

- [7] D. Kirovski and H. S. Malvar, "Spread spectrum watermarking of audio signals," *IEEE Trans. Signal Process.*, vol. 51, no. 4, pp. 1020–1033, Apr. 2003.
- [8] C.-P. Wu, P.-C. Su, and C.-C. Jay Kuo, "Robust audio watermarking for copyright protection," in *Proc. SPIE*, Jul. 1999, vol. 3807, pp. 387–397.
- [9] W. Li and X. Y. Xue, "Audio watermarking based on music content analysis: Robust against time scale modification," in *Proc. 2nd Int. Workshop on Digital Watermarking*, Korea, 2003, pp. 289–300.
- [10] H. O. Kim, B. K. Lee, and N. Y. Lee, "Wavelet-based audio watermarking techniques: Robustness and fast synchronization [Online]. Available: <http://amath.kaist.ac.kr/research/paper/01-11.pdf>
- [11] W. Yong, H. Ji-Wu, and Y. Q. Shi, "Meaningful watermarking for audio with fast resynchronization," *J. Comput. Res. Develop.*, vol. 40, no. 20, pp. 215–220, 2003.
- [12] J. W. Huang, W. Yong, and Y. Q. Shi, "A blind audio watermarking algorithm with self-synchronization," in *Proc. IEEE Int. Symp. Circuits Syst., AZ*, 2002, vol. 3, pp. 627–630.
- [13] M. Kutter and F. A. P. Petitcolas, "A fair benchmark for image watermarking systems," in *Proc. Electron. Imag.*, 1999, vol. 3657, pp. 226–239.

## On the High-SNR Conditional Maximum-Likelihood Estimator Full Statistical Characterization

Alexandre Renaux, *Student Member, IEEE*,  
Philippe Forster, *Member, IEEE*, Eric Chaumette, and  
Pascal Larzabal, *Member, IEEE*

**Abstract**—In the field of asymptotic performance characterization of the conditional maximum-likelihood (CML) estimator, asymptotic generally refers to either the number of samples or the signal-to-noise ratio (SNR) value. The first case has been already fully characterized, although the second case has been only partially investigated. Therefore, this correspondence aims to provide a sound proof of a result, i.e., asymptotic (in SNR) Gaussianity and efficiency of the CML estimator in the multiple parameters case, generally regarded as trivial but not so far demonstrated.

**Index Terms**—Array processing, high signal-to-noise ratio (SNR), maximum likelihood, statistical efficiency.

### I. INTRODUCTION

Parameters estimation of multiple signals impinging on an antenna array is a fundamental problem in signal processing with applications to radar, sonar, digital communication and many other fields. A plethora of algorithms have been proposed in the literature in this sense (see [1]).

Perhaps the most well-known and frequently used model-based approach in signal processing is the maximum-likelihood (ML) technique. When applying the ML technique to a sensors array problem,

Manuscript received September 22, 2005; revised February 22, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Jean Pierre Delmas. This work has been performed in the framework of the European Community Contract no. 507325, NEWCOM.

A. Renaux and P. Larzabal are with Ecole Normale Supérieure de Cachan, SATIE Laboratory, 94235 Cachan Cedex, France (e-mail: renaux@satie.ens-cachan.fr; larzabal@satie.ens-cachan.fr).

P. Forster is with University Paris 10, GEA Laboratory, 92410 Ville d'Avray, France (e-mail: philippe.forster@cva.u-paris10.fr).

E. Chaumette is with Thales Naval France, 92200 Bagneux, France (e-mail: eric.chaumette@fr.thalesgroup.com).

Digital Object Identifier 10.1109/TSP.2006.882072

two main methods have been considered, depending on the model used for the signal waveforms. When the source signals are modeled as Gaussian random processes, a stochastic ML (SML) is obtained. If, on the other hand, when the source signals are modeled as unknown deterministic quantities, the resulting estimator is referred to as the conditional ML (CML) estimator (see, e.g., [2], for a review of the two methods).

Asymptotic statistical performance of these ML methods is an important field of research. For that purpose, the estimation accuracy is generally investigated by means of the Cramér–Rao bound. Since two models are used for the different ML methods, two Cramér–Rao bounds have been derived: the stochastic Cramér–Rao bound when the source signals are modelled as Gaussian random processes and the deterministic Cramér–Rao bound when the source signals are modelled as unknown deterministic quantities (see, e.g., [2], for a review of these two bounds).

