Unconditional maximum likelihood performance at finite number of samples and high signal-to-noise ratio

Alexandre Renaux, Philippe Forster, Eric Boyer, Pascal Larzabal

Unconditional Maximum Likelihood Performance At Finite Number Of Samples And High Signal To Noise Ratio

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

Abstract

This correspondence deals with the problem of estimating signal parameters using an array of sensors. In source localization, two main Maximum Likelihood methods have been introduced: the Conditional Maximum Likelihood method which assumes the source signals nonrandom and the Unconditional Maximum Likelihood method which assumes the source signals random. Many theoretical investigations have been already conducted for the large samples statistical properties. This paper studies the behavior of Unconditional Maximum Likelihood at high Signal to Noise Ratio for finite samples. We first establish the equivalence between the Unconditional and the Conditional Maximum Likelihood Criterions at high Signal to Noise Ratio. Then, thanks to this equivalence we prove the non-Gaussianity and the non-efficiency of the Unconditional Maximum Likelihood estimator. We also rediscover the closed-form expressions of the probability density function and of the variance of the estimates in the one source scenario and we derive a closed-form expression of this estimator variance in the two sources scenario.

Index Terms

Asymptotic performance, Unconditional Maximum Likelihood, finite number of data, high Signal to Noise Ratio, Cramér-Rao bound.

I. Introduction

Direction-Of-Arrival (DOA) estimation using an array of spatially distributed sensors has received a significant attention in the signal processing literature. Initial motivation was the military framework with applications such as radar and sonar. More recently, DOA estimation has also been applied to other frameworks such as friendly communication. For these numerous applications, the resolving power of the algorithm is of the utmost importance. This is why, various algorithms have been proposed in the literature with a resolution which is better than the

Alexandre Renaux, Eric Boyer and Pascal Larzabal are with Ecole Normale Superieure of Cachan, SATIE Laboratory, 61, avenue du President Wilson 94235 Cachan, France. (e-mail: renaux@satie.ens-cachan.fr).

Philippe Forster is with University of Paris 10, GEA Laboratory, 1 Chemin Desvallieres, 92410 Ville d'Avray, France.

This work is partially supported by the European Community Contract no. 507325, NEWCOM

traditional Rayleigh beam-width [1]–[4]. An alternative to these algorithms is the Maximum Likelihood (ML) method which has been extensively studied for its attractive statistical properties. When applying the ML technique to the sensor array problem, two main methods have been considered, depending on the model assumption on the signal waveforms. When the source signals are modelled as Gaussian random processes, a Unconditional ML (UML) is obtained (see [5]–[7]). If, on the other hand,when the source signals are modelled as unknown deterministic quantities, the resulting estimator is referred as the Conditional ML (CML) estimator (see [7]–[9]).

This paper deals with the asymptotic performance of the UML method. The term ”asymptotic” can be understood in two different ways: in the number T of samples and in the Signal to Noise Ratio (SNR). Asymptotic performance in the number T of samples (for finite SNR) have been extensively investigated [7], [10]–[12]. Concerning the asymptotic performance when the SNR tends to infinity (for finite T), few works are available. Under the deterministic model, the CML is Gaussian and efficient (it achieves the conditional Cramér-Rao bound) [13], [14]. The present work is devoted to the analysis of the UML behavior, under the stochastic signals model, when the SNR tends to infinity (for finite T): this is the meaning of asymptotic in this paper. Note that in [15], Athley has observed, with the help of simulation results, that the UML estimates are non-efficient at high SNR. The proposed paper aims to soundly establish the asymptotic non-Gaussianity and the asymptotic non-efficiency (in comparison with the unconditional Cramér-Rao bound) of the UML estimator in the multiple parameters case.

We have already investigated the UML asymptotic behavior for a single source [16]. The proposed paper generalizes these preliminary results to multiple sources case, providing an extended and detailed version of works reported in conference papers [17], [18]. We first show that, at high SNR, Unconditional and Conditional Maximum Likelihood Criterions (UMLC and CMLC) are equivalent in the sense that, with the same observations, they give the same estimates. This preliminary result is the key point for proving that the UML estimates are non-Gaussian and non-efficient when the SNR tends to infinity for any number of sources contrary to the large number of observations case. Finally, we establish a closed-form of the UML estimator variance in the case of two uncorrelated sources for centro-symmetric arrays.

In the sequel, a sample of a random vector $\mathbf{y}$ is denoted $\mathbf{y}(\omega)$ where ω belongs to the event space Ω.

II. Problem setup

Let us consider the classical problem of localizing N narrow-band sources impinging on an array of M sensors. The vector $\mathbf{x}_t(\omega)$ of sensors outputs is given by the following equation [9]:

$$\mathbf{x}_t(\omega) = \mathbf{A}(\boldsymbol{\theta}_0)\mathbf{s}_t(\omega) + \mathbf{n}_t(\omega), \tag{1}$$

where $t = 1, 2, \cdots, T$ and where T is the number of snapshots. $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_N]^T$ denotes the candidate vector of the N DOA's whose exact value is $\boldsymbol{\theta}_0$. $\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \cdots, \mathbf{a}(\theta_N)]$ is the $M \times N$ steering matrix. $\mathbf{s}_t(\omega)$ is the $N \times 1$ vector of the N source signals. $\mathbf{n}_t(\omega)$ is the $M \times 1$ vector of the noise.

In the sequel $\mathbf{N}(\omega) = [\mathbf{n}_1(\omega), \mathbf{n}_2(\omega), \cdots, \mathbf{n}_T(\omega)]$, and $\mathbf{S}(\omega) = [\mathbf{s}_1(\omega), \mathbf{s}_2(\omega), \cdots, \mathbf{s}_T(\omega)]$.

The following assumptions will be used:

A1 The signal $\mathbf{s}_t(\omega)$ is the sample of the random vector $\mathbf{s}_t$ which is complex, circular, Gaussian, temporally white with zero mean and covariance matrix $\mathbf{\Sigma_s} = \mathbb{E}\left[\mathbf{s}\mathbf{s}^H\right]$ where $\mathbb{E}$ denotes the expectation operator.

A2 The noise $\mathbf{n}_t(\omega)$ is the sample of the random vector $\mathbf{n}_t$ which is complex, circular, Gaussian, spatially and temporally white with zero mean and covariance matrix $\mathbf{\Sigma_n} = \mathbb{E}\left[\mathbf{n}\mathbf{n}^H\right] = \sigma^2\mathbf{I}_M$ where $\mathbf{I}_M$ is the $M \times M$ identity matrix.

A3 $\|\mathbf{a}(\theta)\| = \sqrt{M}$.

A4 The number of sources is less than the number of sensors, $M > N$.

Note that the model used in **A1** differs from the conditional model, for which the signal $\mathbf{s}_t$ is deterministic.

III. High SNR Equivalence of the Conditional Criterion and Unconditional Criterion

In this section, we recall the definition of the CMLC and of the UMLC and we prove the equivalence of these two criterions at high SNR in the sense where, with the same observations, they lead to the same estimates.

A. Conditional and Unconditional Maximum Likelihood Criterion

In the conditional model case, the DOA's are obtained by minimization of the concentrated criterion [9]:

$$C_{CML}(\boldsymbol{\theta}) = \frac{1}{M-N} Tr\left\{\mathbf{\Pi}_{\mathbf{A}}^{\perp}(\boldsymbol{\theta})\widehat{\mathbf{\Sigma}}_{\mathbf{x}}\right\}, \tag{2}$$

where $Tr\{.\}$ is the trace operator, where $\widehat{\mathbf{\Sigma}}_{\mathbf{x}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{x}_t(\omega)\mathbf{x}_t^H(\omega)$ is the observations sample covariance matrix and $\mathbf{\Pi}_{\mathbf{A}}^{\perp}(\boldsymbol{\theta}) = \mathbf{I}_M - \mathbf{A}(\boldsymbol{\theta})\left(\mathbf{A}^H(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\right)^{-1}\mathbf{A}^H(\boldsymbol{\theta})$ denotes the orthogonal projector onto the noise subspace. In the sequel, the Moore-Penrose inverse $\left(\mathbf{A}^H(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\right)^{-1}\mathbf{A}^H(\boldsymbol{\theta})$, where $\mathbf{A}(\boldsymbol{\theta})$ is a full-column rank matrix, will be denoted $\mathbf{A}^{\dagger}(\boldsymbol{\theta})$.

