On the Accuracy and Resolvability of Vector Parameter Estimates

Chengfang Ren, Mohammed Nabil El Korso, Jerome Galy, Eric Chaumette, Pascal Larzabal and Alexandre Renaux

Abstract

In this paper we address the problem of fundamental limitations on resolution in deterministic parameters estimation. We introduce a definition of resolvability based on probability and incorporating a requirement for accuracy unlike most existing definitions. Indeed in many application the key problem is to obtain distributions of estimates that are not only distinguishable but also accurate and compliant with a required precision. We exemplify the proposed definition with estimators that produce normal estimates, as in the conditional model for which the Gaussianity and efficiency of maximum likelihood estimators (MLEs) in the asymptotic region of operation (in terms of signal-to-noise ratio and/or in large number of snapshots) is well established, even for a single snapshot. In order to measure the convergence in distribution, we derive a simple test allowing to check whether the conditional MLEs operate in the

Chengfang Ren is with Universite Paris-Sud/LSS 3, Rue Joliot-Curie, 91192 Gif-sur-Yvette, France. Email: cren@lss.supelec.fr

Mohammed Nabil El Korso is with Laboratoire Energétique Mécanique Electromagnétisme (LEME EA 4416), Université Paris Ouest Nanterre La Défense, IUT de Ville d'Avray, France. Email : m.elkorso@u-paris10.fr

Jérome Galy is with Université de Montpellier 2/LIRMM, 161 rue Ada 34392 Montpellier Cedex 5, France. Email: galy@lirmm.fr

Eric Chaumette is with University of Toulouse-ISAE, Department of Electronics, Optronics and Signal, 10 Avenue Edouard Belin, 31055 Toulouse, France. Email: eric.chaumette@isae.fr

Pascal Larzabal is with Université Paris-Sud /SATIE, 61 av. du President Wilson, 94235 Cachan cedex, France. Email: pascal.larzabal@satie.ens-cachan.fr

Alexandre Renaux is with Universite Paris-Sud/LSS 3, Rue Joliot-Curie, 91192 Gif-sur-Yvette, France. Email: renaux@lss.supelec.fr

asymptotic region of operation. Last, we discuss the resolution of two complex exponentials with closely spaced frequencies and compare the results obtained with the ones provided by the various statistical resolution limit released in the open literature.

Index Terms

Statistical resolution limit, performance analysis, Cramer-Rao bound, parameter estimation, minimum probability of error, resolution.

I. INTRODUCTION

In many applications (as target classification in radar or identification of constellation diagrams in telecommunication) the key problem is to obtain distributions of estimates that are not only distinguishable but also accurate and compliant both with a required precision and a given probability of confidence, these two features being requested to assess the performance of subsequent processing (a classifier for instance). Hence the need of a definition of resolvability based on probability and incorporating a requirement for accuracy unlike most existing definitions. Indeed, strictly speaking, the idea of resolvability in the signal processing open literature does not necessarily take into account how well parameters are estimated, only whether multiple signals can be distinguished [1][2][3][4][5][6].

Interestingly enough, there is little contribution to this approach in papers or monographs [7][8][9][10]. Oh and Kashyap [11, Section V] seem to have introduced the first theoretical definition on the resolution based on probability and incorporating a requirement for accuracy, in the simplified case of two signal sources with a single unknown parameter (frequency of two complex exponentials) in order to characterize their separability (resolvability) by a proposed robust estimation method. Later, this work has been substantially extended by Clark [12] to normally distributed vector parameter estimates with covariance matrix equal to the Cramér-Rao bound (CRB). Since the CRB gives the minimum variance attainable by any unbiased estimates, it can be used to find a lower bound on the actual resolution threshold of any normally distributed estimators. Then, at the expense of neglecting the cross-information (cross-covariance) between estimators (as in [11]), Clark has shown that an extension of the definition on the resolution given by [11] to a vector of parameters for each signal source can be obtained with a metric depending on the distance between ellipsoids of constant probability. When the ellipsoids are disjoint the parameters are deemed resolvable. Based on the analysis of the distance problem, a simple method was developed for finding the noise level at which the ellipsoids become tangent, which is the resolution threshold noise level. Unfortunately, in many applications the hypothesis of uncorrelated

estimators is unrealistic and the cross-covariance terms of the associated CRB cannot be neglected (see for instance (29)). To bypass this limitation of Clark's approach, we present a definition of the probability of resolvability based on the cumulative distribution function (c.d.f.) which is the most general way to characterize a random vector, whichever its distribution. Thus, the proposed definition does not depend on the nature of the distributions of estimates (as in [11] and [12]) and allows to take into account their cross-information (unlike in [11] and [12]). As the evaluation (by calculus or by simulation) of a P-dimensional c.d.f. is generally difficult, we propose a general method to generate the lower and upper bounds of confidence intervals on the probability of resolvability for a given required estimation precision. This method is based on the approximation of hyper-rectangles by some ad hoc hypervolumes (generalization of hyper-ellipsoids in non-Euclidean norms) which can be used to substitute the computation of a 1-dimensional c.d.f. for the computation of a P-dimensional c.d.f..

We exemplify the proposed definition (and bounds usefulness) with estimators that produce normal estimates, as in the conditional model for which the Gaussianity and efficiency of MLEs in the asymptotic region of operation (in terms of signal-to-noise ratio (SNR) and/or in large number of snapshots) is well established, even for a single snapshot. Indeed for normal estimates, numerical methods exist (and some closed-form expressions as well in the simplest cases) to compute a confidence interval on the probability of resolvability. Additionally, the conditional model is the privileged model for an active system. In order to measure the convergence in distribution, we have derived a simple test allowing to check whether the conditional maximum likelihood estimators operate in the asymptotic region of operation.

Although the proposed resolvability definition is not limited to a given number of signal sources (see [13] for additional examples), we focus in the present paper on the resolvability of two closely spaced frequencies in order to compare the various definition of statistical resolution limit (SRL) released in the open literature and a SRL deriving from the proposed resolvability definition. The underlying idea is to measure the cost, in terms of scenario requirements (SNR, number of snapshots, number of sensors), of a SRL deriving from a resolvability definition which takes into account not only the ability to distinguish multiple signals (as all definitions do) but also how well parameters are estimated. In other words, we intend to bring to the reader attention that most previously released SRL criteria only take into account the ability to distinguish multiple signals and therefore do not take into account all aspects of performance from an estimation point of view. The significant differences in the region of operation (for a given observation model) requested to fulfill the different SRL criteria should question the reader on which information he looks for: distinguishability only or distinguishability and accuracy, since in order to get both there is a price to pay.

Last, from a practical point of view, the results obtained with the proposed rationale must be regarded as a "lower bound" in terms of actual scenario requirements (SNR, number of snapshots, number of sensors), in the sense that it assumes that the number of signals is known and that all the signals are present as well.

The paper is organized as follows. Section II is tutorial in nature. It presents the proposed resolvability definition and a general scheme to generate the lower and upper bounds of confidence intervals on the probability of resolvability for a given required estimation precision. In Section III we introduce a simple test allowing to check whether the conditional MLEs operate in the asymptotic region of operation in order to exploit its asymptotic Gaussianity and efficiency. Last in section IV, as an application example, we discuss the resolution of two complex exponentials with closely spaced frequencies in order to compare the various definition of SRL available in the open literature.

II. PROBABILITY OF RESOLVABILITY

In the following, $\mathbf{x}$ denotes the random observation vector of dimension N, Ω denotes the observation space and $L^2(\Omega)$ denotes the complex Hilbert space of square integrable functions over Ω. The probability density function (p.d.f.) of $\mathbf{x}$ is denoted $p(\mathbf{x}; \boldsymbol{\Theta})$ and depends on a vector of P real parameters $\boldsymbol{\Theta} = (\theta_1, \ldots, \theta_P) \in \Phi$, where Φ denotes the parameter space. The probability of an event $\mathcal{D} \subset \Omega$ is denoted $\mathcal{P}(\mathcal{D}; \boldsymbol{\Theta})$. Let $\boldsymbol{\Theta}^0$ be a selected value of the parameter $\boldsymbol{\Theta}$, and $\widehat{\mathbf{g}(\boldsymbol{\Theta}^0)}$ an estimator of $\mathbf{g}(\boldsymbol{\Theta}^0)$ where $\mathbf{g}(\boldsymbol{\Theta}) = (g_1(\boldsymbol{\Theta}), \ldots, g_Q(\boldsymbol{\Theta}))^T$ is a vector of Q real-valued functions of $\boldsymbol{\Theta}$. For any selected value $\boldsymbol{\Theta}^0$, $\widehat{\mathbf{g}(\boldsymbol{\Theta}^0)} \triangleq \widehat{\mathbf{g}(\boldsymbol{\Theta}^0)}(\mathbf{x})$ stands for a mapping of the observation space Ω into an estimate of $\mathbf{g}(\boldsymbol{\Theta}^0)$.

A. Approximation of a hyper-rectangle

In the following, $\boldsymbol{\Xi}^- = (\varepsilon_1^-, \ldots, \varepsilon_Q^-)^T$, $\boldsymbol{\Xi}^+ = (\varepsilon_1^+, \ldots, \varepsilon_Q^+)^T$, $\mathbf{s} = (s_1, \ldots, s_Q)^T$ are vectors of $\mathbb{R}^Q$ which components verify $\varepsilon_q^- > 0$, $\varepsilon_q^+ > 0$ and $s_q \geq 1$. Then $R(\mathbf{y}^0, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+)$, $\mathbf{y}^0 \in \mathbb{R}^Q$, denotes the hyper-rectangle (also called an orthotope or a box) of $\mathbb{R}^Q$ containing $\mathbf{y}^0$ defined by:

$$R\left(\mathbf{y}^0, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right) = \left\{ \mathbf{y} : \bigwedge_{q=1}^{Q} \left(y_q \in \left] y_q^0 - \varepsilon_q^-, y_q^0 + \varepsilon_q^+ \right[\right) \right\} \tag{1}$$

or equivalently by:

$$R\left(\mathbf{y}^0, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right) = \left\{ \mathbf{y} : \bigwedge_{q=1}^{Q} \left(\left| y_q - y_q^0 - \frac{d\varepsilon_q}{2} \right| < \overline{\varepsilon_q} \right) \right\} \tag{2}$$

where:

$$
\begin{aligned}
d\varepsilon_q &= \varepsilon_q^+ - \varepsilon_q^- = \left(\boldsymbol{\Xi}^+\right)_q - \left(\boldsymbol{\Xi}^-\right)_q = (d\boldsymbol{\Xi})_q, \\
\overline{\varepsilon_q} &= \frac{\varepsilon_q^+ + \varepsilon_q^-}{2} = \frac{\left(\boldsymbol{\Xi}^+\right)_q + \left(\boldsymbol{\Xi}^-\right)_q}{2} = \left(\overline{\boldsymbol{\Xi}}\right)_q.
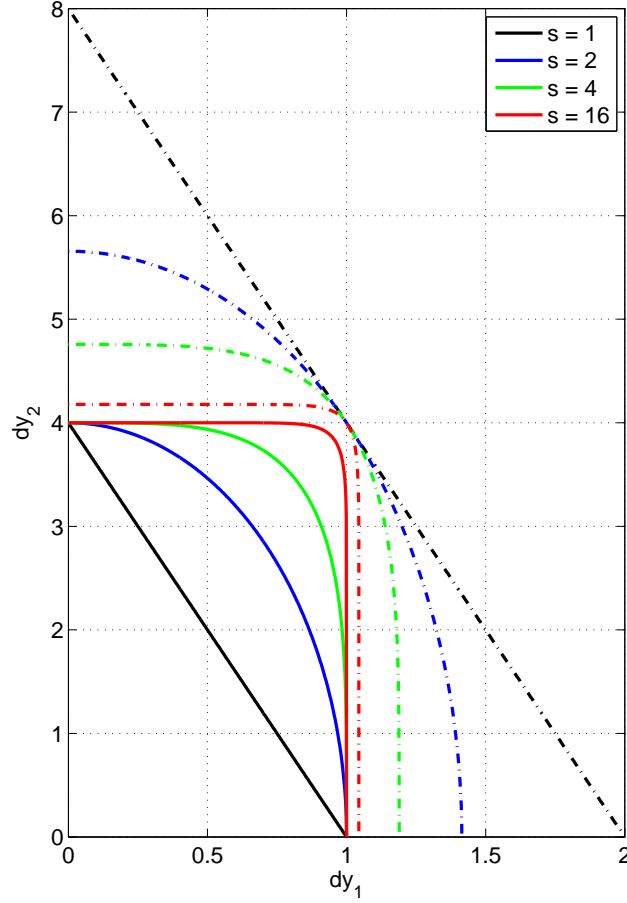\end{aligned}
$$

Fig. 1. Approximation of $R\left(\mathbf{y}^0, \boldsymbol{\Xi}\right) = \{\mathbf{y} : (|dy_1| < 1) \wedge (|dy_2| < 4)\}$, where $dy_1 = y_1 - y_1^0$ and $dy_2 = y_2 - y_2^0$, by $\Sigma_{\mathbf{s}}^{\lambda}\left(\mathbf{y}^0, \boldsymbol{\Xi}\right)$ where $\mathbf{s}^T = (s, s)$ and $\lambda \in \{1, 2\}$, for $s \in \{1, 2, 4, 16\}$, $\boldsymbol{\Xi} = (1, 4)^T$. The solid lines are for $\lambda = 1$ and the dash-dot lines are for $\lambda = 2$.