In the array processing context, the term "asymptotic" can be understood in two different ways: in the number of samples or in the signal-to-noise ratio (SNR) value. At large number of samples, the statistical performance of these ML methods has been fully characterized (see [3]). Concerning the high SNR context, the non-efficiency (in comparison with the stochastic Cramér–Rao bound) and the non-Gaussianity of the SML have been recently proven in [4]. Concerning the CML method in the high-SNR framework, it is generally accepted that this estimator is Gaussian and efficient although, to our knowledge, there is no sound proof of this result in the literature in the multi-parameters case. Indeed, to the best of our knowledge, the CML estimator has been only partially investigated in [5], where the Gaussianity of the CML estimates is proved in the single-parameter case by the way of a Gaussian observation model with parameterized mean. Moreover, the asymptotic efficiency of the CML estimator in the high-SNR case has never been demonstrated. This correspondence aims to complete Kay's result, i.e., to establish the Gaussianity and the efficiency (in comparison with the deterministic Cramér–Rao bound) of the CML estimator in the multiple-parameters case. Moreover, we show how these results still hold for noncircular complex Gaussian noise. Monte Carlo simulations are provided in order to show the accuracy of the analysis.

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  $\mathbf{A}$ . The  $n$ th row and  $m$ th column element of the matrix  $\mathbf{A}$  will be denoted by  $A_{n,m}$ .  $\text{Re}\{A\}$  is the real part of  $A$ , and  $\text{Im}\{A\}$  is the imaginary part of  $A$ . The matrix transpose is indicated by a superscript  $T$  as in  $\mathbf{A}^T$ .  $|\mathbf{A}|$  is the determinant of the square matrix  $\mathbf{A}$ .  $\mathbf{I}_M$  is the identity matrix of order  $M$ .  $E\{\cdot\}$  denotes the expectation operator and  $\|\cdot\|$  the norm. A sample of a random vector  $\mathbf{a}$  is denoted  $\mathbf{a}(\omega)$ , where  $\omega$  belongs to the event space  $\Omega$ .  $o(\cdot)$  and  $o_p(\cdot)$  denote, respectively, the small "o" and the stochastic small "o" notation.

### II. OBSERVATION MODEL AND MAXIMUM-LIKELIHOOD ESTIMATOR

#### A. Observation Model

In the sequel, we consider the following general observation model:

$$\mathbf{x}(\omega) = \mathbf{m}(\mathbf{p}_0) + \mathbf{n}(\omega) \quad (1)$$

where  $\mathbf{x}(\omega)$  is a real  $M \times 1$  sample vector,  $\omega \in \Omega$ ,  $\mathbf{p} = [p_1, \dots, p_P]^T$  is the  $P \times 1$  ( $P \leq M$ ) real vector of unknown deterministic parameters of interest with true value  $\mathbf{p}_0$ ,  $\mathbf{m}$  is a  $M \times 1$  real deterministic vector depending (generally nonlinearly) on  $\mathbf{p}$  which is assumed to be identifiable from  $\mathbf{m}(\mathbf{p})$ .  $\mathbf{n}(\omega)$  is the  $M \times 1$  additive noise vector, which is

a sample of a random Gaussian, zero-mean vector  $\mathbf{n}$  with covariance matrix  $\sigma^2 \mathbf{C}$ .  $\mathbf{C}$  is assumed to be known and  $\sigma^2$  is unknown.<sup>1</sup>

The log-likelihood  $\ln f(\mathbf{x}(\omega) | \mathbf{p}, \sigma^2)$  of the observations is

$$\ln f(\mathbf{x}(\omega) | \mathbf{p}, \sigma^2) = -\frac{1}{2} \ln \left( (2\pi)^M |\sigma^2 \mathbf{C}| \right) - \frac{1}{2\sigma^2} (\mathbf{x}(\omega) - \mathbf{m}(\mathbf{p}))^T \mathbf{C}^{-1} (\mathbf{x}(\omega) - \mathbf{m}(\mathbf{p})). \quad (2)$$