In the stochastic model case, the DOA's are obtained by minimization of the concentrated criterion [9]:

$$C_{UML}(\boldsymbol{\theta}) = \left|\mathbf{A}(\boldsymbol{\theta})\widehat{\mathbf{R}}_{\mathbf{s}}\mathbf{A}^H(\boldsymbol{\theta}) + \hat{\sigma}^2\mathbf{I}_M\right|, \tag{3}$$

with

$$\begin{cases} \widehat{\mathbf{R}}_{\mathbf{s}}(\boldsymbol{\theta}) = \mathbf{A}^{\dagger}(\boldsymbol{\theta})\left(\widehat{\mathbf{\Sigma}}_{\mathbf{x}} - \hat{\sigma}^2(\boldsymbol{\theta})\mathbf{I}_M\right)\mathbf{A}^{\dagger H}(\boldsymbol{\theta}), \\ \hat{\sigma}^2(\boldsymbol{\theta}) = \frac{1}{M-N} Tr\left\{\mathbf{\Pi}_{\mathbf{A}}^{\perp}(\boldsymbol{\theta})\hat{\mathbf{\Sigma}}_{\mathbf{x}}\right\}, \end{cases} \tag{4}$$

where $|.|$ denotes the determinant.

By substituting (2) and (4) into (3) we straightforwardly obtain:

$$C_{UML}(\boldsymbol{\theta}) = \left|\mathbf{\Pi}_{\mathbf{A}}(\boldsymbol{\theta})\widehat{\mathbf{\Sigma}}_{\mathbf{x}}\mathbf{\Pi}_{\mathbf{A}}(\boldsymbol{\theta}) + C_{CML}(\boldsymbol{\theta})\mathbf{\Pi}_{\mathbf{A}}^{\perp}(\boldsymbol{\theta})\right|, \tag{5}$$

where $\mathbf{\Pi}_{\mathbf{A}}(\boldsymbol{\theta}) = \mathbf{A}(\boldsymbol{\theta})\mathbf{A}^{\dagger}(\boldsymbol{\theta})$ denotes the orthogonal projector onto the signal subspace.

B. Equivalence

Proposition 1: At high SNR, the UMLC and the CMLC are equivalent in the sense where the difference of DOA's obtained by minimization of $C_{UML}(\boldsymbol{\theta})$ and $C_{CML}(\boldsymbol{\theta})$ tends to zero in probability when SNR tends to infinity.

Proof: Let $\mathbf{E_s}(\boldsymbol{\theta})$ and $\mathbf{E_n}(\boldsymbol{\theta})$ be the $M \times N$ and $M \times (M-N)$ matrices built with the orthonormal bases of signal and noise subspaces and set $\mathbf{E}(\boldsymbol{\theta})$ be the $M \times M$ matrix such that $\mathbf{E}(\boldsymbol{\theta}) = [\mathbf{E_s}(\boldsymbol{\theta}), \mathbf{E_n}(\boldsymbol{\theta})]$. Equation (5) becomes:

$$\begin{aligned} C_{UML}(\boldsymbol{\theta}) &= \left|\mathbf{E}^H(\boldsymbol{\theta})\left(\boldsymbol{\Pi}_{\mathbf{A}}(\boldsymbol{\theta})\widehat{\boldsymbol{\Sigma}}_{\mathbf{x}}\boldsymbol{\Pi}_{\mathbf{A}}(\boldsymbol{\theta}) + C_{CML}(\boldsymbol{\theta})\boldsymbol{\Pi}_{\mathbf{A}}^{\perp}(\boldsymbol{\theta})\right)\mathbf{E}(\boldsymbol{\theta})\right| \\ &= \left|\begin{array}{cc} \mathbf{E_s}^H(\boldsymbol{\theta})\widehat{\boldsymbol{\Sigma}}_{\mathbf{x}}\mathbf{E_s}(\boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{0} & C_{CML}(\boldsymbol{\theta})\mathbf{I}_{M-N} \end{array}\right|. \end{aligned} \tag{6}$$

The matrix involved in the determinant (6) is block diagonal so that $C_{UML}(\boldsymbol{\theta})$ can also be written as follows by writing down explicitly the dependance of each terms on the noise and $\boldsymbol{\theta}$:

$$C_{UML}(\boldsymbol{\theta}, \mathbf{N}(\omega)) = \left|\mathbf{E_s}^H(\boldsymbol{\theta})\widehat{\boldsymbol{\Sigma}}_{\mathbf{x}}(\boldsymbol{\theta}, \mathbf{N}(\omega))\mathbf{E_s}(\boldsymbol{\theta})\right| C_{CML}(\boldsymbol{\theta}, \mathbf{N}(\omega))^{M-N}. \tag{7}$$

Note that the minimization of $C_{UML}(\boldsymbol{\theta}, \mathbf{N}(\omega))$ is equivalent to the minimization of $\tilde{C}_{UML}(\boldsymbol{\theta}, \mathbf{N}(\omega)) = (C_{UML}(\boldsymbol{\theta}, \mathbf{N}(\omega)))^{\frac{1}{M-N}}$, consequently we will study

$$\tilde{C}_{UML}(\boldsymbol{\theta}, \mathbf{N}(\omega)) = \left|\mathbf{E_s}^H(\boldsymbol{\theta})\widehat{\boldsymbol{\Sigma}}_{\mathbf{x}}(\boldsymbol{\theta}, \mathbf{N}(\omega))\mathbf{E_s}(\boldsymbol{\theta})\right|^{\frac{1}{M-N}} C_{CML}(\boldsymbol{\theta}, \mathbf{N}(\omega)). \tag{8}$$

The right hand side of equation (8) is the product of two terms. Let us set

$$\alpha(\boldsymbol{\theta}, \mathbf{N}(\omega)) = \left|\mathbf{E_s}^H(\boldsymbol{\theta})\widehat{\boldsymbol{\Sigma}}_{\mathbf{x}}(\boldsymbol{\theta}, \mathbf{N}(\omega))\mathbf{E_s}(\boldsymbol{\theta})\right|^{\frac{1}{M-N}}. \tag{9}$$

A Taylor expansion at order zero around $(\boldsymbol{\theta}_0, \mathbf{0})$ of $\alpha(\boldsymbol{\theta}, \mathbf{N}(\omega))$ leads to

$$\alpha(\boldsymbol{\theta}, \mathbf{N}(\omega)) = \alpha(\boldsymbol{\theta}_0, \mathbf{0}) + o(1), \tag{10}$$

where o denotes the small oh notation and where

$$\alpha(\boldsymbol{\theta}_0, \mathbf{0}) = \left|\mathbf{E_s}^H(\boldsymbol{\theta}_0)\mathbf{A}(\boldsymbol{\theta}_0)\widehat{\boldsymbol{\Sigma}}_{\mathbf{s}}\mathbf{A}^H(\boldsymbol{\theta}_0)\mathbf{E_s}(\boldsymbol{\theta}_0)\right| \neq 0, \tag{11}$$

where $\widehat{\boldsymbol{\Sigma}}_{\mathbf{s}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{s}_t(\omega)\mathbf{s}_t^H(\omega)$. Consequently, the first non-null term of a Taylor expansion of $\alpha(\boldsymbol{\theta}, \mathbf{n})$ around $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ and $\mathbf{N}(\omega) = \mathbf{0}$ is $\alpha(\boldsymbol{\theta}_0, \mathbf{0})$. Concerning the term $C_{CML}(\boldsymbol{\theta}, \mathbf{N}(\omega))$, a Taylor expansion at order two around $(\boldsymbol{\theta}_0, \mathbf{0})$ leads to

$$C_{CML}(\boldsymbol{\theta}, \mathbf{N}(\omega)) = C_{CML}(\boldsymbol{\theta}_0, \mathbf{0}) + \boldsymbol{\Delta}^T\mathbf{G} + \frac{1}{2}\boldsymbol{\Delta}^T\ddot{\mathbf{H}}\boldsymbol{\Delta} + o\left(\|\boldsymbol{\Delta}\|^2\right) \tag{12}$$