Let $\Sigma_{\mathbf{s}}^{\lambda}\left(\mathbf{y}^0, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right)$, $\lambda > 0$, be the Q-dimensional hypervolume defined by:

$$\Sigma_{\mathbf{s}}^{\lambda}\left(\mathbf{y}^0, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right) = \left\{\mathbf{y} : \sum_{q=1}^{Q} \left|\frac{y_q - y_q^0 - \frac{d\varepsilon_q}{2}}{\bar{\varepsilon}_q}\right|^{s_q} < \lambda\right\}. \tag{3}$$

$\Sigma_{\mathbf{s}}^{\lambda}\left(\mathbf{y}^0, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right)$ is the generalization of a hyper-ellipsoid obtained when $s_q = 2$. Let us note that $\mathbf{y}^0$ is the centre of $R\left(\mathbf{y}^0, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right)$ and $\Sigma_{\mathbf{s}}^{\lambda}\left(\mathbf{y}^0, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right)$ if and only if $\boldsymbol{\Xi}^- = \boldsymbol{\Xi}^+ = \boldsymbol{\Xi}$. In this case $R\left(\mathbf{y}^0, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right)$ and $\Sigma_{\mathbf{s}}^{\lambda}\left(\mathbf{y}^0, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right)$ are simply denoted $R\left(\mathbf{y}^0, \boldsymbol{\Xi}\right)$ and $\Sigma_{\mathbf{s}}^{\lambda}\left(\mathbf{y}^0, \boldsymbol{\Xi}\right)$. A first result is the existence of lower and upper bounds on $R\left(\mathbf{y}^0, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right)$ in the inclusion sense:

$$\Sigma_{\mathbf{s}}^{1}\left(\mathbf{y}^0, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right) \subset R\left(\mathbf{y}^0, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right) \subset \Sigma_{\mathbf{s}}^{Q}\left(\mathbf{y}^0, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right) \tag{4}$$

as shown in Appendix VI-A. Second, since:

$$\sum_{q=1}^{Q} \left| \frac{y_q - y_q^0 - \frac{d\varepsilon_q}{2}}{\overline{\varepsilon}_q} \right|^{s_q} < Q \Leftrightarrow \sum_{q=1}^{Q} \left| \frac{y_q - y_q^0 - \frac{d\varepsilon_q}{2}}{Q^{\frac{1}{s_q}} \overline{\varepsilon}_q} \right|^{s_q} < 1$$

and $\lim_{s \to \infty} Q^{1/s} = 1$, then:

$$\lim_{\min\left\{\{s_q\}_{q=1}^{Q}\right\} \to \infty} \Sigma_{\mathbf{s}}^{Q} \left(\mathbf{y}^0, \mathbf{\Xi}^-, \mathbf{\Xi}^+\right) - \Sigma_{\mathbf{s}}^{1} \left(\mathbf{y}^0, \mathbf{\Xi}^-, \mathbf{\Xi}^+\right) = 0$$

and:

$$\lim_{\min\left\{\{s_q\}_{q=1}^{Q}\right\} \to \infty} \Sigma_{\mathbf{s}}^{1} \left(\mathbf{y}^0, \mathbf{\Xi}^-, \mathbf{\Xi}^+\right) = R \left(\mathbf{y}^0, \mathbf{\Xi}^-, \mathbf{\Xi}^+\right). \tag{5}$$

Therefore any hyper-rectangle $R\left(\mathbf{y}^0, \mathbf{\Xi}^-, \mathbf{\Xi}^+\right)$ of $\mathbb{R}^Q$ can be lower and upper bounded by Q-dimensional hypervolumes $\Sigma_{\mathbf{s}}^{1}\left(\mathbf{y}^0, \mathbf{\Xi}^-, \mathbf{\Xi}^+\right)$ and $\Sigma_{\mathbf{s}}^{Q}\left(\mathbf{y}^0, \mathbf{\Xi}^-, \mathbf{\Xi}^+\right)$ that converge towards the hyper-rectangle as $\min\left\{\{s_q\}_{q=1}^{Q}\right\}$ increases. This is illustrated by figure (1) in the special case of the approximation of a rectangle ($Q = 2$, more precisely, the upper right quarter of the rectangle) by $\Sigma_{\mathbf{s}}^{\lambda}\left(\mathbf{y}^0, \mathbf{\Xi}\right)$, $\mathbf{s}^T = (s, s)$, $\lambda \in \{1, 2\}$ (solid line for $\lambda = 1$ and dash-dot line for $\lambda = 2$) showing how the increase of s allows very tight bounds (4).

B. Precision in probability sense and bounds

The quality (i.e. the precision) of an estimator $\widehat{\mathbf{g}\left(\mathbf{\Theta}^0\right)}$ can be measured using the following exhaustive canonical risk function:

$$\mathcal{O}_{\mathbf{\Theta}^0}\left(\widehat{\mathbf{g}\left(\mathbf{\Theta}^0\right)}, \mathbf{\Xi}^-, \mathbf{\Xi}^+\right) = \mathcal{P}\left(\widehat{\mathbf{g}\left(\mathbf{\Theta}^0\right)} \in R\left(\mathbf{g}\left(\mathbf{\Theta}^0\right), \mathbf{\Xi}^-, \mathbf{\Xi}^+\right); \mathbf{\Theta}^0\right) \tag{6}$$

where $\mathbf{\Xi}^-$ and $\mathbf{\Xi}^+$ define the left and right errors on the estimation of $\mathbf{g}\left(\mathbf{\Theta}^0\right)$, and $\mathcal{O}_{\mathbf{\Theta}^0}\left(\widehat{\mathbf{g}\left(\mathbf{\Theta}^0\right)}, \mathbf{\Xi}^-, \mathbf{\Xi}^+\right)$[1] is the probability that errors do not exceed $\mathbf{\Xi}^-$ and $\mathbf{\Xi}^+$. Indeed according to (2):

$$\mathcal{O}_{\mathbf{\Theta}^0}\left(\widehat{\mathbf{g}\left(\mathbf{\Theta}^0\right)}, \mathbf{\Xi}^-, \mathbf{\Xi}^+\right) = F_{\mathbf{y}}\left(\overline{\mathbf{\Xi}}\right) \text{ where } (\mathbf{y})_q = \left| \widehat{g_q\left(\mathbf{\Theta}^0\right)}(\mathbf{x}) - g_q\left(\mathbf{\Theta}^0\right) - \frac{d\varepsilon_q}{2} \right|, \tag{7}$$

i.e. $\mathcal{O}_{\mathbf{\Theta}^0}\left(\widehat{\mathbf{g}\left(\mathbf{\Theta}^0\right)}, \mathbf{\Xi}^-, \mathbf{\Xi}^+\right)$ is the c.d.f. of a random vector. This risk function is termed "canonical" since it is deduced naturally from the problem under study: the match between the observations of a random vector and a deterministic vector of interest. This risk function is also *exhaustive*, in the sense that it incorporates all the available information on the problem, in other words the probabilities. In

---

[1]$\mathbf{\Xi}^- \neq \mathbf{\Xi}^+$ takes into account that the p.d.f. of $\widehat{g_q\left(\boldsymbol{\theta}^0\right)}(\mathbf{x})$ is not expected to be symmetric at the vicinity of $g_q\left(\boldsymbol{\theta}^0\right)$ in the general case.

the following $\mathcal{O}_{\boldsymbol{\Theta}^0}\left(\widehat{\mathbf{g}\left(\boldsymbol{\Theta}^0\right)}, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right)$ (6) defines the precision of estimator $\widehat{\mathbf{g}\left(\boldsymbol{\Theta}^0\right)}$ in probability sense which is bounded by (4):

$$
\begin{aligned}
\mathcal{P}\left(\widehat{\mathbf{g}\left(\boldsymbol{\Theta}^0\right)} \in \Sigma_{\mathbf{s}}^1\left(\mathbf{g}\left(\boldsymbol{\Theta}^0\right), \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right); \boldsymbol{\Theta}^0\right) &\leq \mathcal{O}_{\boldsymbol{\Theta}^0}\left(\widehat{\mathbf{g}\left(\boldsymbol{\Theta}^0\right)}, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right) \\
&\leq \mathcal{P}\left(\widehat{\mathbf{g}\left(\boldsymbol{\Theta}^0\right)} \in \Sigma_{\mathbf{s}}^Q\left(\mathbf{g}\left(\boldsymbol{\Theta}^0\right), \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right); \boldsymbol{\Theta}^0\right)
\end{aligned}
\tag{8}
$$

The underlying idea behind the use of probability bounds (8) is to substitute the computation of the c.d.f. of one random variable for the computation of a Q-dimensional integral over a Q-dimensional hyper-rectangle. In the following, for the sake of legibility:

- without loss of generality, $\mathbf{g}\left(\boldsymbol{\Theta}\right) = \boldsymbol{\Theta}$ ($Q = P$),
- $\boldsymbol{\Theta}^0$ (respectively $\widehat{\boldsymbol{\Theta}^0}$) is denoted $\boldsymbol{\Theta}$ (respectively $\widehat{\boldsymbol{\Theta}}$) wherever it is unambiguous.

C. Probability of resolvability

We consider a parameter estimation problem where the parameters of interest are the vectors $\{\boldsymbol{\theta}^m\}_{m=1}^M$, where $\boldsymbol{\theta} \in \mathbb{R}^{P'}$ and $\boldsymbol{\theta}^m \neq \boldsymbol{\theta}^l$, $\forall l \neq m \in [1, M]$. Among all various possible realizations of this setting, the most studied realization in signal processing is that of separating the components of data formed from a linear superposition of individual signals to noisy data (nuisance). Each vector $\boldsymbol{\theta}^m$ parameterizes a single signal of this superposition.

In an ideal estimation problem, the set of estimators $\left\{\widehat{\boldsymbol{\theta}^m}\right\}_{m=1}^M$ would yield the set of unknown parameter vectors $\{\boldsymbol{\theta}^m\}_{m=1}^M$ for all trials : $\widehat{\boldsymbol{\theta}^m} = \boldsymbol{\theta}^m$. However in an actual estimation problem, the set of estimators $\left\{\widehat{\boldsymbol{\theta}^m}\right\}_{m=1}^M$ is a set of random vectors. Therefore given the statistics of the estimates, we desire to know when vector parameters $\{\boldsymbol{\theta}^m\}_{m=1}^M$, which may be closely spaced in $\mathbb{R}^{P'}$, can be resolved by the realizations of $\left\{\widehat{\boldsymbol{\theta}^m}\right\}_{m=1}^M$. So far, most of the definitions of resolvability that have already been proposed (except [12], see Section IV) focuses only on the ability to distinguish the distributions of estimates without any particular requirement on the estimates precision. However in many applications (as target classification in radar or identification of constellation diagram in telecommunication), the key problem is rather to assign the correct vector of estimates to each signal source with a required precision compliant with subsequent processing. This is the problem we want to address by introducing a definition of resolvability that combines the ability to distinguish[2] the distributions of estimates and to measure their precision (see also [12]). The rationale is the following. Let $\boldsymbol{\Theta}^T = \left(\left(\boldsymbol{\theta}^1\right)^T, \ldots, \left(\boldsymbol{\theta}^M\right)^T\right)$ ($P = P'M$), $\left(\boldsymbol{\Xi}^+\right)^T = \left(\left(\boldsymbol{\varepsilon}^{1+}\right)^T, \ldots, \left(\boldsymbol{\varepsilon}^{M+}\right)^T\right)$ and $\left(\boldsymbol{\Xi}^-\right)^T = \left(\left(\boldsymbol{\varepsilon}^{1-}\right)^T, \ldots, \left(\boldsymbol{\varepsilon}^{M-}\right)^T\right)$ where $\boldsymbol{\varepsilon}^{m+} =$

---

[2]A similar idea appears in [14] in connection with the definition of optimally distinguishable distributions in order to obtain a partition in the model space via a partition in the parameter space.

$\left(\varepsilon_1^{m+}, \ldots, \varepsilon_{P'}^{m+}\right)^T$ and $\boldsymbol{\varepsilon}^{m-} = \left(\varepsilon_1^{m-}, \ldots, \varepsilon_{P'}^{m-}\right)$. Any choice of $\boldsymbol{\Xi}^+$ and $\boldsymbol{\Xi}^-$ in such a manner that the set of M hyper-rectangles $R\left(\boldsymbol{\theta}^m, \boldsymbol{\varepsilon}^{m-}, \boldsymbol{\varepsilon}^{m+}\right)$ are disjoint, i.e.

$$R\left(\boldsymbol{\theta}^m, \boldsymbol{\varepsilon}^{m-}, \boldsymbol{\varepsilon}^{m+}\right) \cap R\left(\boldsymbol{\theta}^l, \boldsymbol{\varepsilon}^{l-}, \boldsymbol{\varepsilon}^{l+}\right) = \varnothing, \ l \neq m, \tag{9}$$

defines a resolvability criterion, in the sense that any realization $\left\{\widehat{\boldsymbol{\theta}^m}\right\}_{m=1}^M$ that yields an unique estimate per $R\left(\boldsymbol{\theta}^m, \boldsymbol{\varepsilon}^{m-}, \boldsymbol{\varepsilon}^{m+}\right)$ can be regarded as successful. Nevertheless such a resolvability criterion does not guaranty that for all successful trials $\widehat{\boldsymbol{\theta}^m} \in R\left(\boldsymbol{\theta}^m, \boldsymbol{\varepsilon}^{m-}, \boldsymbol{\varepsilon}^{m+}\right), 1 \leq m \leq M$, since some estimate switches among hyper-rectangles $R\left(\boldsymbol{\theta}^m, \boldsymbol{\varepsilon}^{m-}, \boldsymbol{\varepsilon}^{m+}\right)$ are statistically possible. And this is exactly what we do not want to allow. Therefore we consider as successful only trials for which:

$$\widehat{\boldsymbol{\theta}^m} \in R\left(\boldsymbol{\theta}^m, \boldsymbol{\varepsilon}^{m-}, \boldsymbol{\varepsilon}^{m+}\right), 1 \leq m \leq M, \ \text{s.t.}$$
$$R\left(\boldsymbol{\theta}^m, \boldsymbol{\varepsilon}^{m-}, \boldsymbol{\varepsilon}^{m+}\right) \cap R\left(\boldsymbol{\theta}^l, \boldsymbol{\varepsilon}^{l-}, \boldsymbol{\varepsilon}^{l+}\right) = \varnothing, \ l \neq m$$

leading to the following probability of resolvability $\mathcal{P}_{\text{res}}$:

$$\mathcal{P}_{\text{res}} = \mathcal{O}_{\boldsymbol{\Theta}}\left(\widehat{\boldsymbol{\Theta}}, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right) \ \text{where} \ \left(\boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right) \ \text{s.t.} \tag{10}$$
$$R\left(\boldsymbol{\theta}^m, \boldsymbol{\varepsilon}^{m-}, \boldsymbol{\varepsilon}^{m+}\right) \cap R\left(\boldsymbol{\theta}^l, \boldsymbol{\varepsilon}^{l-}, \boldsymbol{\varepsilon}^{l+}\right) = \varnothing, \ l \neq m.$$