Let us note that the study with the observation  $\check{\mathbf{x}}(\omega) = \check{\mathbf{m}}(\mathbf{p}_0) + \check{\mathbf{n}}(\omega)$ , where  $\check{\mathbf{x}}(\omega)$ ,  $\check{\mathbf{m}}$  and  $\check{\mathbf{n}}(\omega)$  are complex, can be handled by the real model (1). Indeed, by stacking the real and imaginary parts of  $\check{\mathbf{x}}(\omega)$ ,  $\check{\mathbf{m}}$ , and  $\check{\mathbf{n}}(\omega)$ , one obtains

$$\begin{pmatrix} \text{Re} \{ \check{\mathbf{x}}(\omega) \} \\ \text{Im} \{ \check{\mathbf{x}}(\omega) \} \end{pmatrix} = \begin{pmatrix} \text{Re} \{ \check{\mathbf{m}}(\mathbf{p}_0) \} \\ \text{Im} \{ \check{\mathbf{m}}(\mathbf{p}_0) \} \end{pmatrix} + \begin{pmatrix} \text{Re} \{ \check{\mathbf{n}}(\omega) \} \\ \text{Im} \{ \check{\mathbf{n}}(\omega) \} \end{pmatrix} \quad (3)$$

which is similar to the (1). Modifications of (2) are straightforward and lead to an augmented covariance matrix ( $2M \times 2M$ ) instead of  $\mathbf{C}$  taking into account the possible noncircularity of the noise [6].

So, the two following important problems in the array processing context can be statistically characterized in the framework of model (1):

- The CML method (with the notations of [2, eq. (4.16)]):

$$\mathbf{X}_N = \mathbf{A}(\boldsymbol{\theta}_0) \mathbf{S}_N + \mathbf{N}_N \quad (4)$$

by setting

$$\mathbf{m}(\mathbf{p}) = \begin{bmatrix} \text{Re}(\text{vec}(\mathbf{A}(\boldsymbol{\theta}_0) \mathbf{S}_N))^T & \text{Im}(\text{vec}(\mathbf{A}(\boldsymbol{\theta}_0) \mathbf{S}_N))^T \end{bmatrix} \quad (5)$$

where  $\mathbf{p} = [\boldsymbol{\theta}^T, \text{vec}(\mathbf{S}_N)^T]^T$ .

- The so-called known waveforms model (see, e.g., [7]) (with the notations of [7, eq. (8)])

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta}_0) \mathbf{P}(t) \boldsymbol{\alpha} + \mathbf{n}(t) \quad (6)$$

by setting

$$\mathbf{m}(\mathbf{p}) = \begin{bmatrix} \text{Re}(\text{vec}(\mathbf{A}(\boldsymbol{\theta}_0) \mathbf{P}(t) \boldsymbol{\alpha}))^T & \text{Im}(\text{vec}(\mathbf{A}(\boldsymbol{\theta}_0) \mathbf{P}(t) \boldsymbol{\alpha}))^T \end{bmatrix}^T \quad (7)$$

where  $\mathbf{p} = [\boldsymbol{\theta}^T, \boldsymbol{\alpha}^T]^T$ . This model finds applications in active radar and in data-aided processing for mobile communications.

### B. Maximum-Likelihood Estimator

The ML estimate of  $\mathbf{p}$  is given by

$$\hat{\mathbf{p}}(\omega) = \arg \min_{\mathbf{p}} \left\{ (\mathbf{x}(\omega) - \mathbf{m}(\mathbf{p}))^T \mathbf{C}^{-1} (\mathbf{x}(\omega) - \mathbf{m}(\mathbf{p})) \right\}. \quad (8)$$

The solution of (8) is obtained by setting the gradient of the criterion  $(\mathbf{x}(\omega) - \mathbf{m}(\mathbf{p}))^T \mathbf{C}^{-1} (\mathbf{x}(\omega) - \mathbf{m}(\mathbf{p}))$  equal to zero. This leads to the set of  $P$  equations where the dependence on  $\mathbf{n}(\omega)$  is explicitly shown

$$\mathbf{g}(\mathbf{p}, \mathbf{n}(\omega)) |_{\mathbf{p}=\hat{\mathbf{p}}(\omega)} = 0 \quad (9)$$