where $\|.\|$ denotes the norm, where

$$\boldsymbol{\Delta} = \left[(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T, vec(\mathrm{Re}\{\mathbf{N}(\omega)\})^T, vec(\mathrm{Im}\{\mathbf{N}(\omega)\})^T\right]^T, \tag{13}$$

where $\mathrm{Re}\{\}$ and $\mathrm{Im}\{\}$ denotes the real and imaginary part, respectively, and where vec denotes the vec operator. $\mathbf{G}$ is the gradient of $C_{CML}(\boldsymbol{\theta}, \mathbf{N}(\omega))$ at $(\boldsymbol{\theta}_0, \mathbf{0})$

$$\mathbf{G} = \left[\left(\left.\frac{\partial C_{CML}(\boldsymbol{\theta}, \mathbf{N}(\omega))}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}_0, \mathbf{0}}\right)^T, \left(\left.\frac{\partial C_{CML}(\boldsymbol{\theta}, \mathbf{N}(\omega))}{\partial vec(\mathrm{Re}\{\mathbf{N}(\omega)\})}\right|_{\boldsymbol{\theta}_0, \mathbf{0}}\right)^T, \left(\left.\frac{\partial C_{CML}(\boldsymbol{\theta}, \mathbf{N}(\omega))}{\partial vec(\mathrm{Im}\{\mathbf{N}(\omega)\})}\right|_{\boldsymbol{\theta}_0, \mathbf{0}}\right)^T\right]^T, \tag{14}$$

and $\ddot{\mathbf{H}}$ is the Hessian matrix of $C_{CML}(\boldsymbol{\theta}, \mathbf{N}(\omega))$ at $(\boldsymbol{\theta}_0, \mathbf{0})$,

$$\ddot{\mathbf{H}} = \begin{bmatrix} \left.\frac{\partial^2 C_{CML}(\boldsymbol{\theta},\mathbf{N}(\omega))}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right|_{\boldsymbol{\theta}_0,\mathbf{0}} & \left.\frac{\partial^2 C_{CML}(\boldsymbol{\theta},\mathbf{N}(\omega))}{\partial\boldsymbol{\theta}\partial vec^T(\mathrm{Re}\{\mathbf{N}(\omega)\})}\right|_{\boldsymbol{\theta}_0,\mathbf{0}} & \left.\frac{\partial^2 C_{CML}(\boldsymbol{\theta},\mathbf{N}(\omega))}{\partial\boldsymbol{\theta}\partial vec^T(\mathrm{Im}\{\mathbf{N}(\omega)\})}\right|_{\boldsymbol{\theta}_0,\mathbf{0}} \\ \left.\frac{\partial^2 C_{CML}(\boldsymbol{\theta},\mathbf{N}(\omega))}{\partial\boldsymbol{\theta}\partial vec^T(\mathrm{Re}\{\mathbf{N}(\omega)\})}\right|_{\boldsymbol{\theta}_0,\mathbf{0}} & \left.\frac{\partial^2 C_{CML}(\boldsymbol{\theta},\mathbf{N}(\omega))}{\partial vec(\mathrm{Re}\{\mathbf{N}(\omega)\})\partial vec^T(\mathrm{Re}\{\mathbf{N}(\omega)\})}\right|_{\boldsymbol{\theta}_0,\mathbf{0}} & \left.\frac{\partial^2 C_{CML}(\boldsymbol{\theta},\mathbf{N}(\omega))}{\partial vec(\mathrm{Re}\{\mathbf{N}(\omega)\})\partial vec^T(\mathrm{Im}\{\mathbf{N}(\omega)\})}\right|_{\boldsymbol{\theta}_0,\mathbf{0}} \\ \left.\frac{\partial^2 C_{CML}(\boldsymbol{\theta},\mathbf{N}(\omega))}{\partial\boldsymbol{\theta}\partial vec^T(\mathrm{Im}\{\mathbf{N}(\omega)\})}\right|_{\boldsymbol{\theta}_0,\mathbf{0}} & \left.\frac{\partial^2 C_{CML}(\boldsymbol{\theta},\mathbf{N}(\omega))}{\partial vec(\mathrm{Re}\{\mathbf{N}(\omega)\})\partial vec^T(\mathrm{Im}\{\mathbf{N}(\omega)\})}\right|_{\boldsymbol{\theta}_0,\mathbf{0}} & \left.\frac{\partial^2 C_{CML}(\boldsymbol{\theta},\mathbf{N}(\omega))}{\partial vec^T(\mathrm{Im}\{\mathbf{N}(\omega)\})\partial vec(\mathrm{Im}\{\mathbf{N}(\omega)\})}\right|_{\boldsymbol{\theta}_0,\mathbf{0}} \end{bmatrix}. \tag{15}$$

For $(\boldsymbol{\theta},\mathbf{n}) = (\boldsymbol{\theta}_0, \mathbf{0})$, $C_{CML}(\boldsymbol{\theta}, \mathbf{N}(\omega))$ is minimal and null. Consequently,

$$C_{CML}(\boldsymbol{\theta}_0, \mathbf{0}) = 0 \text{ and } \mathbf{G} = \mathbf{0}, \tag{16}$$

and

$$C_{CML}(\boldsymbol{\theta}, \mathbf{N}(\omega)) = \frac{1}{2}\boldsymbol{\Delta}^T\ddot{\mathbf{H}}\boldsymbol{\Delta} + o\left(\|\boldsymbol{\Delta}\|^2\right). \tag{17}$$

Therefore, the first non-null term of its Taylor expansion around $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ and $\mathbf{N}(\omega) = \mathbf{0}$ is $\frac{1}{2}\boldsymbol{\Delta}^T\ddot{\mathbf{H}}\boldsymbol{\Delta}$. Consequently,

$$\begin{aligned} \tilde{C}_{UML}(\boldsymbol{\theta}, \mathbf{N}(\omega)) &= \frac{1}{2}\alpha(\boldsymbol{\theta}_0, \mathbf{0})\boldsymbol{\Delta}^T\ddot{\mathbf{H}}\boldsymbol{\Delta} + o\left(\|\boldsymbol{\Delta}\|^2\right) \\ &= \alpha(\boldsymbol{\theta}_0, \mathbf{0}) C_{CML}(\boldsymbol{\theta}, \mathbf{N}(\omega)) + o\left(\|\boldsymbol{\Delta}\|^2\right). \end{aligned} \tag{18}$$

Consequently, at high SNR, since $\tilde{C}_{UML}(\boldsymbol{\theta}, \mathbf{N}(\omega))$ is the product of $C_{CML}(\boldsymbol{\theta}, \mathbf{N}(\omega))$ by a non-null constant, both criterions provide the same estimates, concluding the proof. ■

IV. Non-Gaussianity of the UML

In the sequel, concerning source signals, we are in the stochastic model framework of assumption **A1** and we note $\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} C_{UML}(\boldsymbol{\theta})$ the UML estimator. The next theorem establishes the asymptotic distribution of $\widehat{\boldsymbol{\theta}}$ and its non-Gaussianity for any number of sources (the single source case has already been reported in [16])

