the $HCRB$ incorporating the non-linear transformation $t_q(y) = y^q$ allows to state that:
in the case of the Gaussian p.d.f. family described in Section 3.1 (including both deterministic and stochastic observation models), there exists at least one non-linear transformation improving the tightness of any lower bound deriving from a linear transformation of the unbiasness constraint (1).
Indeed, all these lower bounds derives from expression (5), which is a generalization of the $HCRB$ to N test-points.

Second, an immediate improvement of tightness of all existing lower bounds can be obtained by using the integral form $t_\theta(y) = y^{q(\theta)}$ of the discrete non-linear transformation $t_\theta(y) = y^q$, for various function $q(\theta)$.

Last, there are probably other non-linear transformations applicable to the Gaussian (or other) p.d.f which should increase the tightness of existing bounds ....

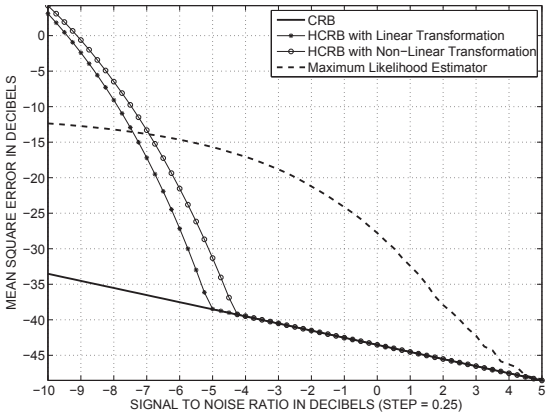


Fig. 1. Comparison of MSE lower bounds versus SNR

4. CONCLUSION

In this paper, we have shown that it is worth looking for non-linear transformations of the unbiasness definition to be combined with linear transformations, in order to derive tighter Barankin bound approximations to improve the threshold prediction. Indeed such non-linear transformations exist for the Gaussian p.d.f.. Additionally, we provide an unified simple framework for the derivation of any Barankin bound approximations based on mixture of linear and non-linear transformations.

5. REFERENCES

[1] E.W. Barankin, "Locally best unbiased estimates", Ann. Math. Stat., vol. 20, no. 4, pp. 477-501, 1949.

[2] W. R. Blischke, A. J. Truelove and P. B. Mundle, "On non-regular estimation . I.Variance bounds for estimators of location parameters", Journal of American Stat. Assoc., vol. 64, no. 327, pp. 1056-1072, Sep. 1969.

[3] E. Chaumette, J. Galy, A. Quinlan, P. Larzabal, "A New Barankin Bound Approximation for the Prediction of the Threshold Region Performance of Maximum-Likelihood Estimators", IEEE Trans. on SP, vol. 56, no. 11, pp. 5319-5333, Nov. 2008

[4] J. Kiefer, "On minimum variance estimators", Ann. Math. Stat., vol. 23, no. 4, pp. 627–629, 1952.

[5] R. McAulay and L.P. Seidman, "A useful form of the Barankin lower bound and its application to PPM threshold analysis", IEEE Trans. on IT, vol. 15, no. 2, pp. 273-279, Mar. 1969.

[6] A. Renaux, P. Forster, E. Chaumette, and P. Larzabal, "On the high SNR CML estimator full statistical characterization," IEEE Trans. on SP, vol. 54, no. 12, pp. 4840-4843, Dec. 2006.

[7] K. Todros and J. Tabrikian, "A new lower bound on the mean-square error of unbiased estimators", in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., pp. 3913-3916, 2008

[8] H.L. Van Trees, *Detection, Estimation and Modulation Theory, Part 1*. New York: Wiley, 1968.

Annexe K

New trends in deterministic lower bounds and SNR threshold estimation : from derivable bounds to conjectural bounds

Proc. of *IEEE* Sensor Array Multichannel Workshop SAM-2010 (special session on lower bound in array processing), Kibutz Ma'ale Hahamisha, Israel

2010 IEEE Sensor Array and Multichannel Signal Processing Workshop

NEW TRENDS IN DETERMINISTIC LOWER BOUNDS AND SNR THRESHOLD ESTIMATION: FROM DERIVABLE BOUNDS TO CONJECTURAL BOUNDS

Eric Chaumette [1], *Alexandre Renaux* [2] *and Pascal Larzabal*[3]

[1] ONERA - DEMR/TSI, The French Aerospace Lab, Chemin de la Hunière, F-91120 Palaiseau, France
[2] Université Paris-Sud 11, L2S, Supelec, 3 rue Joliot Curie, F-91190 Gif-Sur-Yvette, France
[3] SATIE, ENS Cachan, CNRS, UniverSud, 61 av President Wilson, F-94230 Cachan, France