In an ideal estimation problem:

$$\mathcal{P}_{\text{res}} = 1 = \mathcal{O}_{\boldsymbol{\Theta}}\left(\widehat{\boldsymbol{\Theta}}, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right), \forall \left(\boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right) \ \text{s.t.}$$
$$R\left(\boldsymbol{\theta}^m, \boldsymbol{\varepsilon}^{m-}, \boldsymbol{\varepsilon}^{m+}\right) \cap R\left(\boldsymbol{\theta}^l, \boldsymbol{\varepsilon}^{l-}, \boldsymbol{\varepsilon}^{l+}\right) = \varnothing, \ l \neq m.$$

In an actual estimation problem, statistics of the estimates depends on the observation model and of its conditions of operation (SNR of sources, number of independent observations, ...). The requirements on precision (choice of $\boldsymbol{\Xi}_{\text{req}}^-$ and $\boldsymbol{\Xi}_{\text{req}}^+$) and on the probability of resolvability $\mathcal{P}_{\text{res}}$ (10) generally come from subsequent processing for which the set $\left\{\widehat{\boldsymbol{\theta}^m}\right\}_{m=1}^M$ is an input: they are application-defined parameters. Therefore for a given application, vectors of parameters $\{\boldsymbol{\theta}^m\}_{m=1}^M$ will be said "resolved" by estimators $\left\{\widehat{\boldsymbol{\theta}^m}\right\}_{m=1}^M$ with probability $\mathcal{P}_{\text{res}}$ if:

$$\mathcal{O}_{\boldsymbol{\Theta}}\left(\widehat{\boldsymbol{\Theta}}, \boldsymbol{\Xi}_{\text{req}}^-, \boldsymbol{\Xi}_{\text{req}}^+\right) = \mathcal{P}_{\text{res}} \ \text{where} \ \left(\boldsymbol{\Xi}_{\text{req}}^-, \boldsymbol{\Xi}_{\text{req}}^+\right) \ \text{s.t.} \tag{11}$$
$$R\left(\boldsymbol{\theta}^m, \boldsymbol{\varepsilon}_{\text{req}}^{m-}, \boldsymbol{\varepsilon}_{\text{req}}^{m+}\right) \cap R\left(\boldsymbol{\theta}^l, \boldsymbol{\varepsilon}_{\text{req}}^{l-}, \boldsymbol{\varepsilon}_{\text{req}}^{l+}\right) = \varnothing, \ l \neq m.$$

As $\mathcal{P}_{\text{res}} = \mathcal{O}_{\boldsymbol{\Theta}}\left(\widehat{\boldsymbol{\Theta}}, \boldsymbol{\Xi}_{\text{req}}^-, \boldsymbol{\Xi}_{\text{req}}^+\right)$ is generally not computable, the use of bounds (8) whenever they can be computed more easily (for example if $\widehat{\boldsymbol{\Theta}}$ is a Gaussian random vector, see hereinafter) may allow to

check whether resolvability requirements (11) are fulfilled, that is to check whether:

$$\mathcal{P}\left(\widehat{\boldsymbol{\Theta}} \in \Sigma_{\mathbf{s}}^1\left(\boldsymbol{\Theta}, \boldsymbol{\Xi}_{\mathrm{req}}^-, \boldsymbol{\Xi}_{\mathrm{req}}^+\right); \boldsymbol{\Theta}^0\right) \leq \mathcal{P}_{\mathrm{res}} \leq \mathcal{P}\left(\widehat{\boldsymbol{\Theta}} \in \Sigma_{\mathbf{s}}^P\left(\boldsymbol{\Theta}, \boldsymbol{\Xi}_{\mathrm{req}}^-, \boldsymbol{\Xi}_{\mathrm{req}}^+\right); \boldsymbol{\Theta}^0\right) \tag{12}$$

Conversely, the use of bounds (12), whenever they are computable, allows to obtain a confidence interval on the probability of resolvability $\mathcal{P}_{\mathrm{res}}$ for a given required precision, confidence interval that can be taken into account in the assessment of performance of subsequent processing. If $\mathcal{P}\left(\widehat{\boldsymbol{\Theta}} \in \Sigma_{\mathbf{s}}^\lambda\left(\boldsymbol{\Theta}, \boldsymbol{\Xi}_{\mathrm{req}}^-, \boldsymbol{\Xi}_{\mathrm{req}}^+\right); \boldsymbol{\Theta}^0\right)$ is computable for a sub-set of values of $\mathbf{s}$, the tightness of the confidence interval is improved by increasing $\min\left\{\{s_p\}_{p=1}^P\right\}$. The possible exploitation of independence among estimates can improve the tightness of the confidence interval on $\mathcal{P}_{\mathrm{res}}$ as well. Indeed if $\{\boldsymbol{\theta}^m\}_{m=1}^M$ are jointly independent, then:

$$\mathcal{O}_{\boldsymbol{\Theta}}\left(\widehat{\boldsymbol{\Theta}}, \boldsymbol{\Xi}_{\mathrm{req}}^-, \boldsymbol{\Xi}_{\mathrm{req}}^+\right) = \prod_{m=1}^M \mathcal{O}_{\boldsymbol{\theta}^m}\left(\widehat{\boldsymbol{\theta}^m}, \varepsilon_{\mathrm{req}}^{m-}, \varepsilon_{\mathrm{req}}^{m+}\right), \tag{13}$$

leading to:

$$\prod_{m=1}^M \mathcal{P}\left(\widehat{\boldsymbol{\theta}^m} \in \Sigma_{\mathbf{s}}^1\left(\boldsymbol{\theta}^m, \varepsilon_{\mathrm{req}}^{m-}, \varepsilon_{\mathrm{req}}^{m+}\right); \boldsymbol{\theta}^m\right) \leq \mathcal{P}_{\mathrm{res}} \leq \prod_{m=1}^M \mathcal{P}\left(\widehat{\boldsymbol{\theta}^m} \in \Sigma_{\mathbf{s}}^{P'}\left(\boldsymbol{\theta}^m, \varepsilon_{\mathrm{req}}^{m-}, \varepsilon_{\mathrm{req}}^{m+}\right); \boldsymbol{\theta}^m\right) \tag{14}$$

which are tighter bounds than (12). An even more thorough exploitation of independence has been proposed by Clark [12] when $\left\{\widehat{\boldsymbol{\theta}^m}\right\}_{m=1}^M$ are jointly Gaussian with bias vectors $\mathbf{b}^m$ and covariance matrices $\mathbf{C}^m$: $\widehat{\boldsymbol{\theta}^m} \sim \mathcal{N}\left(\boldsymbol{\theta}^m + \mathbf{b}^m, \sigma^2 \mathbf{C}^m\right)$. Then, as noticed by Clark, it is worth substituting ellipsoids for hyper-rectangles in the definition of the precision in probability sense. Indeed if we consider the following ellipsoids of constant probability:

$$E\left(\boldsymbol{\theta}^m, \sigma r\right) = \left\{\boldsymbol{\theta} : \|\boldsymbol{\theta} - (\boldsymbol{\theta}^m + \mathbf{b}^m)\|_{2;\mathbf{C}^m}^2 < \sigma^2 r^2\right\}$$

where $\|\mathbf{u}\|_{2;\mathbf{A}}^2 = \left(\|\mathbf{u}\|_{2;\mathbf{A}}\right)^2 \triangleq \mathbf{u}^T \mathbf{A}^{-1} \mathbf{u}$ and $\mathbf{A}$ is a symmetric positive definite matrix, then $\mathcal{P}\left(\widehat{\boldsymbol{\theta}^m} \in E\left(\boldsymbol{\theta}^m, \sigma r\right); \boldsymbol{\theta}^m\right) = \mathcal{P}(r)$ and:

$$\mathcal{P}\left(\bigcap_{m=1}^M \left(\widehat{\boldsymbol{\theta}^m} \in E\left(\boldsymbol{\theta}^m, \sigma r\right)\right); \boldsymbol{\Theta}\right) = \mathcal{P}(r)^M \tag{15}$$

When the M ellipsoids with a given probability $\mathcal{P}(r)$ are disjoint, the parameters are deemed resolvable with probability $\mathcal{P}_{\mathrm{res}} = \mathcal{P}(r)^M$. Therefore when $\{\boldsymbol{\theta}^m\}_{m=1}^M$ are jointly independent, Clark's approach allows to compute the exact $\mathcal{P}_{\mathrm{res}}$ instead of a confidence interval (14), but at the expense of a more complex method to check that $E\left(\boldsymbol{\theta}^m, \sigma r\right)$ are disjoint (in comparison with hyper-rectangle). Based on the analysis of the distance problem between ellipsoids, Clark provides a method that allows to find the noise level σ^2 at which the ellipsoids become tangent with probability $\mathcal{P}_{\mathrm{res}}$, which is the resolution threshold noise level σ_{res}^2. Unfortunately the assumption of vector of estimates decorrelated (independent) from signal to signal is generally not verified; as a consequence it is generally impossible to predict the accuracy of (15) as an approximation of the exact probability of resolvability.

D. Gaussian p.d.f.

The (lower and upper) bounds on $\mathcal{O}_{\boldsymbol{\Theta}}\left(\widehat{\boldsymbol{\Theta}}, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right)$ given by (8) are easily computable when $\widehat{\boldsymbol{\Theta}}\left(\mathbf{x}\right) - \boldsymbol{\Theta} \sim \mathcal{N}\left(\mathbf{b}\left(\boldsymbol{\Theta}\right), \mathbf{C}\left(\boldsymbol{\Theta}\right)\right)$ and $\mathbf{s} = \mathbf{2}_P$, where $\boldsymbol{\alpha}_P$ is a P-dimensional vector with components equal to $\alpha \in \mathbb{R}$, that is $\widehat{\boldsymbol{\Theta}}$ is a Gaussian estimator of $\boldsymbol{\Theta}$ with bias vector $\mathbf{b}\left(\boldsymbol{\Theta}\right)$ and covariance matrix $\mathbf{C}\left(\boldsymbol{\Theta}\right)$ and we consider the approximation of a hyper-rectangle by a hyper-ellipsoid:

$$\Sigma_{\mathbf{2}_P}^{\lambda}\left(\boldsymbol{\Theta}, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right) = \left\{ \widehat{\boldsymbol{\Theta}}\left(\mathbf{x}\right) : \left\|\mathbf{D}_{\overline{\boldsymbol{\Xi}}}^{-1}\left(\widehat{\boldsymbol{\Theta}}\left(\mathbf{x}\right) - \boldsymbol{\Theta} - d\boldsymbol{\Xi}\right)\right\|_2^2 < \lambda \right\} \tag{16}$$

where $\|\mathbf{y}\|_2^2 = \left(\|\mathbf{y}\|_2\right)^2 = \mathbf{y}^T\mathbf{y}$ and $\mathbf{D}_{\boldsymbol{\alpha}}$ is the diagonal matrix with the components of vector $\boldsymbol{\alpha}$ on the main diagonal: $\left(\mathbf{D}_{\boldsymbol{\alpha}}\right)_{q,p} = \alpha_q \delta_p^q$. As the following unitary transformation of the Gaussian random vector $\widehat{\boldsymbol{\Theta}}$:

$$\widehat{\mathbf{O}}\left(\mathbf{x}\right) = \mathbf{U}_{\overline{\boldsymbol{\Xi}}}^T(\boldsymbol{\Theta})\,\mathbf{D}_{\overline{\boldsymbol{\Xi}}}^{-1}\left(\widehat{\boldsymbol{\Theta}}\left(\mathbf{x}\right) - \boldsymbol{\Theta} - d\boldsymbol{\Xi}\right) \mid \mathbf{D}_{\overline{\boldsymbol{\Xi}}}^{-1}\mathbf{C}\left(\boldsymbol{\Theta}\right)\mathbf{D}_{\overline{\boldsymbol{\Xi}}}^{-1} = \mathbf{U}_{\overline{\boldsymbol{\Xi}}}(\boldsymbol{\Theta})\,\mathbf{D}_{\sigma_{\overline{\boldsymbol{\Xi}}}^2(\boldsymbol{\Theta})}\mathbf{U}_{\overline{\boldsymbol{\Xi}}}^T(\boldsymbol{\Theta})$$

preserves the norm $\|\ \|_2$, then:

$$\mathcal{P}\left(\widehat{\boldsymbol{\Theta}} \in \Sigma_{\mathbf{2}_P}^{\lambda}\left(\boldsymbol{\Theta}, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right); \boldsymbol{\Theta}\right) = \mathcal{P}\left(\left\|\widehat{\mathbf{O}}\left(\mathbf{x}\right)\right\|_2^2 < \lambda; \boldsymbol{\Theta}\right) = F_{Q_P}\left(\lambda; \left\|\boldsymbol{\delta}\left(\boldsymbol{\Theta}, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right)\right\|_2^2, \sigma_{\overline{\boldsymbol{\Xi}}}^2(\boldsymbol{\Theta})\right) \tag{17}$$