Theorem 1: Let $\widetilde{\boldsymbol{\theta}} = \frac{1}{\sigma}\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)$. When SNR tends to infinity, $\widetilde{\boldsymbol{\theta}}$ is non-Gaussian and converges in distribution to $\mathbf{C}(\boldsymbol{\theta}_\mathbf{0})\mathbf{y}$, where $\mathbf{y}$ is a $N \times 1$ Gaussian vector with zero mean and covariance matrix $\mathbf{I}_N$ and $\mathbf{C}(\boldsymbol{\theta}_0)$ is any $N \times N$ random matrix independent of vector $\mathbf{y}$, satisfying:

$$\mathbf{C}(\boldsymbol{\theta}_\mathbf{0})\mathbf{C}^T(\boldsymbol{\theta}_0) = \frac{1}{2T}\left(\mathrm{Re}\left\{\mathbf{H}(\boldsymbol{\theta}_0) \odot \hat{\boldsymbol{\Sigma}}_\mathbf{s}^T\right\}\right)^{-1}, \tag{19}$$

where $\odot$ denotes the Hadamard product (element by element product) and where $\mathbf{H}(\boldsymbol{\theta}_0)$ is a $N \times N$ deterministic matrix which contains the information about the DOA's and about the array structure:

$$\mathbf{H}(\boldsymbol{\theta}_0) = \mathbf{D}^H(\boldsymbol{\theta}_0)\boldsymbol{\Pi}_\mathbf{A}^\perp(\boldsymbol{\theta})\mathbf{D}(\boldsymbol{\theta}_0), \tag{20}$$

with

$$\mathbf{D}(\boldsymbol{\theta}_0) = \left[\left.\frac{d\mathbf{a}(\theta)}{d\theta}\right|_{\theta_1}, \left.\frac{d\mathbf{a}(\theta)}{d\theta}\right|_{\theta_2}, \ldots, \left.\frac{d\mathbf{a}(\theta)}{d\theta}\right|_{\theta_N}\right]. \tag{21}$$

Note that in (19) $T\hat{\boldsymbol{\Sigma}}_\mathbf{s}$ is a $N \times N$ random matrix which follows a complex Wishart distribution with T degrees of freedom and parameter matrix the covariance $\boldsymbol{\Sigma}_\mathbf{s}$ of source signals $\mathbf{s}_t$.

Proof: From proposition 1, we consider that $\hat{\boldsymbol{\theta}}$ is obtained by minimization of $C_{CML}(\boldsymbol{\theta})$ given by (2). Thanks to [14], at high SNR, the conditional distribution $f\left(\widetilde{\boldsymbol{\theta}} \mid \mathbf{S}\right)$ is Gaussian with asymptotic covariance given by the conditional Cramér-Rao bound, see [11]:

$$\mathbf{B}_{COND}(\boldsymbol{\theta}_0) = \frac{1}{2T}\left(\mathrm{Re}\left\{\mathbf{H}(\boldsymbol{\theta}_0) \odot \widehat{\boldsymbol{\Sigma}}_{\mathbf{s}}^T\right\}\right)^{-1}. \tag{22}$$

Let us set $\mathbf{B}_{COND}(\boldsymbol{\theta}_0) = \mathbf{C}(\boldsymbol{\theta}_0)\mathbf{C}^T(\boldsymbol{\theta}_0)$. Therefore, the asymptotic (in SNR) conditional distribution $f\left(\widetilde{\boldsymbol{\theta}} \mid \mathbf{S}\right)$ is the same as the distribution of $\mathbf{C}(\boldsymbol{\theta}_0)\mathbf{y}$, where $\mathbf{y}$ is a Gaussian random vector with zero mean and covariance matrix $\mathbf{I}_N$ and where $\mathbf{C}(\boldsymbol{\theta}_0)$ is a deterministic matrix. Consequently, the asymptotic (in SNR) marginal distribution $f\left(\widetilde{\boldsymbol{\theta}}\right)$ is the same as that of $\mathbf{C}(\boldsymbol{\theta}_0)\mathbf{y}$ where $\mathbf{y}$ is a Gaussian random vector with zero mean and covariance matrix $\mathbf{I}_N$ and where $\mathbf{C}(\boldsymbol{\theta}_0)$ becomes a random matrix since, in (19), $T\widehat{\boldsymbol{\Sigma}}_{\mathbf{s}}$ is complex Wishart distributed with T degrees of freedom, and parameter matrix $\boldsymbol{\Sigma}_{\mathbf{s}}$ the source signals covariance. Since $\mathbf{C}(\boldsymbol{\theta}_0)$ becomes a random matrix, the product $\mathbf{C}(\boldsymbol{\theta}_0)\mathbf{y}$ can not be Gaussian which completes the proof. ■

V. Non-Efficiency of the UML Estimator

In order to proof the non-efficiency of the UML estimator, the comparison between the asymptotic covariance of the UML estimator and the Unconditional Cramér-Rao Bound (UCRB) is provided in this section.

A. Asymptotic covariance of $\widetilde{\boldsymbol{\theta}}$

Corollary 1: Let $cov\left(\widetilde{\boldsymbol{\theta}}\right) = \mathbb{E}\left[\widetilde{\boldsymbol{\theta}}\widetilde{\boldsymbol{\theta}}^T\right]$ be the covariance of $\widetilde{\boldsymbol{\theta}}$. Then, from the above section, we have straightforwardly:

$$\begin{aligned}\lim_{\sigma\to 0} cov\left(\widetilde{\boldsymbol{\theta}}\right) &= \mathbb{E}\left[\mathbf{C}\mathbf{y}\mathbf{y}^T\mathbf{C}^T\right] = \mathbb{E}[\mathbf{C}\mathbf{C}^T],\\ &= \frac{1}{2T}\mathbb{E}\left[\left(\mathrm{Re}\left\{\mathbf{H}(\boldsymbol{\theta}_0)\odot\hat{\boldsymbol{\Sigma}}_{\mathbf{s}}^T\right\}\right)^{-1}\right].\end{aligned} \tag{23}$$

B. Performance bound

According to [11], the UCRB can be written as follows:

$$\mathbf{B}_{UCOND}(\boldsymbol{\theta}_0) = \frac{\sigma^2}{2T}\left(\mathrm{Re}\left\{\mathbf{H}(\boldsymbol{\theta}_0)\odot\left(\boldsymbol{\Sigma}_{\mathbf{s}}\mathbf{A}^H(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\mathbf{A}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{\mathbf{s}}\right)^T\right\}\right)^{-1}, \tag{24}$$

where $\boldsymbol{\Sigma}_{\mathbf{x}}$ is the covariance matrix of the observations.

By using relation (3.20) of [11], it is shown that in (24):

$$\mathbf{A}^H(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\mathbf{A}(\boldsymbol{\theta}_0) = \left(\boldsymbol{\Sigma}_{\mathbf{s}} + \sigma^2\left(\mathbf{A}^H(\boldsymbol{\theta}_0)\mathbf{A}(\boldsymbol{\theta}_0)\right)^{-1}\right)^{-1}, \tag{25}$$

which tends to $\boldsymbol{\Sigma}_{\mathbf{s}}^{-1}$ when σ tends to 0. It follows

$$\lim_{\sigma\to 0}\frac{\mathbf{B}_{UCOND}(\boldsymbol{\theta}_0)}{\sigma^2} = \frac{1}{2T}\left(\mathrm{Re}\left\{\mathbf{H}(\boldsymbol{\theta}_0)\odot\boldsymbol{\Sigma}_{\mathbf{s}}^T\right\}\right)^{-1}. \tag{26}$$

C. Non-efficiency of the UML estimator

In order to prove the non-efficiency of the UML estimator for any number of sources, the following theorem will be of interest. Note that this is a matrix extension of the well-known Jensen's inequality. This theorem has been proved in [19] without the equality condition which will be of particular interest here.

Theorem 2: Let $\mathbf{\Theta}$ be a $N \times N$ real positive definite random matrix. Then

$$\mathbb{E}\left[\mathbf{\Theta}^{-1}\right] \geq \left(\mathbb{E}\left[\mathbf{\Theta}\right]\right)^{-1}, \tag{27}$$

with equality if and only if $\mathbf{\Theta}$ is a constant matrix with probability one. Appendix A details the proof.

Corollary 2: Let us set $\mathbf{\Theta} = 2T\,\mathrm{Re}\left\{\mathbf{H}\left(\boldsymbol{\theta}_0\right) \odot \hat{\mathbf{\Sigma}}_{\mathbf{s}}^T\right\}$ in equations (26) and (23). Equation (27) becomes the Cramér-Rao inequality. Since $\mathbf{\Theta}$ is not a constant matrix with probability one, the inequality is strict and the UML estimator is non-efficient for any number of sources.

VI. Specific Performance Study of the UML Estimator for the Two Sources Scenario

This section is devoted to a deeper statistical investigation of two specific cases frequently met in array processing: the single and two sources case. We remind the probability density function (pdf) and the variance closed form of the UML estimates in the single source case tediously obtained in [16]. For two uncorrelated sources and centro-symmetric arrays, we give a closed-form expression of the UML estimates covariance.