where the elements of  $\mathbf{g} = [g_1(\mathbf{p}, \mathbf{n}(\omega)), \dots, g_P(\mathbf{p}, \mathbf{n}(\omega))]^T$ , for  $i = 1, \dots, P$ , are given by

$$\begin{aligned} g_i(\mathbf{p}, \mathbf{n}(\omega)) &= (\mathbf{x}(\omega) - \mathbf{m}(\mathbf{p}))^T \mathbf{C}^{-1} \frac{\partial \mathbf{m}(\mathbf{p})}{\partial p_i} \\ &= (\mathbf{m}(\mathbf{p}_0) - \mathbf{m}(\mathbf{p}) + \mathbf{n}(\omega))^T \\ &\quad \times \mathbf{C}^{-1} \frac{\partial \mathbf{m}(\mathbf{p})}{\partial p_i}. \end{aligned} \quad (10)$$

<sup>1</sup>The high-SNR analysis is equivalent to an analysis of  $\sigma \rightarrow 0$ .

## III. PERFORMANCE AT HIGH SNR

An improvement on the well-known Kay approach [5] is to resort to the implicit function theorem (see [8, Theorem 9.28]): Indeed, this theorem not only provides a framework for a sound demonstration of the asymptotical Gaussian behavior of CML estimator, but also, the analytical expression of the asymptotic estimator covariance matrix that allows to establish the asymptotic efficiency.

### A. Background: The Implicit Function Theorem

Let  $\mathbf{f}(\mathbf{u}, \mathbf{v}) = [f_1(\mathbf{u}, \mathbf{v}), \dots, f_P(\mathbf{u}, \mathbf{v})]^T$  be a function of  $IR^P \times IR^M \rightarrow IR^P$ . Let us assume the following.

- *Assumption A1*):  $f_i(\mathbf{u}, \mathbf{v})$  for  $i = 1, \dots, P$  are differentiable functions on a neighborhood of the point  $(\mathbf{u}_0, \mathbf{v}_0)$  in  $IR^P \times IR^M$ .
- *Assumption A2*):  $f_i(\mathbf{u}_0, \mathbf{v}_0) = 0$  for  $i = 1, \dots, P$ .
- *Assumption A3*): the  $P \times P$  Jacobian matrix  $\Phi$  of  $f_i(\mathbf{u}, \mathbf{v})$  with respect to  $\mathbf{u}$  is nonsingular at  $(\mathbf{u}_0, \mathbf{v}_0)$ .

Then, there is a neighborhood  $V$  of the point  $\mathbf{v}_0$  in  $IR^M$ , there is a neighborhood  $U$  of the point  $\mathbf{u}_0$  in  $IR^P$ , and there is a unique mapping  $\varphi : V \rightarrow U$  such that  $\varphi(\mathbf{v}_0) = \mathbf{u}_0$  and  $f_i(\varphi(\mathbf{v}), \mathbf{v}) = 0$  for  $i = 1, \dots, P$  and for all  $\mathbf{v}$  in  $V$ . Furthermore,  $\varphi$  is differentiable, and we have

$$\varphi(\mathbf{v}) - \mathbf{u}_0 = -\Phi^{-1} \Psi(\mathbf{v} - \mathbf{v}_0) + \mathbf{r}(\mathbf{v} - \mathbf{v}_0) \quad (11)$$

where the remainder  $\mathbf{r}(\mathbf{v} - \mathbf{v}_0) = o(\|\mathbf{v} - \mathbf{v}_0\|)$  and, by setting  $\mathbf{u} = [u_1, \dots, u_P]^T$  and  $\mathbf{v} = [v_1, \dots, v_M]^T$

$$\Phi = \left[ \frac{\partial f}{\partial u_1} \Big|_{(\mathbf{u}_0, \mathbf{v}_0)}, \dots, \frac{\partial f}{\partial u_P} \Big|_{(\mathbf{u}_0, \mathbf{v}_0)} \right] \quad (12)$$

and

$$\Psi = \left[ \frac{\partial f}{\partial v_1} \Big|_{(\mathbf{u}_0, \mathbf{v}_0)}, \dots, \frac{\partial f}{\partial v_M} \Big|_{(\mathbf{u}_0, \mathbf{v}_0)} \right]. \quad (13)$$