where $\boldsymbol{\delta}\left(\boldsymbol{\Theta}, \boldsymbol{\Xi}^-, \boldsymbol{\Xi}^+\right) = \mathbf{U}_{\overline{\boldsymbol{\Xi}}}^T(\boldsymbol{\Theta})\,\mathbf{D}_{\overline{\boldsymbol{\Xi}}}^{-1}\left(\mathbf{b}\left(\boldsymbol{\Theta}\right) - d\boldsymbol{\Xi}\right)$ and $F_{Q_P}\left(v; \mu, \sigma^2\right)$ is the c.d.f. of a quadratic form for non-central normal variates [15], that is an extension of non-central chi-square with corresponding degrees of freedom in P and positive noncentrality parameters in μ where the power of each component is not constant: $Q_P \sim \sum_{p=1}^{P} \sigma_p^2 \left| z_p + \sqrt{\mu_p} \right|^2$ where $\boldsymbol{\sigma}^2 = \left(\sigma_1^2, \ldots, \sigma_P^2\right)^T$, $\mu = \sum_{p=1}^{P} \mu_p$ and $\mathbf{z} = \left(z_1, \ldots, z_P\right)^T \sim \mathcal{N}\left(\mathbf{0}, \mathbf{I}\right)$. If $\mu = 0$ (unbiased estimates and symmetric errors $\boldsymbol{\Xi}^- = \boldsymbol{\Xi}^+$) then [16]:

$$F_{Q_P}\left(v; 0, \boldsymbol{\sigma}^2\right) = F_P\left(\frac{v}{2}; \boldsymbol{\sigma}^2\right) \tag{18}$$

$$F_P\left(t; \mathbf{a}\right) = \frac{1}{\sqrt{|\mathbf{D_a}|}} \frac{t^{\frac{P}{2}}}{\pi^{\frac{P}{2}}} \sum_{k=0}^{\infty} \frac{w_P^k \left(-t\right)^k}{\Gamma\left(\frac{P}{2} + k + 1\right)}, \quad w_P^k = \sum_{i_1 + \ldots + i_P = k} \frac{\Gamma\left(i_1 + \frac{1}{2}\right) \ldots \Gamma\left(i_n + \frac{1}{2}\right)}{i_1! \ldots i_n!\, a_1^{i_1} \ldots a_P^{i_P}}$$

where $\Gamma\left(t\right)$ is the Gamma function. The general formula for the c.d.f. of Q_P is given in [15]. Unfortunately series expansion of the c.d.f. of a quadratic form for central (18) or non-central normal variates given in [15][16] are not easy to program. Therefore for small P it is worth considering the alternative numerical method consisting in computing the p.d.f. of Q_P by iterative convolutions since:

$$p_{Q_P}\left(v; \mu, \boldsymbol{\sigma}^2\right) = p_{\chi_1^2}\left(v; \mu_1, \sigma_1^2\right) * \ldots * p_{\chi_1^2}\left(v; \mu_P, \sigma_P^2\right)$$

where each χ_1^2 p.d.f. is given by [18]:

$$p_{\chi_1^2}\left(v; \mu_p, \sigma_p^2\right) = \frac{e^{-\frac{v + \mu_p}{2\sigma_p^2}}}{2\sigma_p^2} \left(\frac{v\mu_p}{\sigma_p^4}\right)^{-\frac{1}{4}} I_0\left(\sqrt{\frac{v\mu_p}{\sigma_p^4}}\right) \tag{19}$$

where $I_0(z)$ is the modified Bessel function of the first kind with order zero.

It should be noted that the tightness of the confidence interval (12) on the probability of resolvability $\mathcal{P}_{\text{res}}$ decreases, for a given $\mathbf{s}$, as the product $P'M$ increases. Therefore in some applications it may be necessary to resort to values of $\mathbf{s}$ larger than 2 in order to obtain a tight enough confidence interval, as illustrated in figure (1). Unfortunately in that case we have not found in the open literature a simple numerical method to compute the c.d.f. of $\Sigma_{\mathbf{s}}^{\lambda}\left(\boldsymbol{\Theta}, \boldsymbol{\Xi}_{\text{req}}^{-}, \boldsymbol{\Xi}_{\text{req}}^{+}\right)$ for correlated $\widehat{\boldsymbol{\Theta}}(\mathbf{x})$; nevertheless it is still possible (but computationally expensive) to assess directly $\mathcal{P}_{\text{res}}$ by resorting to algorithms proposed by Genz [17] for numerical evaluation of multivariate normal distributions.

III. Asymptotic performance of conditional model

Historically the first MSE lower bound for deterministic parameters to be derived was the CRB, which was introduced to investigate fundamental limits of a parameter estimation problem or to assess the relative performance of a specific estimator (efficiency) [9]. It has since become the most popular lower bound due to its simplicity of calculation and the fact that in many cases it can be achieved asymptotically (in terms of high SNR [19] and/or large number of snapshots [9]) by MLEs. Additionally, it is well known that in non-linear estimation problems three distinct regions of operation can be observed [20][21][22]. In the asymptotic region, the MSE of MLEs is small and, in many cases, close to the CRB. 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 an additional ambiguity region, also called the transition region. In this region, the MSE of MLEs usually deteriorates rapidly with respect to CRB and exhibits a threshold behaviour corresponding to a "performance breakdown". The nature of this phenomenon is specified by a complicated non-smooth behaviour of the likelihood function in the "threshold" area where it tends to generate outliers.

A. Characterization of the asymptotic region of operation for the conditional model

The choice of focusing on the (Gaussian) conditional model comes from the asymptotic (in terms of SNR [19] and/or in large number of snapshots [9]) Gaussianity and efficiency of CMLEs in the multiple parameters case (what is not true for the unconditional model [23]). Moreover the conditional model is the privileged model for an active system. A typical example of an active system is a localization system such as a radar or a sonar where a known waveform is transmitted, and the signals scattered from the targets of interest are used to estimate their parameters. In an active system, as the waveform parametric

model is known and deterministic (in opposition with a passive system where a probabilistic modelling of the waveform is generally considered), the most accurate statistical prediction for an observation will be obtained when considering the signal amplitudes as deterministic (since it is well known that the complex Gaussian amplitude modelling provide an average unconditional CRB higher than the corresponding conditional CRB [24]). Let us remind that the asymptotic Gaussianity and efficiency of CMLEs in the multiple parameters case has been proved under the assumption that the maximum of the (reduced) log likelihood function belongs to its main lobe [19], or equivalently that the probability of outlier is equal to zero [20][21]. An approximation of the probability of outlier for the three regions of operation of CMLE has already been derived in the form of the probability of the union of events that one of the sidelobe peaks is higher than the mainlobe peak, probability itself approximated by the union bound [21, Section IV]. The main result is that the derived approximation of probability of outlier is quite accurate in modeling the performance of CMLE. Only in the no-information region does the approximation deviate from the simulation results: this is due to the fact that the union bound tightness decreases as the observation model enters the transition region. Nevertheless, in the multi-source case, the computation is rather cumbersome. Indeed, first, one must look for the positions of the sidelobe peaks which are usually not available in closed form and must be calculated by some numerical method. And second, one must compute multiples c.d.f. of noncentral indefinite quadratic form in complex Gaussian random variables.

A far simpler lower bound on the probability of outlier is proposed hereinafter (25). Its advantage is its computational simplicity but at the expense of the information provided: the proposed lower bound is expected to be tight only in the asymptotic region. We focus on the general complex (circular) linear conditional model where the noise correlation matrix is supposed to be known up to a scale factor $\sigma_{\mathbf{n}}^2$ [24]. For the sake of legibility and without loss of generality, we consider the following simplified instance (narrow band signals and spatially white noise):

$$\mathbf{x}_t\left(\mathbf{\Theta}^0\right) = \mathbf{A}\left(\mathbf{\Theta}^0\right)\mathbf{s}_t + \mathbf{n}_t, \quad t \in [1, T] \tag{20}$$

where T is the number of independent observations, M is the number of signal sources, $\mathbf{s}_t = (s_{t,1}, \ldots, s_{t,M})^T$ is the vector of complex amplitudes of the M sources for the tth observation, $\mathbf{A}\left(\mathbf{\Theta}\right) = \left[\mathbf{a}\left(\boldsymbol{\theta}^1\right), \ldots, \mathbf{a}\left(\boldsymbol{\theta}^M\right)\right]$ and $\mathbf{a}\left(\ \right)$ is a vector of N parametric functions depending on the parameters vector $\boldsymbol{\theta}$, $\mathbf{n}_t$ are Gaussian complex circular independent noises with spatially white covariance matrix $\mathbf{C}_{\mathbf{n}} = \sigma_{\mathbf{n}}^2\mathbf{I}_N$ with unknown noise power $\sigma_{\mathbf{n}}^2$, independent from the M sources. Then the reduced log likelihood function $L\left(\mathbf{\Theta}; \mathbf{\Theta}^0\right)$

is given by [24]:

$$L\left(\mathbf{\Theta};\mathbf{\Theta}^0\right) = \sum_{t=1}^{T} \frac{\left\|\mathbf{\Pi}_{\mathbf{A}(\mathbf{\Theta})}\mathbf{x}_t\left(\mathbf{\Theta}^0\right)\right\|^2}{TM} \tag{21}$$

and:

$$\mathcal{P}\left(L\left(\mathbf{\Theta};\mathbf{\Theta}^0\right) \leq \lambda\right) = \mathcal{P}\left(\mathcal{CX}_{MT}^2 \leq \lambda; F\left(\mathbf{\Theta};\mathbf{\Theta}^0\right), \frac{\sigma_{\mathbf{n}}^2}{TM}\right) \tag{22}$$

where $F\left(\mathbf{\Theta};\mathbf{\Theta}^0\right) = \sum_{t=1}^{T}\frac{\left\|\mathbf{\Pi}_{\mathbf{A}(\mathbf{\Theta})}\mathbf{A}(\mathbf{\Theta}^0)\mathbf{s}_t\right\|^2}{TM}$ is a generalized correlation function (aka generalized matched filter) and $\mathcal{P}\left(\mathcal{CX}_K^2 \leq \lambda; \delta, \sigma^2\right)$ denotes the c.d.f. of a non-central complex (circular) chi-square with corresponding degrees of freedom in K and positive noncentrality parameter in δ. Let $\widehat{\mathbf{\Theta}} \triangleq \widehat{\mathbf{\Theta}^0}\left(\mathbf{x}\right)$ denote the CMLE of $\mathbf{\Theta}^0$:

$$\widehat{\mathbf{\Theta}} = \arg\max_{\mathbf{\Theta}}\left\{L\left(\mathbf{\Theta};\mathbf{\Theta}^0\right)\right\} \tag{23}$$

then:

$$\mathbf{\Theta} \in image\left(\widehat{\mathbf{\Theta}}\right) \Rightarrow \exists\{\mathbf{x}_t\}_{t=1}^{T} \mid L\left(\mathbf{\Theta};\mathbf{\Theta}^0\right) > L\left(\mathbf{\Theta}^0;\mathbf{\Theta}^0\right)$$

or conversely:

$$L\left(\mathbf{\Theta};\mathbf{\Theta}^0\right) < L\left(\mathbf{\Theta}^0;\mathbf{\Theta}^0\right), \forall\{\mathbf{x}_t\}_{t=1}^{T} \Rightarrow \mathbf{\Theta} \notin image\left(\widehat{\mathbf{\Theta}}\right)$$

leading to:

$$\mathcal{P}\left(\mathbf{\Theta} \notin image\left(\widehat{\mathbf{\Theta}}\right)\right) \geq \mathcal{P}\left(L\left(\mathbf{\Theta};\mathbf{\Theta}^0\right) < L\left(\mathbf{\Theta}^0;\mathbf{\Theta}^0\right)\right). \tag{24}$$

Then, some additional probability calculus detailed in Appendix VI-A allows to prove that, $\forall\alpha \in \left]0,1\right[$:

$$\mathcal{P}\left(\mathbf{\Theta} \notin image\left(\widehat{\mathbf{\Theta}}\right)\right) \geq 1 - \left(\mathcal{P}\left(\frac{L\left(\mathbf{\Theta}^0;\mathbf{\Theta}^0\right)}{F\left(\mathbf{\Theta}^0;\mathbf{\Theta}^0\right)} < \alpha\right) + \mathcal{P}\left(\frac{L\left(\mathbf{\Theta};\mathbf{\Theta}^0\right)}{F\left(\mathbf{\Theta}^0;\mathbf{\Theta}^0\right)} > \alpha\right)\right) \tag{25}$$

or conversely:

$$\mathcal{P}\left(\mathbf{\Theta} \in image\left(\widehat{\mathbf{\Theta}}\right)\right) \leq \mathcal{P}\left(\frac{L\left(\mathbf{\Theta}^0;\mathbf{\Theta}^0\right)}{F\left(\mathbf{\Theta}^0;\mathbf{\Theta}^0\right)} < \alpha\right) + \mathcal{P}\left(\frac{L\left(\mathbf{\Theta};\mathbf{\Theta}^0\right)}{F\left(\mathbf{\Theta}^0;\mathbf{\Theta}^0\right)} > \alpha\right) \tag{26}$$

Let $\Upsilon_{\mathbf{\Theta}^0}\left(\alpha\right)$ denote the main lobe at α $(0 < \alpha < 1)$:

$$\Upsilon_{\mathbf{\Theta}^0}\left(\alpha\right) = \left\{\mathbf{\Theta} : \frac{F\left(\mathbf{\Theta};\mathbf{\Theta}^0\right)}{F\left(\mathbf{\Theta}^0;\mathbf{\Theta}^0\right)} > \alpha\right\} \tag{27}$$

Then they are many ways of exploiting (25)(26) according to:

• the choice of the main lobe $\Upsilon_{\mathbf{\Theta}^0}\left(\alpha\right)$ including $image\left(\widehat{\mathbf{\Theta}}\right)$ requested to enforce the asymptotic Gaussianity and efficiency of CMLEs,