ABSTRACT

It is well known that in non-linear estimation problems the ML estimator exhibits a threshold effect, i.e. a rapid deterioration of estimation accuracy below a certain SNR or number of snapshots. This effect is caused by outliers and is not captured by standard tools such as the Cramér-Rao bound (CRB). The search of the SNR threshold value can be achieved with the help of approximations of the Barankin bound (BB) proposed by many authors. These approximations may result from linear or non-linear transformation (discrete or integral) of the uniform unbiasedness constraint introduced by Barankin. Additionally, the strong analogy between derivations of deterministic bounds and Bayesian bounds of the Weiss-Weinstein family has led us to propose a conjectural bound which outperforms existing ones for SNR threshold prediction.

Index Terms— Parameter estimation, mean-square-error bounds, SNR threshold

1. INTRODUCTION

Minimal performance bounds allow for calculation of the best performance that may be achieved, in the Mean Square Error (MSE) sense, when estimating a set of model parameters from noisy observations. Historically the first MSE lower bound for deterministic parameters to be derived was the Cramér-Rao Bound (CRB) [4], which has been the most widely used since. Its popularity is largely due to its simplicity of calculation leading to closed-form expressions useful for system analysis and design. Additionally, the CRB can be achieved asymptotically (high SNR and/or large number of snapshots) by Maximum Likelihood Estimators (MLE), and last but not least, it is the lowest bound on the MSE of unbiased estimators, since it derives from a local formulation of unbiasedness in the vicinity of the true parameters [2]. This initial characterization of locally unbiased estimators has been improved first by Bhattacharyya's works [4] which refined the characterization of local unbiasedness, and significantly generalized by Barankin works [1], who established the general form of the greatest lower bound on MSE (BB) taking into account a uniform unbiasedness definition (eq. (1)). Unfortunately the BB is the solution of an integral equation with a generally incomputable analytic solution (eq. (8)).

Therefore, since then, numerous works detailed in [2][3] have been devoted to deriving computable approximations of the BB and have shown that the CRB and the BB can be regarded as key representatives of two general classes of bounds, respectively the Small-Error bounds and the Large-Error bounds. These works have also shown that in non-linear estimation problems three distinct regions of operation can be observed. In the asymptotic region, *i.e.* at a high number of independent snapshots and/or at high SNR, the MSE is small and, in many cases, close to the Small-Error bounds. In the *a priori* performance region where the number of independent snapshots and/or the SNR are very low, the observations provide little information and the MSE is close to that obtained from the prior knowledge about the problem. Between these two extremes, there is a transition region where MSE of MLEs usually deteriorates rapidly with respect to existing MSE lower bounds (Large or Small) and exhibits a threshold behaviour, which corresponds to a "performance breakdown" of the estimators due to the appearance of outliers.

Small-Error bounds are not able to handle the threshold phenomena, whereas it is revealed by Large-Error bounds that can be used to predict the threshold value. On the other hand, Large-Error bounds suffer from their computational cost. Indeed, each BB approximation request the search of an optimum over a set of test points and their tightness depends on the chosen set of test points.

And tightness is the matter, since a more accurate knowledge of the BB allows a better prediction of the SNR threshold value.

Therefore, at least two strategies can be adopted.

The first one is the most consistent with deductive reasoning applied to the unbiasedness paradigm. This strategy, fully mastered from the mathematics and the meaning point of view, provides derivable bounds and relies on the introduction a general class of possible transformations (eq. (11)) of the uniform unbiasedness constraint (eq. (1)), i.e. the mixture of integral linear and non-linear transformations, opening a wide variety of directions in the search of computable tighter BB approximations.

The second one is partially consistent with deductive reasoning since it may be based - see our example Section 3 - on analogies between families of lower bounds without the support of a non-questionnable derivation and interpretation. And yet, but nevertheless this strategy may yield some conjectural bounds tightest than the existing and well-estabished ones.

This work has been partly funded by the European Network of excellence NEWCOM++ under the number 216715

This project is partialy funded by both the Région Ile-de-France and the Digiteo Research Park

2. DERIVABLE LOWER BOUNDS

2.1. Linear transformations of the unbiasedness constraint

For the sake of simplicity we will focus on the estimation of a single real function $g(\theta)$ of a single unknown real deterministic parameter θ. In the following, unless otherwise stated, $\mathbf{x}$ denotes the random observation vector of dimension M, Ω the observations space, and $p(\mathbf{x};\theta)$ the probability density function (p.d.f.) of $\mathbf{x}$ depending on $\theta \in \Theta$, where Θ denotes the parameter space. Let $L^2(\Omega)$ be the real Hilbert space of square integrable functions over Ω.