A. Distribution and Theoretical Variance in the Single Source Case

In the single source case, $\mathbf{\Sigma_s} = \Sigma_1$ and $\mathbf{H} = h_1$. Then $\tilde{\theta}$ is asymptotically distributed as $\sqrt{k}S_{2T}$ where S_{2T} is a Student random variable with $2T$ degrees of freedom and k is given by:

$$k = \lim_{\sigma \to 0} \frac{\mathbf{B}_{UCOND}\left(\boldsymbol{\theta}_0\right)}{\sigma^2} = \frac{1}{2Th_1\Sigma_1}. \tag{28}$$

The asymptotic variance of $\tilde{\theta}$ is then given by:

$$var(\tilde{\theta}) = \frac{T}{T-1}k. \tag{29}$$

As established in theorem 2, for finite T, the UML estimator is not asymptotically efficient since $\frac{T}{T-1} > 1$.

B. Theoretical Variance in the Two Sources Case for Uncorrelated Sources and Centro-symmetric Array

Most arrays met in practice possess a center of symmetry (this is for instance the case of the ULA.). Under this condition which will be assumed in the following, the matrix $\mathbf{H}$ of equation (20) is real and symmetric (see appendix B):

$$\mathbf{H}\left(\boldsymbol{\theta}_0\right) = \begin{pmatrix} h_1 & h_3 \\ h_3 & h_2 \end{pmatrix}. \tag{30}$$

For two uncorrelated sources, $\mathbf{\Sigma_s} = Diag\left\{\Sigma_1, \Sigma_2\right\}$ and the asymptotic covariance of $\tilde{\boldsymbol{\theta}}$ is given by

$$\begin{aligned}\lim_{\sigma\to 0} cov\left(\tilde{\boldsymbol{\theta}}\right) &= \frac{1}{2}\frac{{}_2F_1\left(1,1;2T;\frac{h_3^2}{h_1h_2}\right)}{T-1} Diag\left\{\frac{1}{h_1\Sigma_1},\frac{1}{h_2\Sigma_2}\right\}, \\ &= \frac{T}{T-1}\,{}_2F_1\left(1,1;2T;\frac{h_3^2}{h_1h_2}\right)\mathbf{K}, \end{aligned} \tag{31}$$

where $\mathbf{K} = \lim\limits_{\sigma\to 0}\frac{\mathbf{B}_{UCOND}(\boldsymbol{\theta}_0)}{\sigma^2}$ and where ${}_2F_1\left(a,b;c;\omega\right)$ is the Gauss hypergeometric function defined by its integral representation ([20] pp. 558)

$${}_2F_1\left(a,b;c;\omega\right) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 z^{b-1}\left(1-z\right)^{c-b-1}\left(1-z\omega\right)^{-a}dz, \tag{32}$$

where $\Gamma\left(z\right)$ denotes the Gamma function $\Gamma\left(z\right) = \int_0^\infty t^{z-1}e^{-t}dt$.

The derivation of Eqn. (31) is given in appendix C. As expected, the UML estimator is not asymptotically efficient since $\frac{T}{T-1} > 1$ and ${}_2F_1\left(1,1;2T;\frac{h_3^2}{h_1h_2}\right) \geq 1$.

VII. Simulation Examples

In this section, results of some Monte Carlo simulations concerning the UML estimator are presented. The purpose is to illustrate the applicability of the derived expressions of the pdf and of the variance. In all simulations, the array is an ULA of $M = 10$ sensors with half-wavelength spacing (the beamwidth of the array is equal to 10 degrees). The UML DOA estimation is conducted with $T = 2$ snapshots. We consider the case of two uncorrelated sources with equal power located at 0 degrees and 5 degrees. DOA are given with respect to the broadside. The ML DOA estimation is performed with a Gauss Newton algorithm thanks to a global search over a grid.

We have reported in Figure 1 the evolution of the UML empirical variance, of the theoretical variance (31), and of the UCRB versus SNR. Monte Carlo simulations have been performed with $r = 1000$ independent realizations. Here $\frac{T}{T-1}\,{}_2F_1\left(1,1;2T;\frac{h_3^2}{h_1h_2}\right) = 2.9$. In this asymptotic region, one can notice the good match between theoretical results and simulations. The non-efficiency of UML at high SNR is observed. We also observe the well known threshold effect [21] of the estimator variance when the SNR becomes weak (approximatively 20 dB in this case). This phenomena due to outliers gives the validity domain this asymptotic analysis (see [15] for more details concerning the UML threshold prediction).

Figure 2 gives the histograms of the estimated DOA corresponding to the previous case with Monte Carlo simulations performed with $r = 10000$ independent realizations and a SNR of 30 dB. We also reported the pdf of a Gaussian distribution with the same variance. The non-Gaussianity of the UML estimates is observed. To confirm this ”visual” result, we have used the classical Chi-square test which tests a distribution observed against another theoretical distribution. For the Chi-square fit computation, the data are divided into $k = 15$ bins and the statistical test requires the computation of

$$\Delta = \sum_{i=1}^{k}\frac{\left(O_i - rp_i\right)^2}{rp_i}, \tag{33}$$

where O_i is the observed frequency for bin i and p_i is the candidate probability for bin i. The hypothesis that the data are from a population following the candidate distribution is rejected if

$$\Pr(X \geq \Delta) = \frac{\Gamma\left(\frac{k-1}{2}, \frac{\Delta}{2}\right)}{\Gamma\left(\frac{k-1}{2}\right)} \leq 5\%, \tag{34}$$

where X follows a Chi-square distribution with $k-1$ degrees of freedom. Table I shows that the pdf of the estimates is not Gaussian for a SNR of 30 dB.

VIII. Conclusion

The statistical properties of the UML estimator have been investigated. We have shown that, for any number of sources, this estimator is non-Gaussian and non-efficient at high Signal to Noise Ratio for a finite number of samples. The key point of the analysis is the equivalence between the UML and the CML method at high SNR. Moreover, we have provided the UML estimator covariance closed-form expression for two uncorrelated sources and centro-symmetric array.

Appendix

A. Proof of the UML Non-Efficiency

Lemma 1: Let $\mathbf{\Omega}$ be a $N \times N$ real symmetric positive semidefinite matrix. Then $\forall \mathbf{q}$:

$$\mathbf{q}^T \mathbf{\Omega} \mathbf{q} + \mathbf{q}^T \mathbf{\Omega}^{-1} \mathbf{q} - 2\mathbf{q}^T \mathbf{q} \geq 0, \tag{35}$$

with equality if and only if $\mathbf{q}$ is an eigenvector of $\mathbf{\Omega}$ with eigenvalue one.

Proof: Let us set $\mathbf{\Omega} = \sum_{i=1}^{N} \lambda_i \mathbf{r}_i \mathbf{r}_i^T$ the eigendecomposition of $\mathbf{\Omega}$ on an orthonormal basis $\{\mathbf{r}_i\}_{i=1..N}$ with associated eigenvalues λ_i.Equation (35) can be written:

$$\sum_{i=1}^{N} \left(\lambda_i - 2 + \frac{1}{\lambda_i}\right) \left(\mathbf{q}^T \mathbf{r}_i\right)^2 \geq 0. \tag{36}$$

Noticing that $\lambda_i - 2 + \frac{1}{\lambda_i} \geq 0$ for $\lambda_i > 0$, and that $\lambda_i - 2 + \frac{1}{\lambda_i} = 0$ for $\lambda_i = 1$ the proof is straightforward. ■

Lemma 2: Let $\mathbf{\Omega}$ a $N \times N$ random real symmetric positive semidefinite matrix with probability one such that $\mathbb{E}[\mathbf{\Omega}] = \mathbf{I}_N$. Then there is a vector $\mathbf{q}$ such that

$$\mathbf{q}^T \mathbb{E}\left[\mathbf{\Omega}^{-1}\right] \mathbf{q} - \mathbf{q}^T \mathbf{q} > 0, \tag{37}$$

if and only if $\Pr[\mathbf{\Omega} = \mathbf{I}_N] \neq 1$.