• the acceptable probability of an outlier, in the sense of being outside $\Upsilon_{\mathbf{\Theta}^0}\left(\alpha\right)$: $\mathcal{P}\left(\mathbf{\Theta} \in image\left(\widehat{\mathbf{\Theta}}\right) \mid \mathbf{\Theta} \notin \Upsilon_{\mathbf{\Theta}^0}\left(\alpha\right)\right)$. We opt for a simplified didactic approach where we consider $\Upsilon_{\mathbf{\Theta}^0}\left(\frac{1}{2}\right)$ (i.e. the usual main lobe at $-3dB$)
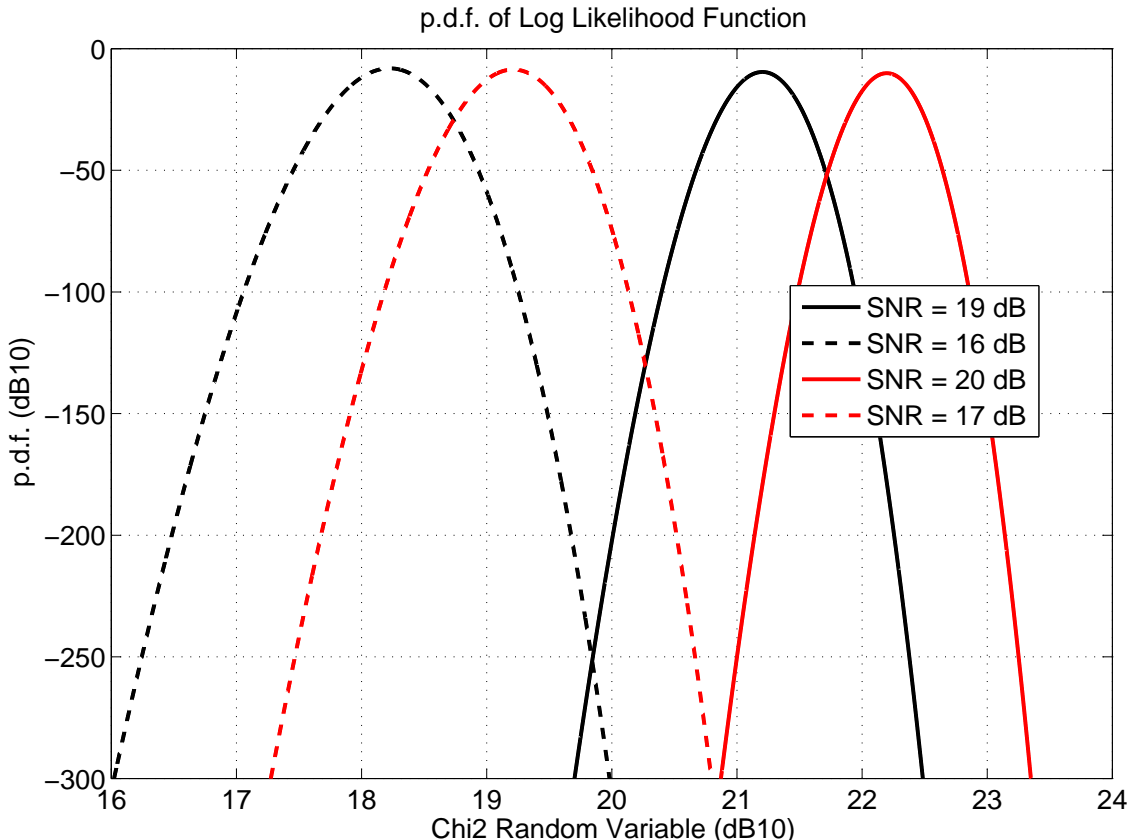
Fig. 2. Loglikelihood p.d.f. (solid lines for $L\left(\Theta^0; \Theta^0\right)$ and dash lines for $L\left(\Theta_{3dB}; \Theta^0\right)$) for two correlated signal sources impinging on a ULA of $N = 10$ sensors, separated by $\delta\theta = \frac{3}{10}\theta_{3dB}$ and $T = 10$ snapshots.

and $\mathcal{P}\left(\Theta \in image\left(\widehat{\Theta}\right) \mid \Theta \notin \Upsilon_{\Theta^0}\left(\frac{1}{2}\right)\right) \simeq 0.$

The quasi-nullity of the probability of an outlier can be demonstrated by computing the p.d.f. of $L\left(\Theta^0; \Theta^0\right)$ and $L\left(\Theta_{3dB}; \Theta^0\right)$ where $F\left(\Theta_{3dB}; \Theta^0\right) = \frac{1}{2}$ and by checking that their supports do not overlap above a certain p.d.f. threshold value, as small as possible (10^{-30} in the present paper). The underlying hypothesis is to consider that below this threshold value, any p.d.f. can be rounded to zero. As the p.d.f. of $\mathcal{CX}^2_{MT}\left(F\left(\Theta; \Theta^0\right), \frac{\sigma^2}{TM}\right)$ is an increasing function in $F\left(\Theta; \Theta^0\right)$, it is sufficient to check that the p.d.f. of $L\left(\Theta^0; \Theta^0\right)$ and $L\left(\Theta_{3dB}; \Theta^0\right)$ do not overlap to ensure that this property is valid for any $\Theta \notin \Upsilon_{\Theta^0}\left(\frac{1}{2}\right)$, what proves that (under the approximations above mentioned) $image\left(\widehat{\Theta}\right) \subset \Upsilon_{\Theta^0}\left(\frac{1}{2}\right)$. We illustrate the soundness of the proposed approach with the application example given in [19] where two ($M = 2$) fully correlated sources separated by $\frac{3}{10}$ of the beamwidth impinge on a uniform linear array (ULA) of $N = 10$ sensors with half-wavelength spacing. The high SNR Gaussianity of the CMLE

was checked via a Lilliefors goodness-of-fit test [25] and was observed for a SNR at output of the single source matched filter higher than approximately $19dB$ (that is a SNR of $9dB$ at sensor level [19]). Figure 2 displaying the p.d.f. of the log likelihood functions $L\left(\mathbf{\Theta}^0;\mathbf{\Theta}^0\right)$ (solid lines) and $L\left(\mathbf{\Theta}_{3dB};\mathbf{\Theta}^0\right)$ (dash lines), for $SNR \in \{19, 20\}\, dB$, shows that Gaussianity is checked for a SNR of $20dB$ according to the proposed approach (and p.d.f. threshold value) which is in accordance with the Lilliefors goodness-of-fit test.

B. Asymptotic performance of conditional model

Then, in the asymptotic region of operation:

$$\widehat{\mathbf{\Theta}} \sim \mathcal{N}\left(\mathbf{\Theta}^0, \mathbf{CRB}\left(\mathbf{\Theta}^0\right)\right) \tag{28}$$

where, for the general linear observation model (20) [24][26][3]:

$$\mathbf{CRB}\left(\mathbf{\Theta}\right) = \frac{\sigma_{\mathbf{n}}^2}{2T} \operatorname{Re}\left\{\mathbf{H}\left(\mathbf{\Theta}\right) \odot \left(\widehat{\mathbf{R}}_{\mathbf{s}}^T \otimes \mathbf{1}_{P\times P}\right)\right\}^{-1}$$

$$\mathbf{H}\left(\mathbf{\Theta}\right) = \begin{bmatrix} \mathbf{H}\left(\mathbf{\Theta}\right)_{1,1} & \ldots & \mathbf{H}\left(\mathbf{\Theta}\right)_{1,P} \\ \vdots & \ddots & \vdots \\ \mathbf{H}\left(\mathbf{\Theta}\right)_{P,1} & \ldots & \mathbf{H}\left(\mathbf{\Theta}\right)_{P,P} \end{bmatrix}, \quad \mathbf{H}\left(\mathbf{\Theta}\right)_{p,q} = \frac{\partial\mathbf{A}(\mathbf{\Theta})}{\partial\theta_p}^H \mathbf{\Pi}_{\mathbf{A}(\mathbf{\Theta})}^\perp \frac{\partial\mathbf{A}(\mathbf{\Theta})}{\partial\theta_q}, \quad \widehat{\mathbf{R}}_{\mathbf{s}} = \sum_{t=1}^T \frac{\mathbf{s}_t\mathbf{s}_t^H}{T} \tag{29}$$

and $\mathbf{1}_{P\times P}$ is a $P \times P$ matrix of ones. Additionally it has been proved that for each source [26][30]: the highest (worst) variance is obtained when the sources amplitudes are fully correlated and the lowest (best) variance is obtained when the sources amplitudes are uncorrelated.

IV. STATISTICAL RESOLUTION LIMIT FOR TWO CLOSELY SPACED FREQUENCIES

A reference application example is the multiple tones estimation problem where in (20):

$$\mathbf{a}\left(\theta\right) = \left[1, \ldots, e^{j2\pi n\theta}, \ldots, e^{j2\pi(N-1)\theta}\right]^T \tag{30}$$

Although the proposed resolvability criterion (11) is not limited to a given number of signal sources (tones herein) (see [13] for additional examples), we focus in the present paper on the resolvability of two closely spaced frequencies.

---

[3]More general CRB expressions can be found in [27][28][29].

A. Background on statistical resolution limit (SRL) criteria

The resolvability of closely spaced signals, in terms of their parameters of interest, for a given scenario (e.g., for a given number of snapshots, for a given SNR and/or for a given number of sensors) is an old and challenging problem which was recently updated by Smith [4], Shahram and Milanfar [5], Liu and Nehorai [1], Amar and Weiss [2] and El Korso et al [31]. Historically, the statistical resolution limit (SRL) has been empirically defined as the minimum distance between the parameters of interest that allows '*distinguishing*' two closely spaced noisy signals. The concept of SRL does not necessarily take into account how well parameters are estimated, only whether multiple signals can be distinguished. Several criteria were introduced to describe and derive the SRL. We can sum up most of them in four different categories.

(1) The first one, and most probably the oldest one, is an algorithm dependent criterion based on the cost function of the considered parametric scheme. It was mainly introduced and applied using the concept of mean null spectrum. Let us considered the problem of distinguishing two signals parameterized by their frequencies, namely f_1 and f_2. In 1974, Cox states that these two signals are resolved, w.r.t. a given parametric algorithm, *if the mean of the null spectrum at the frequencies f_1 and f_2 is lower than the mean of the null spectrum at the frequency midpoint $\frac{f_1+f_2}{2}$* [3]. Two decades after, Sharman and Durrani proposed another commonly used criterion in [6] 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*. These intuitive criteria were successfully applied to derive the SRL for a specific algorithm, as MUSIC and Min-Norm scheme, in many different situations (the reader can refer to [3], [6], [32], [33], [34], [35]).

(2) The second approach is based on detection theory. This approach reformulates the problem of distinguishing two signals as a binary hypothesis test. More precisely, the feature of interest of one signal is tested against the feature of interest of the other signal. Let us consider the spectral analysis problem where the main frequency is considered as the feature of interest. In this context, Sharman and Milanfar [5] proposed to use a binary hypothesis test to define/derive the SRL. The hypothesis $\mathcal{H}_1$ represents the case in which two signals are resolvable/distinguishable, i.e., $f_1 - f_2 \neq 0$, whereas, the hypothesis $\mathcal{H}_0$ represents the case where the two signals coalise/combine into one signal, i.e., $f_1 - f_2 = 0$. To solve such a hypothesis test, many strategies can be considered. The commonly used one is the generalized likelihood ratio test (GLRT). This choice is mainly motivated by the known fact that, under certain conditions, the GLRT is asymptotically uniformly most powerful test among all the invariant statistical tests [36], which is one of the strongest statements of optimality that one could expect to obtain [37]. Using this

strategy, the main difficulty is to derive the SRL (i.e., the separation $\delta = f_1 - f_2$ which resolves the binary hypothesis test) as a function of the probability of false alarm $\mathcal{P}_{\text{fa}}$ and the probability of detection $\mathcal{P}_{\text{d}}$. This approach was recently applied to several parametric problems (the reader can refer to [1], [5], [31], [38] and references in [31] for more details). Finally note that, in this context, another strategy was applied by Amar and Weiss in [2]. The authors have proposed to determine the SRL of complex sinusoids with nearby frequencies using the Bayesian approach for a given correct decision probability instead of considering $\mathcal{P}_{\text{fa}}$ and $\mathcal{P}_{\text{d}}$ which are relevant generally in a non Bayesian detection strategy.

(3) The third approach is based on the estimation accuracy. Unlike the first approach and like the second one, this approach is independent from the estimation algorithm and uses mainly tools as lower bounds (LBs) on the mean square error (MSE). LBs on MSE characterize the ultimate performance of an unbiased estimator in terms of MSE [39], [40]. Due to its simplicity and tightness under certain conditions (see Section III) the CRB is the LB mainly used practically to assess the ultimate estimation accuracy in terms of MSE. In this way, Lee [41], [42], Yau and Bressler [26], and Smith [4] proposed different criteria to describe the SRL. In [26, Section VIII][41], authors state that *two signals are said to be resolvable w.r.t. the frequencies if the maximum standard deviation of each frequency estimate is less than half the difference between f_1 and f_2.* Under conditions for which the CRB is a tight bound [10], the standard deviations σ_{f_1} and σ_{f_2}, of $\hat{f}_1$ and $\hat{f}_2$ using MLEs, can be well approximated by $\sqrt{\text{CRB}(f_1)}$ and $\sqrt{\text{CRB}(f_2)}$, respectively. Leading to the following separation '*limit*', δ, which represents the SRL: $\delta = 2\max\left\{\sqrt{\text{CRB}(f_1)}, \sqrt{\text{CRB}(f_2)}\right\}$. The second strategy within this estimation accuracy approach was formulated by Smith in [4] (also alluded to in [26, Section VIII.A]) and states that *two signals are resolvable w.r.t. their frequencies if the difference between the frequencies, is greater than the standard deviation of the frequencies difference estimation.* Considering that the mild conditions are met for the CRB, the separation '*limit*' δ representing the SRL in the Smith sense, is then given as the solution of the following equation: $\delta^2 = \text{CRB}(\delta)$ where $\delta = f_1 - f_2$. Which can be written as $\delta^2 = \text{CRB}(f_1) + \text{CRB}(f_2) - 2\text{CRB}(f_1, f_2)$.

(4) The fourth approach is based on the definition of a probability of resolvability [11][12] of parameter vectors by estimators as detailed in Section I.