In the search for a lower bound on the MSE of unbiased estimators, two fundamental properties of the problem at hand, introduced by Barankin [1], must be noticed. The first property is that the MSE of a particular estimator $\widehat{g(\theta^0)}(\mathbf{x}) \in L^2(\Omega)$ of $g(\theta^0)$, where θ^0 is a selected value of the parameter θ, is a norm associated with a particular scalar product $\langle \, | \, \rangle_{\theta^0}$:

$$\begin{aligned} MSE_{\theta^0}\left[\widehat{g(\theta^0)}\right] &= \left\| \widehat{g(\theta^0)}(\mathbf{x}) - g(\theta^0) \right\|^2_{\theta^0}, \\ \langle g(\mathbf{x}) \mid h(\mathbf{x}) \rangle_{\theta^0} &= E_{\theta^0}\left[g(\mathbf{x})^* h(\mathbf{x}) \right]. \end{aligned}$$

The second property is that an unbiased estimator $\widehat{g(\theta^0)}(\mathbf{x})$ of $g(\theta)$ should be uniformly unbiased, i.e. for all possible values of the unknown parameter $\theta \in \Theta$ it must satisfy:

$$E_\theta\left[\widehat{g(\theta^0)}(\mathbf{x})\right] = g(\theta) = E_{\theta^0}\left[\widehat{g(\theta^0)}(\mathbf{x})\,\nu(\mathbf{x};\theta)\right], \quad (1)$$

where $\nu(\mathbf{x};\theta) = \frac{p(\mathbf{x};\theta)}{p(\mathbf{x};\theta^0)}$ denotes the Likelihood Ratio (LR). As a consequence, the locally-best (at θ^0) unbiased estimator is the solution of a norm minimization under linear constraints

$$\min\left\{ MSE_{\theta^0}\left[\widehat{g(\theta^0)}\right] \right\} \text{ under } E_{\theta^0}\left[\widehat{g(\theta^0)}(\mathbf{x})\,\nu(\mathbf{x};\theta)\right] = g(\theta),$$

solution that can be obtained by using the norm minimization lemma

$$\begin{aligned} &\min\left\{ \mathbf{u}^H\mathbf{u} \text{ under } \mathbf{c}_k^H\mathbf{u} = v_k, 1 \le k \le K \right\} = \mathbf{v}^H\mathbf{G}^{-1}\mathbf{v} \\ &\mathbf{u}_{opt} = \sum_{k=1}^{K} \alpha_k \mathbf{c}_k, \quad \boldsymbol{\alpha} = \mathbf{G}^{-1}\mathbf{v}, \quad \mathbf{G}_{n,k} = \mathbf{c}_n^H\mathbf{c}_k \end{aligned} \quad (2)$$

Unfortunately, as shown hereinafter, if Θ contains a continuous subset of $\mathbb{R}$, then the norm minimization under a set of an infinite number of linear constraints (1) leads to an integral equation (8) with no analytical solution in general. Therefore, since the original work of Barankin [1], many studies [2, and references therein][3] have been dedicated to the derivation of "computable" lower bounds approximating the MSE of the locally-best unbiased estimator (BB). All these approximations derive from sets of discrete or integral linear transform of the "Barankin" constraint (1), and accordingly of the LR, and can be obtained using the following simple rationale.

Let $\boldsymbol{\theta}^N = \left(\theta^1, \ldots, \theta^N\right)^T \in \Theta^N$ be a vector of N test points, $\boldsymbol{\nu}(\mathbf{x};\boldsymbol{\theta}^N) = \left(\nu(\mathbf{x};\theta^1), \ldots, \nu(\mathbf{x};\theta^N)\right)^T$ be the vector of LR associated to $\boldsymbol{\theta}^N$, $\xi(\theta) = g(\theta) - g(\theta^0)$ and $\boldsymbol{\xi}(\boldsymbol{\theta}^N) = \left(\xi(\theta^1), \ldots, \xi(\theta^N)\right)^T$. Any unbiased estimator $\widehat{g(\theta^0)}(\mathbf{x})$ satisfying (1) must comply with

$$E_{\theta^0}\left[\left(\widehat{g(\theta^0)}(\mathbf{x}) - g(\theta^0)\right)\boldsymbol{\nu}\left(\mathbf{x};\boldsymbol{\theta}^N\right)\right] = \boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right), \quad (3)$$

and with any subsequent linear transformation of (3). Therefore, any given set of K ($K \le N$) independent linear transformations of (3):

$$E_{\theta^0}\left[\left(\widehat{g(\theta^0)}(\mathbf{x}) - g(\theta^0)\right)\mathbf{h}_k^T\boldsymbol{\nu}\left(\mathbf{x};\boldsymbol{\theta}^N\right)\right] = \mathbf{h}_k^T\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right), \quad (4)$$