Proof: Let us set $\zeta_{\mathbf{q}} = \mathbf{q}^T \mathbf{\Omega} \mathbf{q} + \mathbf{q}^T \mathbf{\Omega}^{-1} \mathbf{q} - 2\mathbf{q}^T \mathbf{q}$. Since $\mathbb{E}[\mathbf{\Omega}] = \mathbf{I}_N$, we have $\mathbb{E}\left[\zeta_{\mathbf{q}}\right] = \mathbf{q}^T \mathbb{E}\left[\mathbf{\Omega}^{-1}\right] \mathbf{q} - \mathbf{q}^T \mathbf{q}$. Consequently, proving lemma 2 amounts to prove that $\exists \mathbf{q}$ such that $\mathbb{E}\left[\zeta_{\mathbf{q}}\right] > 0$ if and only if $\Pr[\mathbf{\Omega} = \mathbf{I}_N] \neq 1$. From (35) $\zeta_{\mathbf{q}}$ is a nonnegative random variable. Thus, $\mathbb{E}\left[\zeta_{\mathbf{q}}\right] > 0$ if and only if $\Pr\left[\zeta_{\mathbf{q}} = 0\right] \neq 1$. From lemma 1, $\Pr\left[\zeta_{\mathbf{q}} = 0\right] = \Pr[\mathbf{\Omega} \mathbf{q} = \mathbf{q}]$. Consequently, $\forall \mathbf{q} \Pr\left[\zeta_{\mathbf{q}} = 0\right] = 1 \Longleftrightarrow \forall \mathbf{q} \Pr[\mathbf{\Omega} \mathbf{q} = \mathbf{q}] = 1 \Longleftrightarrow \Pr[\mathbf{\Omega} = \mathbf{I}_N] = 1$.

This completes the proof. ■

Finally, with the notations of theorem 2, let us set $\mathbf{\Omega} = \mathbb{E}[\mathbf{\Theta}]^{-1/2} \mathbf{\Theta} \mathbb{E}[\mathbf{\Theta}]^{-1/2}$. Theorem 2 follows from lemma 2.

B. *Study of matrix* $\mathbf{H}$ *centro-symmetric sensor arrays*

We prove in this appendix that $\mathbf{H}$ (see equation (20)) is a real symmetric matrix. It is obvious that $\mathbf{H}$ is an hermitian matrix. Therefore we must prove that $\mathbf{H}$ is a real matrix under the assumption that the array has a center of symmetry.

$$\begin{aligned}\mathbf{H}(\boldsymbol{\theta}_0) &= \mathbf{D}^H(\boldsymbol{\theta}_0)\,\boldsymbol{\Pi}_{\mathbf{A}}^{\perp}(\boldsymbol{\theta})\,\mathbf{D}(\boldsymbol{\theta}_0)\\ &= \mathbf{D}^H(\boldsymbol{\theta}_0)\,\mathbf{D}(\boldsymbol{\theta}_0) - \left(\mathbf{A}^H(\boldsymbol{\theta}_0)\,\mathbf{D}(\boldsymbol{\theta}_0)\right)^H\left(\mathbf{A}^H(\boldsymbol{\theta}_0)\,\mathbf{A}(\boldsymbol{\theta}_0)\right)^{-1}\left(\mathbf{A}^H(\boldsymbol{\theta}_0)\,\mathbf{D}(\boldsymbol{\theta}_0)\right).\end{aligned} \tag{38}$$

The i-th element of the steering vector is[1]:

$$a_i(\theta_k) = e^{j\frac{2\pi}{\lambda}\mathbf{v}_i^T\mathbf{u}(\theta_k)}, \tag{39}$$

where $\mathbf{v}_i$ is the coordinate vector of the i-th sensor and $\mathbf{u}(\theta_k)$ is the unit vector pointing towards the k-th source. Therefore the i-th element of $\mathbf{d}(\theta_k) = \left.\frac{d\mathbf{a}(\theta)}{d\theta}\right|_{\theta_k}$ (see (21)) is:

$$d_i(\theta_k) = j\frac{2\pi}{\lambda}\mathbf{v}_i^T \underbrace{\left.\frac{d\mathbf{u}}{d\theta}\right|_{\theta_k}}_{\dot{\mathbf{u}}(\theta_k)} e^{j\frac{2\pi}{\lambda}\mathbf{v}_i^T\mathbf{u}(\theta_k)}. \tag{40}$$

If the array has a center of symmetry, the sensors can be labelled so that $\mathbf{v}_i = -\mathbf{v}_{M-i+1}$. The m-th row and n-th column element of each term in (38) is

$$\begin{cases} \left.\mathbf{A}^H(\boldsymbol{\theta}_0)\,\mathbf{A}(\boldsymbol{\theta}_0)\right|_{m,n} = \sum_{i=1}^{M} e^{j\frac{2\pi}{\lambda}\mathbf{v}_i^T(\mathbf{u}(\theta_n)-\mathbf{u}(\theta_m))}, \\ \left.\mathbf{D}^H(\boldsymbol{\theta}_0)\,\mathbf{D}(\boldsymbol{\theta}_0)\right|_{m,n} = \sum_{i=1}^{M} \left(\frac{2\pi}{\lambda}\right)^2\left(\mathbf{v}_i^T\dot{\mathbf{u}}(\theta_m)\right)\left(\mathbf{v}_i^T\dot{\mathbf{u}}(\theta_n)\right)e^{j\frac{2\pi}{\lambda}\mathbf{v}_i^T(\mathbf{u}(\theta_n)-\mathbf{u}(\theta_m))}, \\ \left.\mathbf{A}^H(\boldsymbol{\theta}_0)\,\mathbf{D}(\boldsymbol{\theta}_0)\right|_{m,n} = \sum_{i=1}^{M} j\frac{2\pi}{\lambda}\left(\mathbf{v}_i^T\dot{\mathbf{u}}(\theta_n)\right)e^{j\frac{2\pi}{\lambda}\mathbf{v}_i^T(\mathbf{u}(\theta_n)-\mathbf{u}(\theta_m))}. \end{cases} \tag{41}$$

$\left.\mathbf{A}^H(\boldsymbol{\theta}_0)\,\mathbf{A}(\boldsymbol{\theta}_0)\right|_{m,n}$ is a sum of two by two complex conjugates with the same magnitude[2]. Therefore, $\left.\mathbf{A}^H(\boldsymbol{\theta}_0)\,\mathbf{A}(\boldsymbol{\theta}_0)\right|_{m,n} \in \mathbb{R}$. Similarly, $\left.\mathbf{D}^H(\boldsymbol{\theta}_0)\,\mathbf{D}(\boldsymbol{\theta}_0)\right|_{m,n} \in \mathbb{R}$ since $\left(\mathbf{v}_i^T\dot{\mathbf{u}}(\theta_m)\right)\left(\mathbf{v}_i^T\dot{\mathbf{u}}(\theta_n)\right) = \left(\mathbf{v}_{M-i+1}^T\dot{\mathbf{u}}(\theta_m)\right)\left(\mathbf{v}_{M-i+1}^T\dot{\mathbf{u}}(\theta_n)\right)$. And $\left.\mathbf{A}^H(\boldsymbol{\theta}_0)\,\mathbf{D}(\boldsymbol{\theta}_0)\right|_{m,n} \in \mathbb{R}$ since:

$$\left.\mathbf{A}^H(\boldsymbol{\theta}_0)\,\mathbf{D}(\boldsymbol{\theta}_0)\right|_{m,n} = j\frac{2\pi}{\lambda}\sum_{i=1}^{\frac{M}{2}}\left(\mathbf{v}_i^T\dot{\mathbf{u}}(\theta_n)\right)\underbrace{\left(e^{j\frac{2\pi}{\lambda}\mathbf{v}_i^T(\mathbf{u}(\theta_n)-\mathbf{u}(\theta_m))} - e^{-j\frac{2\pi}{\lambda}\mathbf{v}_i^T(\mathbf{u}(\theta_n)-\mathbf{u}(\theta_m))}\right)}_{\text{imaginary number}}. \tag{42}$$

Therefore $\mathbf{H}$ is a real symmetric matrix.