B. Comparison of SRL predictions

The aim of this section is to compare the SRL provided by (11) with the existing ones given by Lee-Yaw-Bressler [26][41], Smith [4], Shahram-Milanfar-Liu-Nehorai [1][5][43] and Amar-Weiss [2]. The underlying idea is to measure the cost, in terms of scenario requirements (SNR, number of snapshots,

number of sensors), of the proposed SRL definition which takes into account not only the ability to distinguish multiple signals (as all definitions do) but also how well parameters are estimated. To get an idea, let us consider the problem of predicting the SNR required (according to a given criterion) to achieve a given resolution limit δ, $\delta = \theta_2 - \theta_1$, $\theta_1 < \theta_2$. As a setting, let:

• $N = 32$, $\delta = \frac{1}{N}\frac{1}{k}$ where $k \in \{2, 4, 8, 16, 32\}$,

• the tones amplitude be fully correlated (equal amplitudes but with a phase shift different from zero [43]) $\mathbf{s}_t^T = \sqrt{\frac{SNR}{N}}\left(1, e^{j\frac{\pi}{8}}\right)$ where the SNR is computed at output of the single source matched filter (note that it may not be the worst correlation case [26]).

As an instance of the resolvability criterion (11), we consider the combination of symmetric and isotropic required estimation precisions $\mathbf{\Xi}_{\text{req}}^+ = \mathbf{\Xi}_{\text{req}}^- = \mathbf{\Xi}_{\text{req}} = \left[\frac{\delta}{2}, \frac{\delta}{2}\right]^T$ defining disjoint intervals (1-dimensional hyper-rectangles) and $\mathcal{P}_{\text{res}} \in \{0.68, 0.95\}$. For two Gaussian unbiased estimators, the bounds of the confidence interval on the probability of resolvability $\mathcal{P}_{\text{res}}$ (12) can be computed for $\mathbf{s}^T = (2, 2)$ (17) by numerical integration of the p.d.f. of a quadratic form for central normal variates (18) which has then the following closed-form expression (derived in Appendix VI-C for sake of completeness):

$$p_{Q_2}\left(v; 0, \boldsymbol{\sigma}^2\right) = \frac{e^{-\frac{v}{2}\frac{1}{2}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)}}{2\sigma_1\sigma_2} I_0\left(\frac{v}{2}\frac{1}{2}\left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}\right)\right)$$

Then the main features of each high resolution scenario and of the proposed probabilistic approach associated to a given value of δ can be described with the help of 3 figures (exemplified for $k = 4$):

• figure (3): the output of the single source matched filter
$\left(\frac{1}{N}\left\|\mathbf{a}\left(\boldsymbol{\theta}\right)^H \mathbf{x}_t\left(\boldsymbol{\theta}^0\right)\right\|^2\right)$ which could be the first step in a practical implementation of the CMLE (Clean algorithm, Alternating Projection algorithm) [24].

• figure (4): the probability (lower and upper) bounds (PLB and PUB) defined by (12) for $\mathbf{s}^T = (2, 2)$ as a function of the SNR. These bounds allow to determine the SNR interval containing the SNR_{res} for which $\mathcal{P}_{\text{res}}$ is obtained since: $SNR\left(PUB = \mathcal{P}_{\text{res}}\right) \leq SNR_{\text{res}} \leq SNR\left(PLB = \mathcal{P}_{\text{res}}\right)$.

• figure (5): the p.d.f. of $L\left(\boldsymbol{\Theta}^0; \boldsymbol{\Theta}^0\right)$ and $L\left(\boldsymbol{\Theta}_{3dB}; \boldsymbol{\Theta}^0\right)$ (22) for $SNR\left(PUB = \mathcal{P}_{\text{res}}\right)$ and $SNR\left(PLB = \mathcal{P}_{\text{res}}\right)$ to prove that within $[SNR\left(PUB = \mathcal{P}_{\text{res}}\right), SNR\left(PLB = \mathcal{P}_{\text{res}}\right)]$ the condition of asymptotic region of operation for CMLE is valid, that is $\mathcal{P}\left(\boldsymbol{\Theta} \in image\left(\widehat{\boldsymbol{\Theta}}\right) \mid \boldsymbol{\Theta} \notin \Upsilon_{\boldsymbol{\Theta}^0}\left(\frac{1}{2}\right)\right) \simeq 0$.

The comparison of the SNR required to achieve a given resolution limit δ is displayed on figure (6), where LYB-SNR is the one obtained with the Lee-Yaw-Bressler criterion (LYB-criterion) [26][41], S-SNR is the one obtained with the Smith criterion (S-criterion) [4], SMLN-SNR is the one obtained with the GLRT approach (SMLN-criterion) where $\mathcal{P}_d = 0.95$ and $\mathcal{P}_{fa} = 0.1$ [1][5][43], and AW-SNR is the
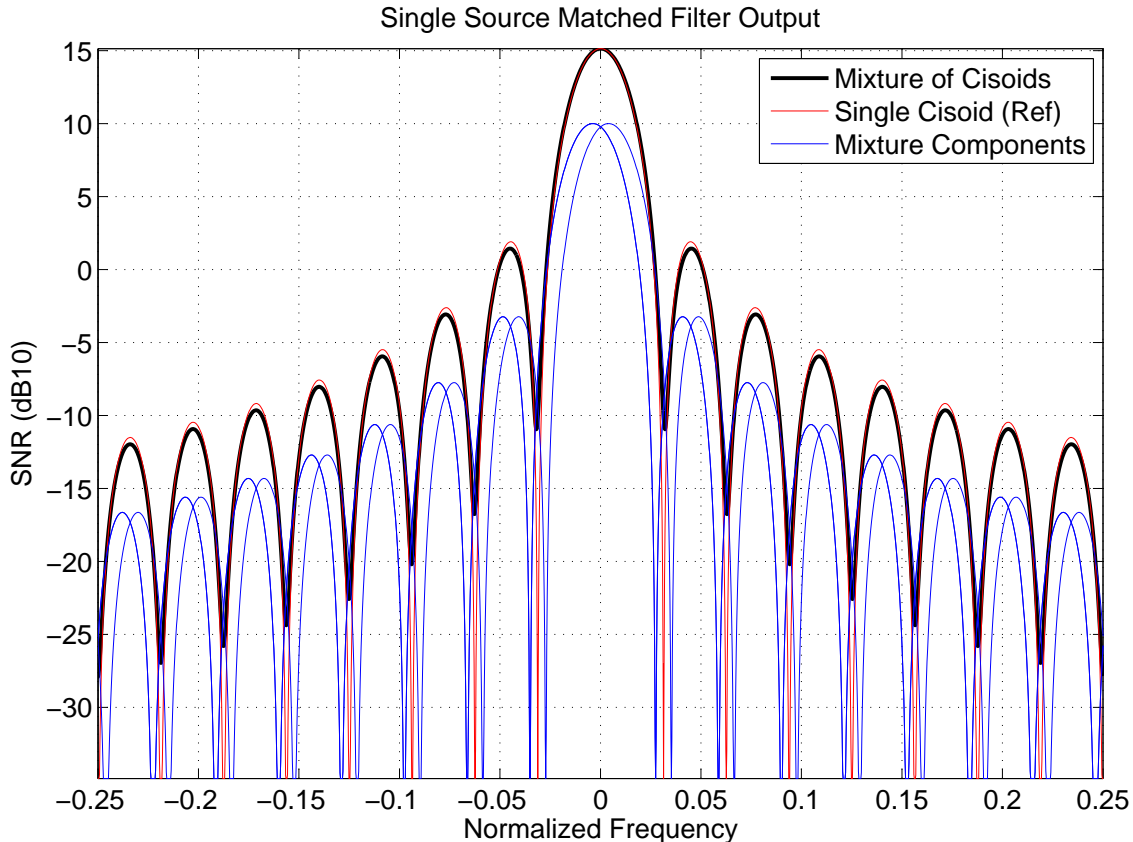
Fig. 3. Single source matched filter output: 2 complex exponentials separated by $\delta = \frac{1}{4N}$, $N = 32$

one obtained with the Bayesian hypothesis test approach (AW-criterion) [2] where $\mathcal{P}d = 0.95$. This figure raises the following comments:

• there is no bound on SNR_{res} for the proposed approach when $k = 2$. The reason is simple: the condition of asymptotic region of operation for CMLEs is not valid within $[SNR(PUB = \mathcal{P}_{\text{res}}), SNR(PLB = \mathcal{P}_{\text{res}})]$, $\mathcal{P}_{\text{res}} \in \{0.68, 0.95\}$, when $k = 2$. Therefore the probability (lower and upper) bounds defined by (12)(17) are not trustworthy and should not be taken into account.

• in the setting considered, the LYB-criterion takes into account the same estimation errors $\Xi_{\text{req}} = \frac{\delta}{2} \mathbf{1}_M$, except that the LYB-criterion expresses requirements on marginal probability of errors, that is $\mathcal{P}\left(\left| \widehat{\theta}_i - \theta_i \right| \le \frac{\delta}{2} \right) = 0.68$, $i \in \{1, 2\}$, where $\widehat{\theta}_i$ is the MLE of θ_i in the asymptotic region of operation. Since $\mathcal{P}(A \cap B) \le \min \{\mathcal{P}(A), \mathcal{P}(B)\}$, these requirements are less demanding than the requirement on the joint probability of error with same occurrence probability as formulated by (12)(17) when $\mathcal{P}_{\text{res}} = 0.68$; hence the smaller (or equal) required LYB-SNR to achieve a given resolution limit δ. The difference
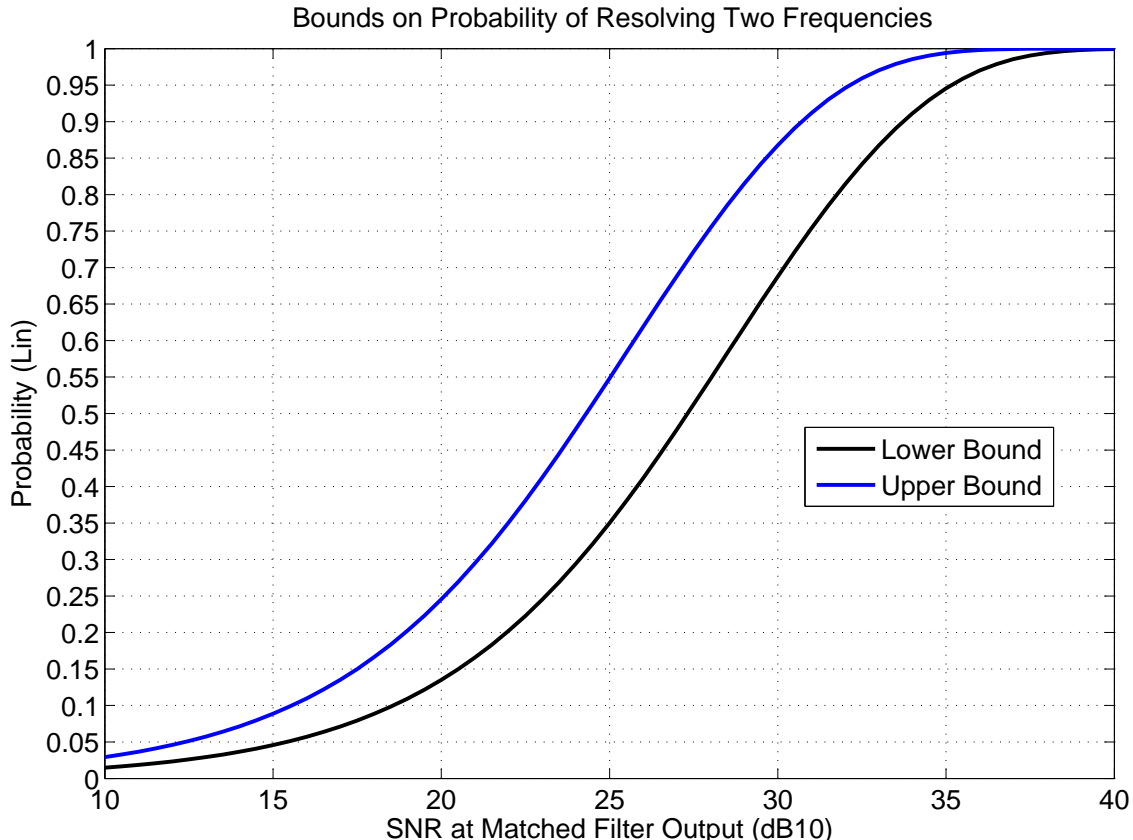
Fig. 4. Bounds (12) on the probability of resolving two complex exponentials separated by $\delta = \frac{1}{4N}$, $N = 32$, $\mathbf{\Xi}_{\text{req}}^T = \left[\frac{\delta}{2}, \frac{\delta}{2}\right]$, $\mathbf{s}^T = (2,2)$, $\mathcal{P}_{\text{res}} = 0.95$.

between SNR_{res} and LYB-SNR increases as $\mathcal{P}_{\text{res}}$ increases above 0.68, as shown when $\mathcal{P}_{\text{res}} = 0.95$.

• as already noticed in [26, Section VIII], the LYB-criterion implies the S-criterion which finally states that the SNR required (S-SNR) is the one for which $\mathcal{P}\left(\left|\left(\widehat{\theta}_2 - \widehat{\theta}_1\right) - \delta\right| \leq \delta\right) = 0.68$, or equivalently the one for which $\mathcal{P}\left(\widehat{\theta}_2 \leq \widehat{\theta}_1\right) = 0.16$ and $\mathcal{P}\left(\widehat{\theta}_2 \geq \left(\widehat{\theta}_1 + \delta\right) + \delta\right) = 0.16$. The second form of the S-criterion clearly highlights that this criterion is mainly focused on parameters distinguishability as it only requires a loose estimation precision for each parameter. Hence the lower S-SNR values in comparison with LYB-SNR and $[SNR\left(PUB = 0.68\right), SNR\left(PLB = 0.68\right)]$.