### B. Application to the Criterion Gradient

Let us now detail the structure of  $\Phi$  and  $\Psi$  for the specific function  $\mathbf{g}$  of (10). For the elements of  $\Psi$ , we have

$$\Psi_{i,j} = \frac{\partial g_i(\mathbf{u}, \mathbf{v})}{\partial v_j} \Big|_{(\mathbf{u}_0, \mathbf{v}_0)} = \frac{\partial \mathbf{v}^T}{\partial v_j} \mathbf{C}^{-1} \frac{\partial \mathbf{m}}{\partial u_i} \Big|_{(\mathbf{u}_0, \mathbf{v}_0)} \quad (14)$$

with  $1 \leq i \leq P$  and  $1 \leq j \leq M$ . Concerning the elements of  $\Phi$ , we have

$$\Phi_{i,j} = \frac{\partial g_i(\mathbf{u}, \mathbf{v})}{\partial u_j} \Big|_{(\mathbf{u}_0, \mathbf{v}_0)} = -\frac{\partial \mathbf{m}^T}{\partial u_j} \mathbf{C}^{-1} \frac{\partial \mathbf{m}}{\partial u_i} \Big|_{\mathbf{u}_0} \quad (15)$$

with  $1 \leq i, j \leq P$ .

In other words

$$\Psi = \frac{\partial \mathbf{m}^T}{\partial \mathbf{u}} \mathbf{C}^{-1} \frac{\partial \mathbf{v}}{\partial \mathbf{v}^T} \quad \text{and} \quad \Phi = -\frac{\partial \mathbf{m}^T}{\partial \mathbf{u}} \mathbf{C}^{-1} \frac{\partial \mathbf{m}}{\partial \mathbf{u}^T}. \quad (16)$$

One can notice that  $\Phi$  is linked to  $\Psi$  by

$$\Phi = -\Psi \mathbf{C} \Psi^T \quad (17)$$

since  $\partial \mathbf{v} / \partial \mathbf{v}^T = \mathbf{I}_M$ .

Let us now apply the implicit function theorem by setting  $\mathbf{v} = \mathbf{n}(\omega)$ ,  $\varphi(\mathbf{v}) = \hat{\mathbf{p}}(\omega)$ ,  $\mathbf{u}_0 = \mathbf{p}_0$  and  $\mathbf{v}_0 = \mathbf{0}$ . Let us note that A2) is obviously satisfied since, without noise, the criterion  $(\mathbf{x}(\omega) - \mathbf{m}(\mathbf{p}))^T \mathbf{C}^{-1} (\mathbf{x}(\omega) - \mathbf{m}(\mathbf{p}))$  is minimum and equal to  $\mathbf{0}$  for  $\mathbf{p} = \mathbf{p}_0$ . A1) and A3) will be assumed to be satisfied in the sequel.

Consequently, the implicit function theorem leads to

$$\hat{\mathbf{p}}(\omega) - \mathbf{p}_0 = -\Phi^{-1} \Psi \mathbf{n}(\omega) + \mathbf{r}(\mathbf{n}(\omega)), \quad \forall \omega \in \Omega. \quad (18)$$

### C. Asymptotic Gaussianity of the ML Estimator

Let us set  $\mathbf{n}_k$  a sequence of a Gaussian random noise with zero means and covariance matrices  $\sigma_k^2 \mathbf{C}$ .  $\sigma_k$  is a sequence such as  $\sigma_k \rightarrow 0$  when  $k \rightarrow \infty$ . Let us set  $\hat{\mathbf{p}}_k(\omega)$  and  $\mathbf{n}_k(\omega)$  two sequences of the random vector  $\hat{\mathbf{p}}_k$  and  $\mathbf{n}_k$ , respectively.  $\hat{\mathbf{p}}_k(\omega)$  is the corresponding solution of (9). Let us set  $\Delta \mathbf{p}_k(\omega) = (1/\sigma_k)(\hat{\mathbf{p}}_k(\omega) - \mathbf{p}_0)$  a sequence of the random vector  $\Delta \mathbf{p}_k$ . Then, the high-SNR analysis of the ML estimator is given by the behavior of the random vector  $\Delta \mathbf{p}_k$  when  $k \rightarrow \infty$ . From (18), we have

$$\Delta \mathbf{p}_k(\omega) = -\Phi^{-1} \Psi \frac{\mathbf{n}_k(\omega)}{\sigma_k} + \frac{\mathbf{r}(\mathbf{n}_k(\omega))}{\sigma_k}, \quad \forall \omega \in \Omega \quad (19)$$

or equivalently in terms of random vectors  $\mathbf{n}_k$  and  $\Delta \mathbf{p}_k$

$$\Delta \mathbf{p}_k = -\Phi^{-1} \Psi \frac{\mathbf{n}_k}{\sigma_k} + \frac{\mathbf{r}(\mathbf{n}_k)}{\sigma_k}. \quad (20)$$