[1] λ is the wave-length of emitted signals.

[2] If the number of sensors is odd $\mathbf{v}_{\frac{M}{2}+1} = 0$.

C. Theoretical Variance in the Two Sources Case

According to (23), and with the assumption that the array has a center of symmetry, i.e. $\mathbf{H}(\boldsymbol{\theta}_0)$ becomes a real symmetric matrix:

$$\begin{aligned}\lim_{\sigma\to 0} cov\left(\widetilde{\boldsymbol{\theta}}\right) &= \frac{1}{2T}\mathbb{E}\left[\left(\mathbf{H}(\boldsymbol{\theta}_0)\odot \mathrm{Re}\left(\hat{\boldsymbol{\Sigma}}_{\mathbf{s}}^T\right)\right)^{-1}\right] \\ &= \frac{1}{2}\mathbb{E}\left[\left(\mathbf{H}(\boldsymbol{\theta}_0)\odot \mathrm{Re}(\mathbf{W})\right)^{-1}\right], \end{aligned} \tag{43}$$

where $\mathbf{W}$ is a $N \times N$ random matrix which follows a complex Wishart distribution with T degrees of freedom and parameter matrix the covariance, $\boldsymbol{\Sigma}_{\mathbf{s}} = Diag\{\mathit{\Sigma}_1, \mathit{\Sigma}_2\}$, of source signals $\mathbf{s}_t$.

Under assumptions **A1** and uncorrelated sources, $\mathbf{W}_R = \mathrm{Re}\{\mathbf{W}\}$ is a $N\times N$ symmetric positive definite random matrix which follows a real Wishart distribution with $2T$ degrees of freedom and parameter matrix the covariance $\frac{1}{2}\boldsymbol{\Sigma}_{\mathbf{s}}$. From the Cholesky factorization, $\mathbf{W}_R = \mathbf{D}\mathbf{D}^T$, with:

$$\mathbf{D} = \begin{pmatrix} \rho_1 & 0 \\ \alpha & \rho_2 \end{pmatrix}. \tag{44}$$

The elements of $\mathbf{D}$ are independent and satisfy [22]:

$$\begin{cases} \rho_1 \sim \sqrt{\frac{\mathit{\Sigma}_1}{2}\chi^2(2T)}, \\ \rho_2 \sim \sqrt{\frac{\mathit{\Sigma}_2}{2}\chi^2(2T-1)}, \\ \alpha \sim \mathcal{N}\left(0, \frac{\mathit{\Sigma}_2}{2}\right), \end{cases} \tag{45}$$

where $\mathcal{N}(0,\varepsilon)$ is a Gaussian distribution with mean value 0 and variance ε and where $\chi^2(P)$ is a Chi-square distribution with P degrees of freedom.

The covariance of $\widetilde{\boldsymbol{\theta}}$ is given by:

$$\begin{aligned}\lim_{\sigma\to 0} cov\left(\widetilde{\boldsymbol{\theta}}\right) &= \begin{pmatrix} var(\tilde{\theta}_1) & \Psi \\ \Psi & var(\tilde{\theta}_2) \end{pmatrix} \qquad (46) \\ &= \frac{1}{2}\mathbb{E}\left[\left(\mathbf{H}(\boldsymbol{\theta}_0)\odot \mathbf{W}_R\right)^{-1}\right] \\ &= \frac{1}{2}\mathbb{E}\left[\frac{1}{\Phi}\begin{pmatrix} h_2\left(\rho_2^2+\alpha^2\right) & -h_3\rho_1\alpha \\ -h_3\rho_1\alpha & h_1\rho_1^2 \end{pmatrix}\right], \qquad (47)\end{aligned}$$

where $var(\tilde{\theta}_1)$ (respectively $var(\tilde{\theta}_2)$) is the variance of the first source (respectively the second source), Ψ is the cross-correlation and $\Phi = h_1 h_2 \rho_1^2\left(\rho_2^2+\alpha^2\right) - \left(h_3\rho_1\alpha\right)^2$.

From (47),

$$\begin{aligned} var(\tilde{\theta}_1) &= \frac{1}{2}\mathbb{E}\left[\frac{h_2\left(\rho_2^2+\alpha^2\right)}{h_1 h_2 \rho_1^2\left(\rho_2^2+\alpha^2\right) - \left(h_3\rho_1\alpha\right)^2}\right] \\ &= \frac{1}{2h_1}\mathbb{E}\left[\frac{1}{\rho_1^2}\frac{1}{1-\frac{h_3^2}{h_1h_2}\frac{\alpha^2}{\alpha^2+\rho_2^2}}\right], \end{aligned} \tag{48}$$

where $\alpha^2 \sim \frac{\Sigma_2}{2}\chi^2(1)$ and the ratio $\frac{\alpha^2}{\alpha^2+\rho_2^2} = Z$ follows a beta distribution with 1 and $2T-1$ degrees of freedom which is independent of $Y = \rho_1^2$. Therefore, equation (48) becomes:

$$var(\tilde{\theta}_1) = \frac{1}{2h_1}\mathbb{E}\left[\frac{1}{Y}\right]\mathbb{E}\left[\frac{1}{1-\frac{h_3^2}{h_1h_2}Z}\right] = \frac{I_1 I_2}{2h_1}. \tag{49}$$

$I_1 = \mathbb{E}\left[\frac{1}{Y}\right]$ and $I_2 = \mathbb{E}\left[\frac{1}{1-\frac{h_3^2}{h_1h_2}Z}\right]$ satisfy:

$$\begin{cases} I_1 = \int_0^\infty \frac{1}{y}\Pi_{\chi^2}(y)\,dy, \\ I_2 = \int_0^1 \frac{1}{1-\frac{h_3^2}{h_1h_2}z}\Pi_\beta(z)\,dz, \end{cases} \tag{50}$$

where $\Pi_{\chi^2}(y)$ and $\Pi_\beta(z)$ are respectively the probability density functions of a chi-square random variable $\frac{\Sigma_1}{2}\chi^2(2T)$ and of a beta random variable $\beta(1, 2T-1)$:

$$\begin{cases} \Pi_{\chi^2}(y) = \frac{1}{2^{T-1}\Gamma(T)\Sigma_1^T}y^{T-1}e^{-\frac{y}{\Sigma_1}}, \\ \Pi_\beta(z) = (2T-1)(1-z)^{2(T-1)}. \end{cases} \tag{51}$$

When $T \geq 2$, I_1 converges: it is a Gamma function. I_2 is the integral representation of a Gauss hypergeometric function ([20] pp. 556-565)

$${}_2F_1(a,b;c;\omega) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 z^{b-1}(1-z)^{c-b-1}(1-z\omega)^{-a}dz, \tag{52}$$

where $a=1$, $b=1$, $c=2T$ and $\omega = \frac{h_3^2}{h_1h_2}$. Note that (52) is finite:

$$\begin{cases} \text{for all } (a,b,c) \text{ if } -1<\omega<1, \\ \text{for } c>a+b \text{ if } \omega = \pm 1. \end{cases} \tag{53}$$

In our case, $\mathbf{H}$ is a semi positive definite matrix, then $|\mathbf{H}| \geq 0 \Leftrightarrow \omega = \frac{h_3^2}{h_1h_2} \leq 1$. It signifies that I_2 is finite for $T \geq 2$.

Finally:

$$\begin{cases} I_1 = \frac{1}{(T-1)\Sigma_1}, \\ I_2 = {}_2F_1\left(1,1;2T;\frac{h_3^2}{h_1h_2}\right). \end{cases} \tag{54}$$

and

$$var(\tilde{\theta}_1) = \frac{{}_2F_1\left(1,1;2T;\frac{h_3^2}{h_1h_2}\right)}{2(T-1)h_1\Sigma_1}. \tag{55}$$

Similarly:

$$var(\tilde{\theta}_2) = \frac{{}_2F_1\left(1,1;2T;\frac{h_3^2}{h_1h_2}\right)}{2(T-1)h_2\Sigma_2}. \tag{56}$$

It can be easily shown that $\Psi = 0$ (see equation (46)): it is the integral from minus infinity to plus infinity of an odd function of the variable α.