• in order to obtain tractable solutions, the AW-SNR and the SMLN-SNR reported here are derived from [2] and [43] in the case of a known central frequency and a known signals power or amplitude. After

some proper modifications, one obtains:

$$\text{AW-SNR} = \frac{0.675\pi N}{\delta^2(N-1)(N+1)} \tag{31}$$

$$\text{SMLN-SNR} = \lambda\frac{2}{\delta^2\varrho^2} \tag{32}$$

where $\varrho = \|s_{t,2} - s_{t,1}\|\,\|\dot{a}\|$, $\dot{a}$ denotes the derivative of a w.r.t. θ and λ is numerically computed as the solution of $\mathcal{Q}_{\chi_1^2}^{-1}(\mathcal{P}_{fa};0,1) = \mathcal{Q}_{\chi_1^2}^{-1}(\mathcal{P}_d;0,\lambda)$ for a given $\mathcal{P}_{fa}$ and $\mathcal{P}_d$, where $\mathcal{Q}_{\chi_1^2}(v;0,\sigma^2)$ denotes the right tail of the distribution of a central χ_1^2 with mean σ^2.

• Let us remind that both the SMLN-criterion and the AW-criterion exploit the historical definition of resolvability, that is the problem of distinguishing whether the observation consists of a single signal or of two signals, irrespective of any requirement on the precision of parameters estimation. Thus on a scale representing the ability of a criterion to incorporate a requirement on the precision of parameters estimation, we may say that the SMLN-criterion and the AW-criterion are based on an approach of resolvability at the opposite of our proposed approach (and the one of LYB), and that the S-criterion defines an intermediate approach.

The significant differences in the SNR requested to fulfill the different SRL criteria are representative of the position of each approach on this scale: indeed each approach allows to discern the presence of two signals but the estimates of the signal parameters may not be sufficiently accurate according to a prescribed precision (hyper-rectangles). Hence the question on which information one looks for: distinguishability only or distinguishability and accuracy, since in order to get both there is a price to pay (possibly overstressed in our example[4]).

V. CONCLUSION

We have proposed a resolvability definition which takes into account not only the ability to distinguish multiple signals (as all definitions do) but also how well parameters are estimated. In order to widen the practical use of this definition, we have also introduced a general scheme to generate the lower and upper bounds of confidence intervals on the probability of resolvability for a given required estimation precision. The application of this approach to the conditional model has shown that significant differences in the region of operation (SNR, number of snapshots, number of sensors) requested to fulfill the

---

[4]The hypotheses of a known central frequency and a known signals power or amplitude in (31,32) introduce an *a priori* information that has not been taken into account in the conditional CRB considered (29), therefore higher than the one incorporating this *a priori* information.
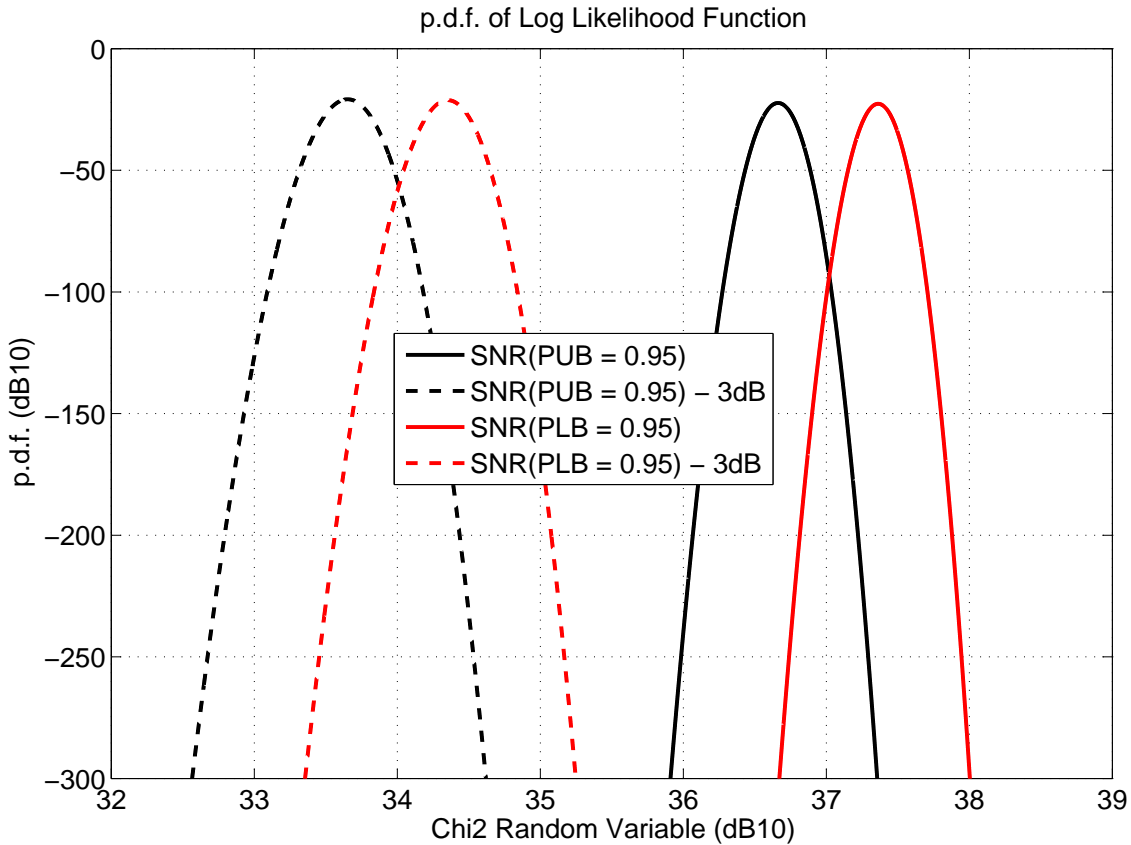
Fig. 5. Loglikelihood p.d.f. at limits of SNR interval allowing the resolvability of two complex exponentials separated by $\delta = \frac{1}{4N}$, $N = 32$, $\mathcal{P}_{\text{res}} = 0.95$

different SRL criteria are expected according to which information is looked for: distinguishability only or distinguishability and accuracy. Possible application of the proposed approach to the unconditional model will be pursued in future work, since asymptotic efficiency of MLEs does no longer hold for small number of observations [23].

VI. APPENDIX

A. Approximation of a hyper-rectangle

As:

$$\left| y_q - y_q^0 - \frac{d\varepsilon_q}{2} \right| < \overline{\varepsilon_q} \Leftrightarrow \left| \frac{y_q - y_q^0 - \frac{d\varepsilon_q}{2}}{\overline{\varepsilon_q}} \right| < 1$$
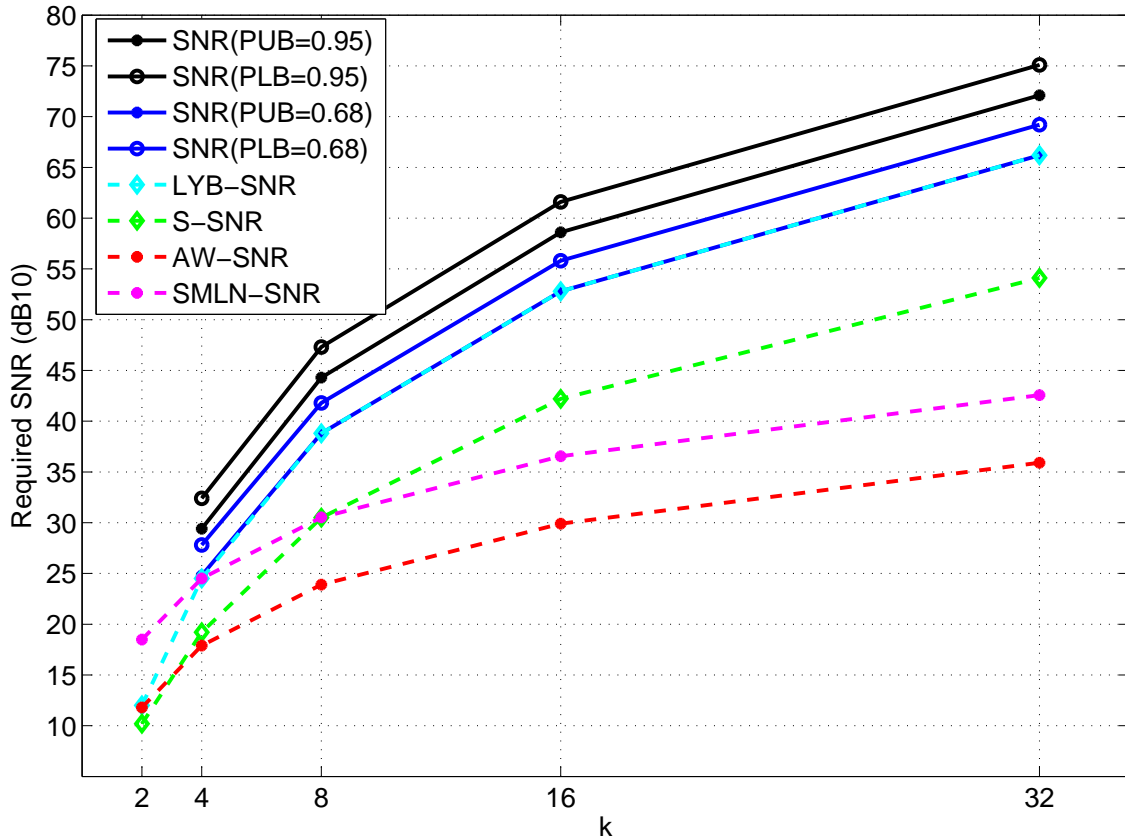
Fig. 6. Required SNR to achieve a given SRL for two complex exponentials separated by $\delta = \frac{1}{kN}$, $N = 32$, $k \in \{2, 4, 8, 16, 32\}$

then, $\forall s_q \geq 1$:

$$\left| y_q - y_q^0 - \frac{d\varepsilon_q}{2} \right| < \overline{\varepsilon_q} \Rightarrow \left| \frac{y_q - y_q^0 - \frac{d\varepsilon_q}{2}}{\overline{\varepsilon_q}} \right|^{s_q} < 1$$

and:

$$\bigwedge_{q=1}^{Q} \left(\left| y_q - y_q^0 - \frac{d\varepsilon_q}{2} \right| < \overline{\varepsilon_q} \right) \Rightarrow \sum_{q=1}^{Q} \left| \frac{y_q - y_q^0 - \frac{d\varepsilon_q}{2}}{\overline{\varepsilon_q}} \right|^{s_q} < Q$$

that is $R\left(\mathbf{y}^0, \mathbf{\Xi}^-, \mathbf{\Xi}^+\right) \subset \Sigma_{\mathbf{s}}^{Q}\left(\mathbf{y}^0, \mathbf{\Xi}^-, \mathbf{\Xi}^+\right)$. Conversely:

$$\sum_{q=1}^{Q} \left| \frac{y_q - y_q^0 - \frac{d\varepsilon_q}{2}}{\overline{\varepsilon_q}} \right|^{s_q} < 1 \Rightarrow \bigwedge_{q=1}^{Q} \left(\left| \frac{y_q - y_q^0 - \frac{d\varepsilon_q}{2}}{\overline{\varepsilon_q}} \right|^{s_q} < 1 \right).$$

As, $\forall s_q \geq 1$:

$$\left| \frac{y_q - y_q^0 - \frac{d\varepsilon_q}{2}}{\overline{\varepsilon_q}} \right|^{s_q} < 1 \Leftrightarrow \left| \frac{y_q - y_q^0 - \frac{d\varepsilon_q}{2}}{\overline{\varepsilon_q}} \right| < 1$$

therefore:

$$\sum_{q=1}^{Q} \left| \frac{y_q - y_q^0 - \frac{d\varepsilon_q}{2}}{\overline{\varepsilon}_q} \right|^{s_q} < 1 \Rightarrow \bigwedge_{q=1}^{Q} \left(\left| y_q - y_q^0 - \frac{d\varepsilon_q}{2} \right| < \overline{\varepsilon}_q \right)$$

that is $\Sigma_{\mathbf{s}}^1 \left(\mathbf{y}^0, \mathbf{\Xi}^-, \mathbf{\Xi}^+ \right) \subset R \left(\mathbf{y}^0, \mathbf{\Xi}^-, \mathbf{\Xi}^+ \right)$. QED.