$\mathbf{h}_k \in \mathbb{R}^N, 1 \le k \le K$, provides with a lower bound on the MSE (2):

$$MSE_{\theta^0}\left[\widehat{g(\theta^0)}\right] \ge \boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T \widetilde{\mathbf{G}}_{\mathbf{H}_K}\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right), \quad (5)$$

where $\widetilde{\mathbf{G}}_{\mathbf{H}_K} = \mathbf{H}_K\left(\mathbf{H}_K^T\mathbf{R}_{\boldsymbol{\nu}}\mathbf{H}_K\right)^{-1}\mathbf{H}_K^T$, $\mathbf{H}_K = [\mathbf{h}_1 \ldots \mathbf{h}_K]$ and $(\mathbf{R}_{\boldsymbol{\nu}})_{n,m} = E_{\theta^0}\left[\nu(\mathbf{x};\theta^n)\,\nu(\mathbf{x};\theta^m)\right]$. The BB is obtained by taking the supremum of (5) over all the existing degrees of freedom $\left(N, \boldsymbol{\theta}^N, K, \mathbf{H}_K\right)$. Moreover, for a given vector of test points $\boldsymbol{\theta}^N$, the lower bound (5) reaches its maximum iff the matrix $\mathbf{H}_K$ is invertible ($K = N$), which represents a bijective transformation of the set of the N initial constraints (3):

$$MSE_{\theta^0}\left[\widehat{g(\theta^0)}\right] \ge \boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T \widetilde{\mathbf{G}}_{\mathbf{I}_N}\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right) \ge \boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right)^T \widetilde{\mathbf{G}}_{\mathbf{H}_K}\boldsymbol{\xi}\left(\boldsymbol{\theta}^N\right),$$

where $\mathbf{I}_N$ is the identity matrix with dimension N. All known bounds on the MSE deriving from the Barankin Bound is a particular implementation of (5), including the most general formalism introduced lately in [3]. Indeed, the limit of (4) where $N \to \infty$ and $\boldsymbol{\theta}^N$ uniformly samples Θ leads to the linear integral constraint:

$$\begin{aligned} E_{\theta^0}\left[\left(\widehat{g(\theta^0)}(\mathbf{x}) - g(\theta^0)\right)\eta(\mathbf{x},\tau)\right] &= \Gamma_h(\tau), \quad (6) \\ \eta(\mathbf{x},\tau) = \int_\Theta h(\tau,\theta)\,\nu(\mathbf{x};\theta)\,d\theta, \quad \Gamma_h(\tau) &= \int_\Theta h(\tau,\theta)\,\xi(\theta)\,d\theta, \end{aligned}$$

where each $\mathbf{h}_k = \left(h(\tau_k,\theta^1), \ldots, h(\tau_k,\theta^N)\right)^T$ is the vector of samples of a parametric function $h(\tau,\theta), \tau \in \Lambda \subset \mathbb{R}$, integrable over $\Theta, \forall \tau \in \Lambda$. Then, for any subset of K values of τ, $\{\tau_k\}_{1\le k\le K}$, the subset of the associated K linear integral constraints (6) leads to the following lower bound (2):

$$\begin{cases} MSE_{\theta^0}\left[\widehat{g(\theta^0)}(\mathbf{x})\right] \ge MSE_{\theta^0}\left[\widehat{g(\theta^0)}_{lmvu}(\mathbf{x})\right] \\ MSE_{\theta^0}\left[\widehat{g(\theta^0)}_{lmvu}(\mathbf{x})\right] = \boldsymbol{\Gamma}_h^T\mathbf{R}_\eta^{-1}\boldsymbol{\Gamma}_h = \boldsymbol{\Gamma}_h^T\left(\frac{\mathbf{a}}{\lambda}\right) \\ \widehat{g(\theta^0)}_{lmvu}(\mathbf{x}) - g(\theta^0) = \sum_{k=1}^{K}\frac{a_k}{\lambda}\eta(\mathbf{x},\tau_k) \\ \mathbf{R}_\eta\left(\frac{\mathbf{a}}{\lambda}\right) = \boldsymbol{\Gamma}_h \end{cases} \quad (7)$$