According to (26) the UCRB in the two sources case is:

$$\lim_{\sigma\to 0}\frac{1}{\sigma^2}\mathbf{B}_{UCOND}(\boldsymbol{\theta}_0) = \frac{1}{2T}Diag\left\{\frac{1}{h_1\Sigma_1}, \frac{1}{h_2\Sigma_2}\right\}. \tag{57}$$

Therefore, using (46), (55), (56) and (57) one obtain (31).

REFERENCES

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

[2] R. Roy, A. Paulraj, and T. Kailath, "ESPRIT a subspace rotation approach to estimation of parameters of cisoids in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 1340–1342, 1986.

[3] R. Kumaresan and D. W. Tufts, "Estimating the angles of arrival of multiple plane waves," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 19, no. 1, pp. 134–138, 1983.

[4] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Trans. Signal Processing*, vol. 39, pp. 1110–1121, May 1991.

[5] J. F. Böhme, "Estimation of source parameters by maximum likelihood and non linear regression," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 731–734, 1984.

[6] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood and the Cramer Rao bound," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 720–741, May 1989.

[7] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood and the Cramer Rao bound: further results and comparisons," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 2140–2150, 1990.

[8] J. F. Böhme, "Estimation of spectral parameters of correlated signals in wavefields," *Signal Processing*, vol. 10, pp. 329–337, 1986.

[9] 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.), ch. 4, pp. 99–151, Berlin: Springer-Verlag, 1993.

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

[11] P. Stoica and A. Nehorai, "Performances study of conditional and unconditional direction of arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1783–1795, Oct. 1990.

[12] B. Ottersten, M. Viberg, and T. Kailath, "Analysis of subspace fitting and ML techniques for parameter estimation from sensors array data," *IEEE Trans. Signal Processing*, vol. 40, pp. 590–599, Mar. 1992.

[13] S. M. Kay, *Fundamentals of Statistical Signal Processing*, vol. 1. NJ: Prentice Hall, 1993.

[14] A. Renaux, P. Forster, E. Chaumette, and P. Larzabal, "On the high SNR conditional maximum-likelihood estimator full statistical characterization," *to appear in IEEE Trans. on Signal Processing*, 2006.

[15] F. Athley, "Threshold region performance of Maximum Likelihood direction of arrival estimators," *IEEE Trans. Signal Processing*, vol. 53, pp. 1359–1373, Apr. 2005.

[16] P. Forster, E. Boyer, and P. Larzabal, "Nonefficiency of stochastic beamforming bearing estimates at high SNR and finite number of samples," *IEEE Signal Processing Lett.*, vol. 11, pp. 509–512, May 2004.

[17] 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*, vol. 2, (Montreal, QC), pp. 121–124, May 2004.

[18] A. Renaux, P. Forster, and E. Boyer, "Asymptotic non efficiency of a maximum likelihood estimator at finite number of samples," in *Proceedings EUSIPCO*, (Vienna, AT), Sept. 2004.

[19] B. Z. Bobrovsky, E. Mayer-Wolf, and M. Zakai, "Some classes of global Cramer Rao bounds," *Ann. Statistics*, vol. 15, pp. 1421–1438, 1987.

[20] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. New York: Dover Publications, inc., 1972.

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

[22] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. New York: Wiley, 1984.

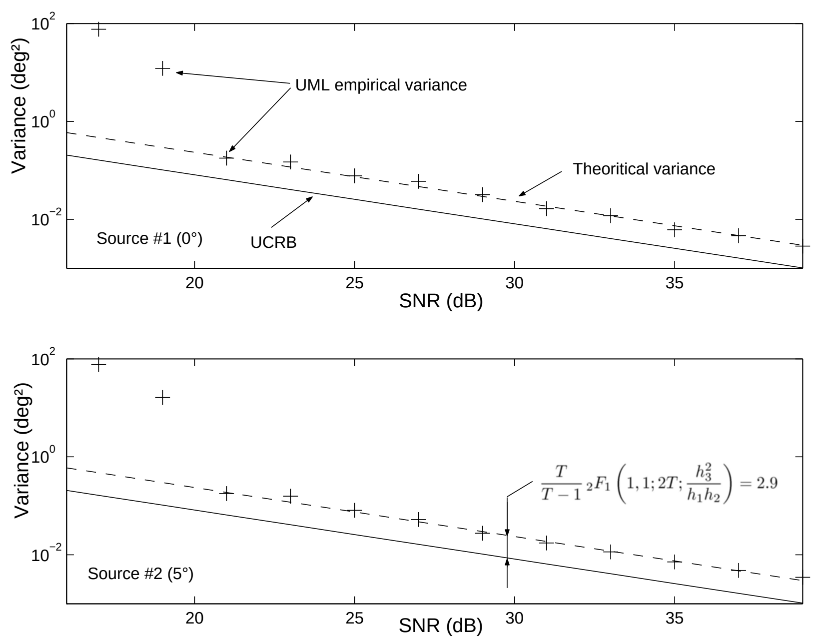

Fig. 1. Asymptotic variance of the UML estimator in the two sources case. $\boldsymbol{\theta}_0 = [0°, 5°]$, $T = 2$ snapshots and $M = 10$ sensors.

Two sources case	Gaussian pdf: source 1	Gaussian pdf: source 2
Δ	$1.32\ 10^{22}$	$4\ 10^{24}$
$\Pr(X \geq \Delta)$	0%	0%
Hypothesis	rejected	rejected

TABLE I

CHI-SQUARE TEST IN THE TWO SOURCES CASE. $k = 15$ BINS, $r = 10000$ REALIZATIONS. $\boldsymbol{\theta_0} = [0°, 5°]$, $M = 10$ SENSORS, $T = 2$ SNAPSHOTS, AND SNR$= 30$ DECIBEL.

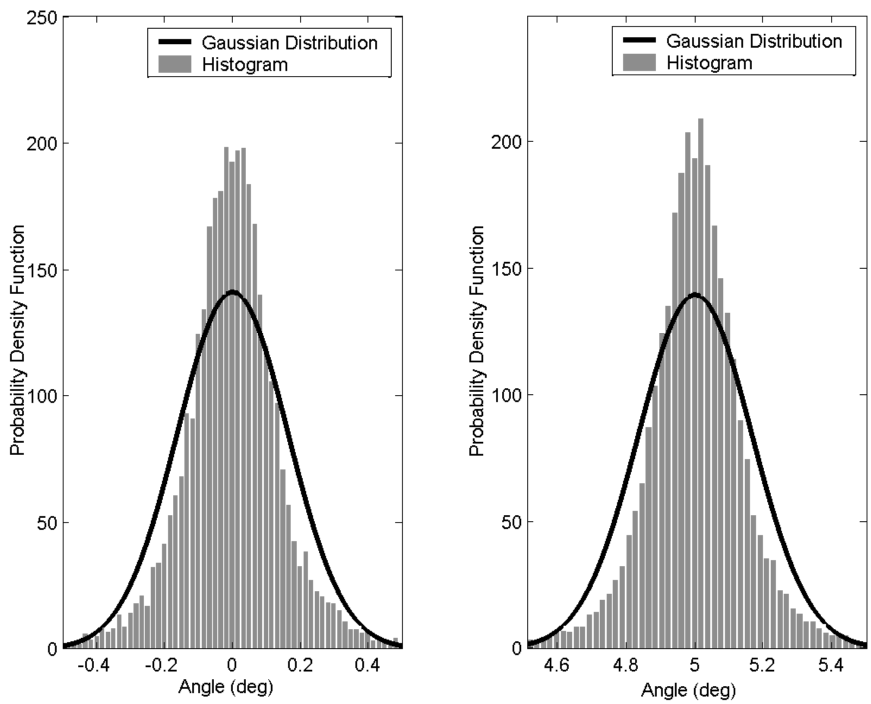

Fig. 2. Histogram of UML estimates in the two sources case. $\boldsymbol{\theta_0} = [0°, 5°]$, $M = 10$ sensors, $T = 2$ snapshots, and SNR$= 30$ dB.