By using [9, Lemma 2.12 (i)], we have

$$\mathbf{r}(\mathbf{n}_k) = o_p(\|\mathbf{n}_k\|). \quad (21)$$

We will now study the two terms of the right-hand side of (20). For that purpose, note that  $\mathbf{n}_k/\sigma_k$  follows a Gaussian distribution with zero mean and covariance matrix  $\mathbf{C}$ . The term  $(\mathbf{r}(\mathbf{n}_k)/\sigma_k) = (o_p(\|\mathbf{n}_k\|)/\sigma_k)$  can be rewritten as

$$\frac{o_p(\|\mathbf{n}_k\|)}{\sigma_k} = \frac{\|\mathbf{n}_k\|}{\sigma_k} \frac{o_p(\|\mathbf{n}_k\|)}{\|\mathbf{n}_k\|} \quad (22)$$

where  $o_p(\|\mathbf{n}_k\|)/\|\mathbf{n}_k\|$  converges in probability to zero by definition and where  $\|\mathbf{n}_k\|/\sigma_k$  follows a chi law which does not depend on  $k$ . Consequently,  $o_p(\|\mathbf{n}_k\|)/\sigma_k$  converges in probability to zero (see [10, pp. 122]).

Concerning the term  $-\Phi^{-1} \Psi (\mathbf{n}_k/\sigma_k)$ , it follows a Gaussian distribution with zero mean and covariance matrix  $\mathbf{\Gamma}$  equal to

$$\begin{aligned} \mathbf{\Gamma} &= \Phi^{-1} \Psi \mathbf{C} (\Phi^{-1} \Psi)^T \\ &= \Phi^{-1} \Psi \mathbf{C} \Psi^T \Phi^{-T} \\ &= -\Phi^{-1} = (\Psi \mathbf{C} \Psi^T)^{-1} \end{aligned} \quad (23)$$

thanks to (17). Consequently,  $\Delta \mathbf{p}_k$  converges in distribution to a zero-mean Gaussian random vector with covariance  $\mathbf{\Gamma} = (\Psi \mathbf{C} \Psi^T)^{-1}$  when  $k \rightarrow \infty$  or equivalently when  $\sigma_k \rightarrow 0$ . This proves the asymptotic Gaussianity of the ML estimator at high SNR.

### D. Asymptotic Efficiency of the ML Estimator

The closed-form of the error covariance matrix obtained by the means of the implicit function theorem allows to establish the asymptotic efficiency of the ML estimator by a direct comparison with the Cramér–Rao lower bound (CRLB). For that purpose, let us compute the Fisher information matrix (FIM), which reduces to a block diagonal form thanks to the decoupling between parameter  $\mathbf{p}$  and  $\sigma^2$  [11]

$$\mathbf{FIM}(\mathbf{p}_0, \sigma^2) = \begin{pmatrix} \mathbf{FIM}(\mathbf{p}_0) & \mathbf{0} \\ \mathbf{0} & \mathbf{FIM}(\sigma^2) \end{pmatrix}. \quad (24)$$

Consequently, the CRLB for  $\mathbf{p}_0$  is given by

$$\begin{aligned} \mathbf{CRLB}(\mathbf{p}_0) &= \mathbf{FIM}^{-1}(\mathbf{p}_0) \\ &= - \left( E \left[ \frac{\partial^2 \ln f(\mathbf{x}; \mathbf{p}, \sigma^2)}{\partial \mathbf{p} \partial \mathbf{p}^T} \right] \right)^{-1} \\ &= 2\sigma^2 \left( E \left[ \frac{\partial^2 (\mathbf{x} - \mathbf{m}(\mathbf{p}))^T \mathbf{C}^{-1} (\mathbf{x} - \mathbf{m}(\mathbf{p}))}{\partial \mathbf{p} \partial \mathbf{p}^T} \right] \right)^{-1} \\ &= -\sigma^2 \left( E \left[ \frac{\partial \mathbf{g}(\mathbf{p}, \mathbf{n})}{\partial \mathbf{p}^T} \right] \right)^{-1} \\ &= -\sigma^2 \Phi^{-1} = \sigma^2 (\Psi \mathbf{C} \Psi^T)^{-1}. \end{aligned} \quad (25)$$