where $(\mathbf{R}_\eta)_{k,k'} = E_{\theta^0}\left[\eta(\mathbf{x},\tau_k)\,\eta(\mathbf{x},\tau_{k'})\right]$ and $(\boldsymbol{\Gamma}_h)_k = \Gamma(\tau_k)$. Therefore, when $K \to \infty$ and the set $\{\tau_k\}_{1\le k\le K}$ uniformly samples Λ, by setting $\frac{1}{\lambda} = d\tau = \tau_{k+1} - \tau_k$, $\boldsymbol{\beta} = \frac{\mathbf{a}}{\lambda}$, the integral form of the above lower bound appears straightforwardly:

$$\begin{cases} MSE_{\theta^0}\left[\widehat{g(\theta^0)}_{lmvu}(\mathbf{x})\right] = \int_\Lambda \Gamma_h(\tau)\,\beta(\tau)\,d\tau \\ \widehat{g(\theta^0)}_{lmvu}(\mathbf{x}) - g(\theta^0) = \int_\Lambda \eta(\mathbf{x},\tau)\,\beta(\tau)\,d\tau \\ \int_\Lambda K_h(\tau',\tau)\,\beta(\tau)\,d\tau = \Gamma_h(\tau') \end{cases}, \quad (8)$$

$$\begin{aligned} K_h(\tau,\tau') &= E_{\theta^0}\left[\eta(\mathbf{x},\tau)\,\eta(\mathbf{x},\tau')\right] \\ &= \iint_\Theta h(\tau,\theta)\,R_{\boldsymbol{\nu}}(\theta,\theta')\,h(\tau',\theta')\,d\theta d\theta', \end{aligned}$$

$$R_{\boldsymbol{\nu}}(\theta,\theta') = E_{\theta^0}\left[\frac{p(\mathbf{x};\theta)}{p(\mathbf{x};\theta^0)}\frac{p(\mathbf{x};\theta')}{p(\mathbf{x};\theta^0)}\right] = \int_\Omega \frac{p(\mathbf{x};\theta)p(\mathbf{x};\theta')}{p(\mathbf{x};\theta^0)}d\mathbf{x},$$

which is exactly the main result introduced in [3] and is a generalization of the Kiefer Bound [4] ($K = 2$). Note that if $h(\tau,\theta) = \delta(\tau - \theta)$ (limit case of $\mathbf{H}_N = \mathbf{I}_N$ where $N = K \to \infty$) then $K_h(\tau,\tau') = R_{\boldsymbol{\nu}}(\tau,\tau')$ and (8) becomes the simplest expression of the exact Barankin Bound [2, (10)]. As mentioned above, in most practical cases, it is impossible to find either the limit of (7) or an analytical solution of (8) to obtain an explicit form of the exact Barankin Bound on the MSE, which somewhat limits its interest.

Nevertheless this formalism allows to use discrete (4) or integral (6) linear transforms of the LR, possibly non-invertible, possibly optimized for a set of p.d.f. (such as the Fourier transform in [3]) in order to get a tight approximation of the BB.

2.2. Non-linear transformations of the unbiasedness constraint

Let us consider the set of estimation problems characterized by a p.d.f. for which there exists a real valued function t such that:

$$t\left(p\left(\mathbf{x};\theta\right)\right)=k\left(\theta,t\right)p\left(\mathbf{x};\gamma\left(\theta,t\right)\right),\quad k\left(\theta,t\right)=\int_{\Omega}t\left(p\left(\mathbf{x};\theta\right)\right)d\mathbf{x} \tag{9}$$

Then an unbiased estimator satisfying (1) satisfies as well [5],$\forall\theta\in\Theta$:

$$E_{\theta^0}\left[\left(\widehat{g\left(\theta^0\right)}\left(\mathbf{x}\right)-g\left(\theta^0\right)\right)\frac{t\left(p\left(\mathbf{x};\theta\right)\right)}{p\left(\mathbf{x};\theta^0\right)}\right]= k\left(\theta,t\right)\left[g\left(\gamma\left(\theta,t\right)\right)-g\left(\theta^0\right)\right].$$

Moreover, if there exists a set of functions t_θ satisfying (9), then we can update the definition of $\nu\left(\mathbf{x};\theta\right)$ and $\xi\left(\theta\right)$ in (6) according to:

$$\nu\left(\mathbf{x};\theta\right)=\frac{t_\theta\left(p\left(\mathbf{x};\theta\right)\right)}{p\left(\mathbf{x};\theta^0\right)},\quad \xi\left(\theta\right)=k\left(\theta,t_\theta\right)\left[g\left(\gamma\left(\theta,t_\theta\right)\right)-g\left(\theta^0\right)\right], \tag{10}$$

and all the results released in the previous Section still hold, the linear integral transformation becoming a mixture of linear and non-linear integral transformations:

$$\begin{aligned}\eta\left(\mathbf{x},\tau\right)&=\int_{\Theta}h\left(\tau,\theta\right)\frac{t_\theta\left(p\left(\mathbf{x};\theta\right)\right)}{p\left(\mathbf{x};\theta^0\right)}d\theta,\\ \Gamma_h\left(\tau\right)&=\int_{\Theta}h\left(\tau,\theta\right)k\left(\theta,t_\theta\right)\left[g\left(\gamma\left(\theta,t_\theta\right)\right)-g\left(\theta^0\right)\right]d\theta.\end{aligned} \tag{11}$$