B. Lower bound on $\mathcal{P} \left(L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) < L \left(\mathbf{\Theta}^0; \mathbf{\Theta}^0 \right) \right)$

$\forall \beta > 0$:

$$
\begin{aligned}
\mathcal{P} \left(L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) > L \left(\mathbf{\Theta}^0; \mathbf{\Theta}^0 \right) \right) &= \mathcal{P} \left(\left(L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) > L \left(\mathbf{\Theta}^0; \mathbf{\Theta}^0 \right) \right) \wedge \left(\left(\beta \geq L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) \right) \vee \left(L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) > \beta \right) \right) \right) \\
&= \mathcal{P} \left(\left(\beta \geq L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) \right) \wedge \left(L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) > L \left(\mathbf{\Theta}^0; \mathbf{\Theta}^0 \right) \right) \right) + \\
&\quad \mathcal{P} \left(\left(L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) > L \left(\mathbf{\Theta}^0; \mathbf{\Theta}^0 \right) \right) \wedge \left(L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) > \beta \right) \right) \\
&= \mathcal{P} \left(\beta \geq L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) > L \left(\mathbf{\Theta}^0; \mathbf{\Theta}^0 \right) \right) + \\
&\quad \mathcal{P} \left(\left(L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) > L \left(\mathbf{\Theta}^0; \mathbf{\Theta}^0 \right) \right) \wedge \left(L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) > \beta \right) \right)
\end{aligned}
$$

Therefore:

$$\mathcal{P} \left(L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) > L \left(\mathbf{\Theta}^0; \mathbf{\Theta}^0 \right) \right) \leq \mathcal{P} \left(L \left(\mathbf{\Theta}^0; \mathbf{\Theta}^0 \right) < \beta \right) + \mathcal{P} \left(L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) > \beta \right)$$

what yields (25):

$$\mathcal{P} \left(L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) < L \left(\mathbf{\Theta}^0; \mathbf{\Theta}^0 \right) \right) \geq 1 - \left(\mathcal{P} \left(L \left(\mathbf{\Theta}^0; \mathbf{\Theta}^0 \right) < \beta \right) + \mathcal{P} \left(L \left(\mathbf{\Theta}; \mathbf{\Theta}^0 \right) > \beta \right) \right)$$

C. Derivation of $p_{Q_2} \left(v; 0, \left(\sigma_1^2, \sigma_2^2 \right)^T \right)$

Since $Q_2 = \sigma_1^2 z_1^2 + \sigma_2^2 z_2^2$ where $\mathbf{z} = (z_1, z_2)^T \sim \mathcal{N} \left(\mathbf{0}, \mathbf{I} \right)$, then:

$$p_{Q_2} \left(v; 0, \boldsymbol{\sigma}^2 \right) = \int_0^v p_{\chi_1^2}(v - x; 0, \sigma_1^2) p_{\chi_1^2}(x; 0, \sigma_2^2) dx = \frac{1}{2\pi \sigma_1 \sigma_2} e^{-\frac{v}{2\sigma_1^2}} \int_0^v \frac{1}{\sqrt{v - x} \sqrt{x}} e^{-ax} dx$$

where $a = \frac{1}{2} \left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2} \right)$. After a first change of variable $x = \frac{v}{2} - u$ yielding:

$$\int_0^v \frac{e^{-ax}}{\sqrt{(v - x) x}} dx = \frac{2}{v} e^{-\frac{a}{2} v} \int_{-\frac{v}{2}}^{\frac{v}{2}} \frac{e^{au}}{\sqrt{1 - \left(2\frac{u}{v} \right)^2}} du,$$

a second change of variable $\sin \theta = 2\frac{u}{v}$ allows to show that:

$$\int_0^v \frac{1}{\sqrt{(v - x) x}} e^{-ax} dx = e^{-\frac{a}{2} v} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} e^{\frac{av}{2} \sin \theta} d\theta.$$

Therefore:

$$p_{Q_2}\left(v;0,\boldsymbol{\sigma}^2\right) = \frac{e^{-\frac{v}{2\sigma_1^2}}e^{-\frac{va}{2}}}{2\sigma_1\sigma_2}\left(\frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}e^{\frac{av}{2}\sin\theta}d\theta\right)$$

Additionally as:

$$\frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}e^{z\sin\theta}d\theta = \frac{1}{\pi}\int_0^{\pi}e^{-z\cos(u)}du = I_0\left(-z\right) = I_0\left(z\right)$$

one finally has:

$$p_{Q_2}\left(v;0,\boldsymbol{\sigma}^2\right) = \frac{e^{-\frac{v}{2}\frac{1}{2}\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)}}{2\sigma_1\sigma_2}I_0\left(\frac{v}{2}\frac{1}{2}\left(\frac{1}{\sigma_2^2}-\frac{1}{\sigma_1^2}\right)\right)$$

REFERENCES

[1] Z. Liu and A. Nehorai, "Statistical angular resolution limit for point sources", IEEE Trans. on SP, 55(11): 5521-5527, 2007

[2] A. Amar and A. Weiss, "Fundamental limitations on the resolution of deterministic signals", IEEE Trans. on SP, 56(11): 5309-5318, 2008

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

[4] S. T. Smith, "Statistical resolution limits and the complexified Cramér Rao bound", IEEE Trans. on SP, 53(5): 1597-1609, 2005

[5] M. Shahram and P. Milanfar, "On the resolvability of sinusoids with nearby frequencies in the presence of noise", IEEE Trans. on SP, 53(7): 2579-2585, 2005

[6] K. Sharman and T. Durrani, "Resolving power of signal subspace methods for fnite data lengths", in Proc IEEE ICASSP, 1995

[7] S.M. Kay, *Fundamentals of Statistical Signal Processing: estimation theory*, Prentice-Hall, 1993

[8] L.L. Scharf, *Detection, Estimation and Time Series Analysis*, Adison-Wesley, 1991

[9] H. L. VanTrees, *Detection, Estimation and Modulation Theory: vol. 1*, Wiley, 1968

[10] E. L. Lehmann, *Theory of Point Estimation*, Wiley, 1983

[11] S. Oh and R. Kashyap, "A robust approach for high resolution frequency estimation", IEEE Trans. on SP, 39(3): 627-643, 1991

[12] M. P. Clark, "On the resolvability of normally distributed vector parameter estimates", IEEE Trans. on SP, 43(12): 2975-2981, 1995

[13] C. Ren, J. Galy, E. Chaumette, P. Larzabal and A. Renaux, "High resolution techniques for radar : myth or reality?", in Proc. Eurasip EUSIPCO, 2013

[14] J. Rissanen, *Information and Complexity in Statistical Modeling*, Springer Verlag, 2007

[15] B. K. Shah and C. G. Khatri, "Distribution of a Definite Quadratic Form for Non-Central Normal Variates", Ann. Math. Statist. 32(3): 883-887, 1961

[16] J. Pachares, "Note on the Distribution of a Definite Quadratic Form", Ann. Math. Statist. 26(1): 128-131, 1955

[17] A. Genz, "Numerical computation of multivariate normal probabilities", J. Comput. Graph. Statist. 1:141-149, 1992

[18] R. Muirhead, *Aspects of Multivariate Statistical Theory (2nd Edition)*, Wiley, 2005

[19] A. Renaux, P. Forster, E. Chaumette, P. Larzabal, "On the High SNR CML Estimator Full Statistical Characterization", IEEE Trans. on SP, vol 54(12): 4840-4843, 2006

[20] D.C. Rife and R.R. Boorstyn, "Single tone parameter estimation from discrete-time observations", IEEE Trans. on IT, 20(5): 591-598, 1974

[21] Athley, "Threshold region performance of maximum likelihood direction of arrival estimators", IEEE Trans. on SP, 53(4): 1359-1373, 2005

[22] C. D. Richmond, "Capon Algorithm Mean Squared Error Threshold SNR Prediction and Probability of Resolution", IEEE Trans. on SP, 53(8): 2748-2764, 2005

[23] 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 Trans. on SP, vol 55(5): 2358-2364, 2007

[24] H.L. Van Trees, *Optimum Array Processing*, Wiley-Interscience, 2002

[25] H. Lilliefors, "On the Kolmogorov-Smimov test for normality with mean and variance unknown", J. Amer. Stat. Assoc., 62(318): 399-402, 1967

[26] S. F. Yau and Y. Bresler, "Worst Case Cramér-Rao Bounds for Parametric Estimation of Superimposed Signals", IEEE Trans. on SP, 40(12): 2973-2986, 1992

[27] M.C. Wicks, E.L. Mokole, S.D. Blunt, R.S. Schneible, V.J. Amuso, *Principles of Waveform Diversity and Design*, SciTech Publishing, 2010

[28] T. Menni, E. Chaumette, P. Larzabal and J. P. Barbot, "CRB for Active Radar", in Proc. Eurasip EUSIPCO, 2011

[29] T. Menni, J. Galy, E. Chaumette and P. Larzabal, "On the versatility of constrained Cramér-Rao bound for estimation performance analysis and design of a system of measurement", IEEE Trans. on AES, accepted for publication, 2013

[30] T. Menni, E. Chaumette, P. Larzabal and J. P. Barbot, "New results on Deterministic Cramér-Rao bounds for real and complex parameters", IEEE Trans. on SP, 60(3): 1032-1049, 2012

[31] 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", Journal on ASP, Special Issue, June 2011

[32] M. Kaveh and A. Barabell, "The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise", IEEE Trans. on. ASSP, 34(2): 331-341, 1986

[33] S. Marcos, *Les Méthodes à Haute Résolution: Traitement d'Antenne et Analyse Spectrale*, Hermes, 1998

[34] H. Abeida and J.-P. Delmas, "Statistical Performance of MUSIC-like Algorithms in Resolving Noncircular Sources", IEEE Trans. on SP, 56(6): 4317-4329, 2008

[35] A. Ferreol, P. Larzabal, M. Viberg, "Statistical Analysis of the MUSIC Algorithm in the Presence of Modeling Errors, Taking Into Account the Resolution Probability", IEEE Trans. on. ASSP, 58(8): 4156-4166, 2010

[36] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, Springer-Verlag, 2005

[37] S. M. Kay, *Fundamentals of Statistical Signal Processing : Detection Theory*, Prentice Hall, 1998

[38] M. Shahram and P. Milanfar, "Imaging below the diffraction limit: A statistical analysis", IEEE Trans. on IP, 13(5): 677-689, 2004

[39] 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, 56(11): 5319-5333, 2008

[40] K. Todros and J. Tabrikian, "General Classes of Performance Lower Bounds for Parameter Estimation-Part I: Non-Bayesian Bounds for Unbiased Estimators", IEEE Trans. on IT, 56(10): 5064-5082, 2010

[41] H. B. Lee, "The Cramér-Rao bound on frequency estimates of signals closely spaced in frequency", IEEE Trans. on SP, 40(6): 1507-1517, 1992

[42] H. B. Lee, "The Cramér-Rao bound on frequency estimates of signals closely spaced in frequency (unconditional case)", IEEE Trans. on SP, 42(6): 1569-1572, 1994

[43] M. N. El Korso and R. Boyer and A. Renaux and S. Marcos, "'On the Asymptotic Resolvability Of Two Point Sources in Known Subspace Interference Using a GLRT-Based Framework", Elsevier Signal Processing, 92(10): 2471-2483, 2012

Fig. 7. Chengfang Ren

Chengfang Ren was born in Zhejiang, province of China, in 1987. He came in France since 1997 and obtained the French nationality in 2010. He received the Agrégation in applied Physics, the M.Sc. degrees in electrical engineering from École Normale Supérieure de Cachan, France, in 2011, and 2012, respectively. He is currently working toward the Ph.D. degree at the Laboratory of Signals and Systems (L2S), University of Paris-Sud, France. His research interests include performance bounds and resolution limits for parameter estimation applied to statistical signal processing.

Mohammed Nabil El Korso was born in Oran, Algeria. He received the M.Sc. in Electrical Engineering from the National Polytechnic School, Algeria in 2007. He obtained the Master Research degree in Signal and Image Processing from Paris-Sud XI University/Supélec, France in 2008. In, 2011, he obtained his PhD degree from Paris-Sud XI University. Between 2011 and 2012, he was a research scientist in the Communication Systems Group at Technische Universität Darmstadt, Germany. He was a temporary assistant professor at l'école normale supérieure de Cachan, between 2012 and 2013. Currently, he is assistant professor at Université Paris Ouest Nanterre la Défense and a member of LEME (EA4416) laboratory. His research interests include statistical signal processing, estimation/detection theory with applications to array signal processing.

Jérôme Galy was born in 1970 at Toulouse (France). He studied Electronics and Signal Processing at ENAC high school (Toulouse, France), where he obtained a diploma of engineer in 1993. He received the PhD in Signal Processing from ENSICA high school (Toulouse, France) in 1998. He has been since an Assistant Professor in electronics at University of Montpellier II (France) and a researcher at the LIRMM (Laboratoire d'Informatique, de Robotique et de Micro-Electronique de Montpellier). Main domains of interest are related to signal processing and its applications such as radar or microelectronics (Functional Electrical Stimulation).

Fig. 8. Mohammed Nabil El Korso

Fig. 9. Jérôme Galy

Eric Chaumette was born in 1965 at Chartres (France). He studied Electronics and Signal Processing both at ENAC (Toulouse, France), where he obtained a diploma of engineer in 1989, and at Toulouse University where he obtained a M.Sc. degrees in Signal Processing in 1989. From 1990 to 2007, he was with Thales in the radar studies departments and received a PhD degree in 2004 at laboratory SATIE, CNRS, Ecole Normale Supérieure de Cachan, France. From 2007 to 2013, he was with the Electromagnetic and Radar Division of the French Aerospace Lab (ONERA), Palaiseau, France, as a research engineer. He is now with the Department of Electronics, Optronics and Signal of ISAE (Institut Supérieur de l'Aéronautique et de l'Espace). Main domains of interest are related to radar scene modelling, detection and estimation theory and navigation.

Pascal Larzabal (M'93) was born in the Basque country in the south of France in 1962. He entered the Ecole Normale Supérieure of Cachan (France) in 1985 where he received the Agrégation in Electrical Engineering in 1988. He received the PhD in 1992 and the "habilitation à diriger les recherches" in 1998. He is now Professor of Electrical Engineering at University of ParisSud 11, France. He teaches electronic, signal processing, control and mathematics. From 1998 to 2003 he was the director of Electrical Engineering IUP of the University of ParisSud 11. From March 2003 to March 2007 he was at the

Fig. 10.   Eric Chaumette

Fig. 11.   Pascal Larzabal

head of electrical engineering department in the Institut Universitaire de Technologie of Cachan. Since January 2007 he is the director of the laboratory SATIE (UMR CNRS 8029, Ecole Normale Supérieure de Cachan). Since 1993 he is at the head of the Signal Processing Group of the laboratory SATIE. His research concerns estimation in array processing and spectral analysis for wavefront identification, radars, communications, tomography and medical imaging. His recent works concern estimator performances, associated minimal bounds, modelling error and geographical positioning.

Alexandre Renaux received the Ph.D. degrees in electrical engineering from École Normale Supérieure de Cachan, Cachan, France, in 2006. From 2006 to 2007, he was a Postdoctoral Research Associate at the Department of Electrical and Systems Engineering , Washington University in St. Louis, St. Louis, MO, USA. Since 2007, he has been an Assistant Professor (and currently Associate Professor since 2011) at the Physics Department of University Paris-Sud, Orsay, France. His research interests are in the field of estimation theory and lower bounds on the mean square error applied to statistical signal processing.

Fig. 12.   Alexandre Renaux