Finally, the ML estimator is efficient at high SNR since

$$\frac{1}{\sigma^2} \mathbf{CRLB}(\mathbf{p}_0) = \mathbf{\Gamma}. \quad (26)$$

## IV. SIMULATION RESULTS

In this section, simulations are performed to illustrate the validity of our analysis in the framework of the CML estimator described in [2]. The CRLB is computed according to [2, eq. (4.68)]. In all simulations, the array is a uniform linear array of  $M = 10$  sensors with half-wavelength spacing. Direction-of-arrivals (DOA) are given with respect to the broadside. Monte Carlo simulations have been performed with  $r = 1000$  independent realizations. The ML DOA estimation is performed with a Gauss–Newton algorithm thanks to a global search over a grid.

### A. Efficiency

Let us consider the case of two equi-powered sources located at  $0^\circ$  and  $3^\circ$  (the array beamwidth is equal to  $10^\circ$ ). The CML DOA estimation is performed with  $T = 10$  snapshots. Fig. 1 displays the behavior of the CML empirical variance and the CRLB versus SNR. As expected, the efficiency of the CML estimator at high SNR is observed.

We also observe the well known threshold effect [12] of the estimator variance when the SNR becomes weak (approximately 5 dB in this case). This phenomena due to outliers gives the validity domain in term of efficiency of this asymptotic analysis (see [13] for more details concerning the CML threshold prediction).

### B. Gaussianity

In order to emphasize the high SNR Gaussianity of the CML estimator, we have used a Lilliefors goodness-of-fit test [14]. This test evaluates the hypothesis that a sample  $[y_1 \cdots y_L]$  has a normal distribution with unspecified mean and variance against the alternative hypothesis that the sample does not have a normal distribution. This test is close to the well-known Kolmogorov–Smirnov test, which requires the specification of the mean and the variance.

The Lilliefors test statistic  $T_{stat}$  is the maximum vertical distance between the empirical cumulative distribution function  $F(s)$  of the score series

$$z_k = \left( y_k - \frac{1}{L} \sum_{i=1}^L y_i \right) / \sqrt{\frac{1}{L} \sum_{j=1}^L \left( y_j - \frac{1}{L} \sum_{i=1}^L y_i \right)^2} \quad (27)$$

and the cumulative distribution function of the standard normal distribution  $Q(s)$ . In other words

$$T_{stat} = \sup_{s \in \mathbb{R}} |F(s) - Q(s)|. \quad (28)$$

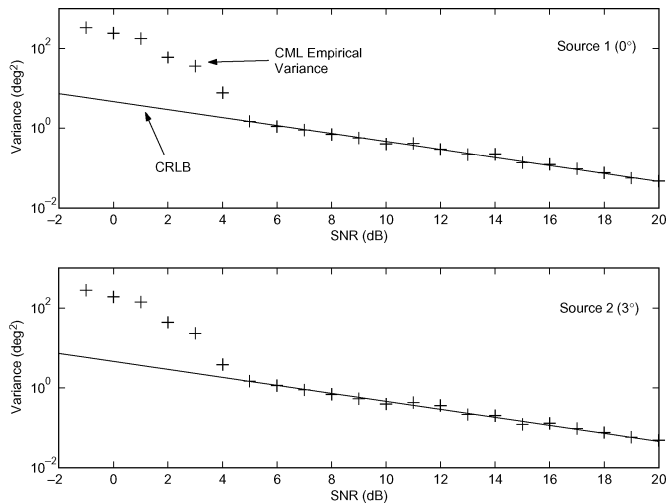


Fig. 1. CRLB and empirical variance of the CML estimator versus SNR for two sources.  $M = 10$  sensors and  $T = 10$  snapshots.