At first sight, the proposed rationale does not seem appealing, since a non-linear transformation of type (9) is unlikely to exist whatever the form of the p.d.f., although the linear transformation of the LR (6) is always possible. Fortunately, it is applicable to a subset of M-dimensional complex circular Gaussian p.d.f.:

$$p\left(\mathbf{x};\boldsymbol{\theta}\right)=p\left(\mathbf{x};\mathbf{m}\left(\boldsymbol{\theta}\right),\mathbf{C}\left(\boldsymbol{\theta}\right)\right)=\frac{e^{-\left(\mathbf{x}-\mathbf{m}\left(\boldsymbol{\theta}\right)\right)^H\mathbf{C}\left(\boldsymbol{\theta}\right)^{-1}\left(\mathbf{x}-\mathbf{m}\left(\boldsymbol{\theta}\right)\right)}}{\pi^M\left|\mathbf{C}\left(\boldsymbol{\theta}\right)\right|}$$

Indeed, the transformation $t_q\left(y\right)=y^q$ can be applied to the observation model resulting from a mixture of deterministic and stochastic signals in presence of Gaussian interference [5]. In this case $\mathbf{m}\left(\boldsymbol{\theta}\right)=\mathbf{m}\left(\boldsymbol{\varepsilon}\right)$, $\mathbf{C}\left(\boldsymbol{\theta}\right)=\boldsymbol{\Psi}\left(\boldsymbol{\zeta}\right)\mathbf{C_s}\boldsymbol{\Psi}\left(\boldsymbol{\zeta}\right)^H+\mathbf{C_n}$, $\boldsymbol{\theta}=\left[\boldsymbol{\varepsilon}^T,\boldsymbol{\zeta}^T,vec\left(\mathbf{C_s}\right)^T,vec\left(\mathbf{C_n}\right)^T\right]^T$.

3. CONJECTURAL LOWER BOUNDS

Although initially introduced by resorting to the covariance inequality, the Bayesian bounds of the Weiss-Weinstein family have been lately revisited by authors in [6] who have shown that these bounds are also solutions of a norm minimization under linear constraints (see [6]§III.B) analogous to the one introduce in section 2. Therefore any deterministic lower bounds have a corresponding Bayesian bound: Cramér-Rao bound, Bhattacharyya bound, Hammersley-Chapman-Robbins bound, ... Our idea is to argue from analogy from the Bayesian bounds towards the deterministic bounds to explore new possible bounds. As an example, in the case of a single unknown parameter θ - for sake of simplicity -, the Bayesian Weiss-Weinstein bound is associated with the linear constraint:

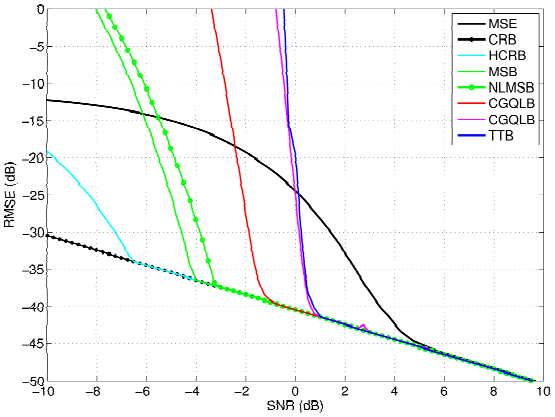


Fig. 1. Comparison of MSE lower bounds versus SNR ($M=8,\theta=0$)

$$\begin{aligned}\iint_{\Theta,\Omega}\left(\widehat{\theta}\left(\mathbf{x}\right)-\theta\right)\left[\frac{p(\mathbf{x},\theta+\delta)^q}{p(\mathbf{x},\theta)^q}-\frac{p(\mathbf{x},\theta-\delta)^{1-q}}{p(\mathbf{x},\theta)^{1-q}}\right]p\left(\mathbf{x},\theta\right)d\mathbf{x}d\theta\\ =-\delta\iint_{\Theta,\Omega}\frac{p(\mathbf{x},\theta-\delta)^{1-q}}{p(\mathbf{x},\theta)^{1-q}}p\left(\mathbf{x},\theta\right)d\mathbf{x}d\theta\end{aligned}$$