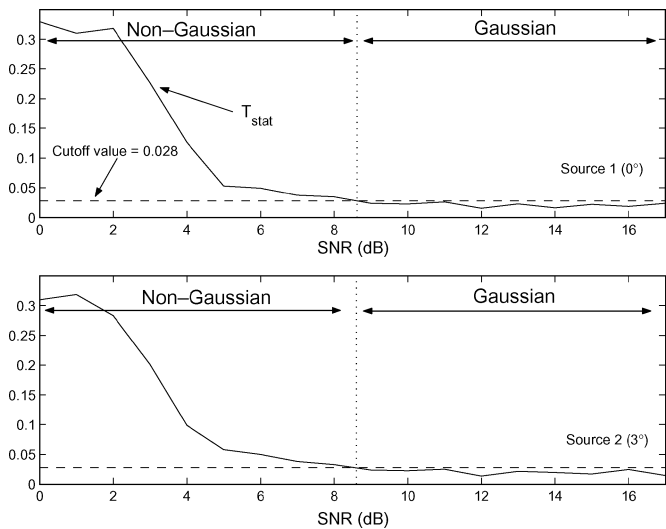


Fig. 2. Lilliefors test  $T_{stat}$  and cutoff value for the CML estimates versus SNR.  $M = 10$  sensors,  $T = 10$  snapshots,  $r = 1000$  runs.

The Gaussianity hypothesis is rejected at the significance level  $\alpha$  (equal to 5% in the following example) if  $T$  exceeds the  $1 - \alpha$  quantile in a Lilliefors table of quantiles [15] equals to  $0.886/\sqrt{r}$  for  $r > 30$ . Here, this cutoff value is  $0.886/\sqrt{r} = 0.028$ .

The test has been performed for the aforementioned scenario (two sources located at  $0^\circ$  and  $3^\circ$ , 10 snapshots, 10 sensors, 1000 runs). Fig. 2 displays the behavior of  $T_{stat}$  versus SNR. The Gaussianity of the CML estimator is observed for an SNR higher than approximately 9 dB.

V. CONCLUSION

In this correspondence, we have completed initial Kay works on the CML characterization at high SNR. Thanks to the implicit function theorem, we have provided a sound proof of its asymptotic Gaussianity and efficiency in the multiple parameters case.

REFERENCES

- [1] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, 1996.
- [2] 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. Haykin, J. Litva, and T. J. Shepherd, Eds. Berlin, Germany: Springer-Verlag, 1993, ch. 4, pp. 99–151.
- [3] P. Stoica and A. Nehorai, "Performances study of conditional and unconditional direction of arrival estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 1783–1795, Oct. 1990.
- [4] 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. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, Montreal, QC, Canada, May 2004, vol. 2, pp. 121–124.
- [5] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1993, vol. 1.
- [6] A. V. D. Bos, "The multivariate complex normal distribution—A generalization," *IEEE Trans. Inf. Theory*, vol. 41, pp. 537–539, Mar. 1995.
- [7] J. Li and R. T. Compton, "Maximum likelihood angle estimation for signals with known waveforms," *IEEE Trans. Signal Process.*, vol. 41, no. 9, pp. 2850–2862, Sep. 1993.
- [8] W. Rudin, *Principles of Mathematical Analysis*. New York: McGraw-Hill, 1976.
- [9] A. W. van der Vaart, *Asymptotic Statistics*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
- [10] C. R. Rao, *Linear Statistical Inference and Its Applications*. New York: Wiley, 2002.
- [11] A. Zeira and A. Nehorai, "Frequency domain Cramer–Rao bound for Gaussian processes," *IEEE Trans. Signal Process.*, vol. 38, no. 6, pp. 1063–1066, Jun. 1990.
- [12] H. L. V. Trees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968, vol. 1.
- [13] F. Athley, "Threshold region performance of maximum likelihood direction of arrival estimators," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1359–1373, Apr. 2005.
- [14] H. Lilliefors, "On the Kolmogorov–Smirnov test for normality with mean and variance unknown," *J. Amer. Stat. Assoc.*, vol. 62, pp. 399–402, 1967.
- [15] W. J. Conover, *Practical Nonparametric Statistics*. New York: Wiley, 1980.