where $p\left(\mathbf{x},\theta\right)=p\left(\mathbf{x}\mid\theta\right)p\left(\theta\right)=p\left(\mathbf{x};\theta\right)p\left(\theta\right)$, $q\in[0,1]$. The corresponding linear constraint for deterministic estimation is (drawn from examples in [6]§III.B):

$$\begin{aligned}\int_{\Omega}\left(\widehat{\theta^0}\left(\mathbf{x}\right)-\theta^0\right)\left[\frac{p\left(\mathbf{x};\theta^0+\delta\right)^q}{p\left(\mathbf{x};\theta^0\right)^q}-\frac{p\left(\mathbf{x};\theta^0-\delta\right)^{1-q}}{p\left(\mathbf{x};\theta^0\right)^{1-q}}\right]p\left(\mathbf{x};\theta^0\right)d\mathbf{x}\\ =-\delta\int_{\Omega}\frac{p\left(\mathbf{x};\theta^0-\delta\right)^{1-q}}{p\left(\mathbf{x};\theta^0\right)^{1-q}}p\left(\mathbf{x};\theta^0\right)d\mathbf{x}\end{aligned} \tag{12}$$

leading to the deterministic Weiss-Weinstein bound (WWB):

$$MSE_{\theta^0}\left[\widehat{\theta^0}\right]\geq\sup_{q,\delta}\left\{\frac{\delta^2E_{\theta^0}\left[\left[\frac{p\left(\mathbf{x};\theta^0-\delta\right)^{1-q}}{p\left(\mathbf{x};\theta^0\right)^{1-q}}\right]^2\right]}{E_{\theta^0}\left[\left(\frac{p\left(\mathbf{x};\theta^0+\delta\right)^q}{p\left(\mathbf{x};\theta^0\right)^q}-\frac{p\left(\mathbf{x};\theta^0-\delta\right)^{1-q}}{p\left(\mathbf{x};\theta^0\right)^{1-q}}\right)^2\right]}\right\} \tag{13}$$

The WWB (13) is a bound for estimators satisfying (12). The conjecture is that unbiased estimators satisfy (12) as well. It is true where $q=0$ or $q=1$ since then (12) amounts to the Hammersley-Chapman-Robbins constraint. Unfortunately so far, we have not been able to prove that (12) derives from the mixture of integral linear and non-linear transformations of the unbiasedness constraint. And yet, but nevertheless simulations performed for the single tone threshold analysis clearly shows that the WWB (13) is a very tight bound for unbiased estimators; at least in this application case.

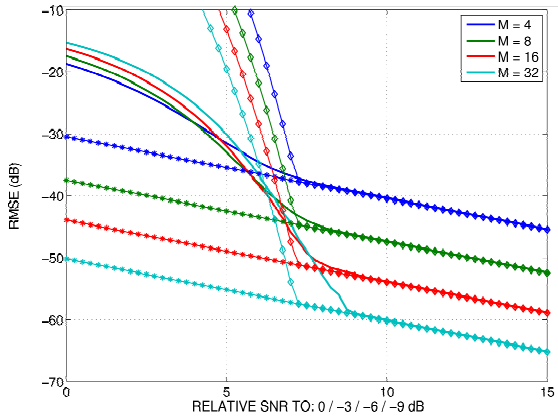


Fig. 2. MSE, CRB and WWB (Sup over $q \in [0,1]$) versus SNR

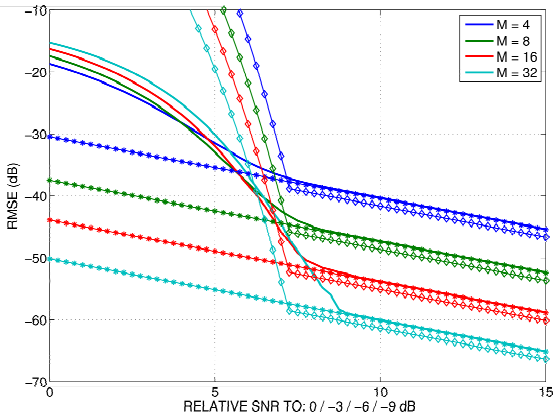


Fig. 3. MSE, CRB and WWB $\left(q = \frac{1}{2}\right)$ versus SNR

4. CONCLUSION: SINGLE TONE THRESHOLD ANALYSIS

A reference problem in threshold analysis is the estimation of a single tone $\theta \in]-0.5, 0.5[$ for a deterministic observation model:

$$\mathbf{x} = a\boldsymbol{\psi}\left(\theta^0\right) + \mathbf{n}, \quad \boldsymbol{\psi}(\theta) = \left[1, ..., e^{j(M-1)2\pi\theta}\right]^T \qquad p\left(\mathbf{x};\theta\right) = \frac{e^{-\|\mathbf{x}-a\boldsymbol{\psi}(\theta)\|^2}}{\pi^M} \tag{14}$$

where a^2 is the known SNR ($a > 0$) and $\mathbf{n}$ is a complex circular Gaussian noise, with zero mean and a known covariance matrix $\mathbf{C_n} = \mathbf{I}$. In the simulations:

- $\delta \in]0, 0.5[$, $\hat{\theta}_{ML} = \max_{\theta} \left\{ Re\left[\boldsymbol{\psi}(\theta)^H \mathbf{x}\right] \right\}$.
- the HCRB [2] is the simplest approximation of the BB (5) based on 2 test-points $\boldsymbol{\theta}^2 = \left(\theta^0, \theta^0 + \delta\right)^T$ + supremum on δ,
- the MSB [2] is the simplest approximation of the BB based on 3 test-points $\boldsymbol{\theta}^3 = \left(\theta^0, \theta^0 + \delta, \theta^0 - \delta\right)^T$ + supremum on δ,
- the NLMSB [5] is the nonlinear generalisation (10) of the MSB based on 3 test-points + supremum on δ and $q \in]0.5, 2[$,
- the CGQLB [2] is the generalization of the CRB based on 3 test-points $\boldsymbol{\theta}^3 = \left(\theta^0, \theta^0 + \delta, \theta^0 - \delta\right)^T$ + supremum on δ,
- the TTB [3] is the combination of $\mathrm{CRB}\left(\theta^0\right)$ and of (5) where $N = 1024$, $K = 32$ and $\mathbf{H}_K$ is an ad hoc submatrix of FFT matrix of dimension N.

All these lower bounds are displayed on figure (1) and compared with the MSE of the MLE estimator (5×10^5 trials) for $M = 8$ and $\theta^0 = 0$. The first occurrence of the CGQLB is obtained for δ lying on a discretization of $]-0.5, 0.5[$ with a step of $1/1024$. The second one is obtained for a step of $1/(1024 * 128)$. The purpose of the 2 cases is to show that it is generally difficult to compare tightness of bounds which are based on subsets of constraints that are not included one in each other. For each bound, tightness may depend on specific optimization parameters.
Additionally, the tightness of CGQLB and TTB (or any existing bound) could be improved by updating their associated linear constraints with the non-linear transformation (10) as we did for the MSB, which is a topic for future work.
Finally, a much more impartial criterion could be the computation time (and possibly the memory load).
Nevertheless, all these bounds seem to provide a still too coarse prediction of the SNR threshold value (underestimated by at least 5 dB), an imperfection mostly compensated by the WWB (13) as shown on figure (2). This figure clearly shows that the WWB is not only a lower bound for unbiased estimators whatever the value of M (also checked for $M = 2, 64, 128$), but it is an extremely tight lower bounds, far tighter that all the existing ones.
Moreover, the deterministic WWB seems to share the same property as its Bayesian analogue one, i.e. to be nearly the tightest for $q = 0.5$ as shown on figure (3). Under that form, the WWB is as simple to implement as the HCRB.
Such a simple and tight bound really deserves to be derived!

5. REFERENCES

[1] E.W. Barankin, "Locally best unbiased estimates", Ann. Math. Stat., vol. 20, no. 4, pp. 477-501, 1949.

[2] E. Chaumette, J. Galy, A. Quinlan, P. Larzabal, "A New Barankin Bound Approximation for the Prediction of the Threshold Region Performance of Maximum-Likelihood Estimators", IEEE Trans. on SP, vol. 56, no. 11, pp. 5319-5333, Nov. 2008

[3] K. Todros and J. Tabrikian, "A new lower bound on the mean-square error of unbiased estimators", in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., pp. 3913-3916, 2008

[4] H.L. Van Trees, *Detection, Estimation and Modulation Theory, Part 1*. New York: Wiley, 1968.

[5] E. Chaumette, A. Renaux, P. Larzabal, "Lower bounds on the mean square error derived from mixture of linear and non-linear transformations of the unbiasness definition", in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., pp. 3045-3048, 2009

[6] A. Renaux, P. Forster, P. Larzabal, C.D. Richmond, A. Nehorai, "A Fresh Look at the Bayesian Bounds of the Weiss-Weinstein Family", IEEE Trans. on SP, vol. 56, no. 11, pp. 5334-5352, Nov. 2008