ON THE RESOLUTION PROBABILITY OF CONDITIONAL AND UNCONDITIONAL MAXIMUM LIKELIHOOD DOA ESTIMATION

Xavier Mestre[1], *Pascal Vallet*[2], *Philippe Loubaton*[3],

[1]Centre Tecnològic de Telecomunicacions de Catalunya, Castelldefels, Barcelona (Spain)
[2]Institut Polytechnique de Bordeaux (ENSEIRB-MATMECA), Bordeux (France)
[3]IGM Université Paris-Est Marne-la-Vallée, Paris (France)

ABSTRACT

The outlier production mechanism of maximum likelihood direction-of-arrival estimators is investigated. The objective is to provide an accurate description of the probability of resolution for both conditional and unconditional maximum likelihood methods in the small sample size regime. To that effect, the asymptotic behavior of these two cost functions is analyzed assuming that both the number of antennas and the number of available snapshots increase without bound at the same rate, so that both quantities are comparable in magnitude. The finite dimensional distributions of both conditional and unconditional cost functions are shown to be Gaussian in this asymptotic regime, and a closed form expression of the corresponding asymptotic covariance matrices is provided.

Index Terms— Conditional ML, Unconditional ML, DoA estimation, random matrix theory, central limit theorem.

1. INTRODUCTION

Among all direction of arrival (DoA) estimation techniques, Maximum Likelihood (ML) methods offer the highest advantage in terms of both precision and spatial resolution. Unfortunately, ML algorithms typically require high dimensional search techniques that cause a significant increase in the required computational complexity, especially when compared to other simpler algorithms based on one-dimensional search (such as subspace-based approaches). Still, ML methods continue to be the most attractive solution in applications based on offline processing or in architectural solutions where computational resources are not scarce.

Let $\mathbf{y}(n)$ denote an $M \times 1$ complex vector containing the samples received by the M antenna array at the nth sampling instant, where $n = 1, \ldots, N$ and N denotes the number of available snapshots. For simplicity, we will consider here the case where $N > M$, i.e. the sample size is higher than the observation dimension. Assume that the array is receiving the signals transmitted by K sources from angles of arrival $\bar{\theta} = \left[\bar{\theta}(1), \ldots, \bar{\theta}(K)\right]^T$, so that we can model $\mathbf{y}(n)$ as

$$\mathbf{y}(n) = \mathbf{A}(\bar{\theta})\mathbf{s}(n) + \mathbf{n}(n)$$

where $\mathbf{s}(n)$ contains the source signals at the nth sampling instant, $\mathbf{A}(\bar{\theta})$ is a $M \times K$ matrix containing as columns the steering vectors associated with the true DoAs of the $K < M$ sources, and $\mathbf{n}(n)$ are noise samples, which will be assumed to be independent and identically distributed (i.i.d.) Gaussian circularly symmetric random variables with zero mean and variance σ^2.

Traditionally, two different ML methods for DoA estimation have coexisted in the literature, each one derived from a different modelling assumption regarding the nature of the source signals $\mathbf{s}(n)$: the "conditional" (or deterministic) model, which assumes that the received signals are deterministic unknowns, and the "unconditional" (or stochastic) model, which assumes that the source signals are random variables [1]. In this last case, the column vectors $\mathbf{s}(n)$ are typically assumed to be i.i.d. in the time domain, following a circularly symmetric Gaussian distribution with zero mean and covariance $\mathbf{P}_s = \mathbb{E}\left[\mathbf{s}(n)\mathbf{s}^H(n)\right]$. For each of these two models, one can derive a different Maximum Likelihood estimator. From now on, these two methods will be referred to as conditional ML (CML) and unconditional ML (UML) estimators, respectively. In both cases, the estimated angles are determined as $\hat{\theta} = \arg\min_{\theta \in \Theta} \hat{\eta}\left(\theta\right)$ where Θ is a domain for which $\mathbf{A}(\theta)$ has full column rank and where $\hat{\eta}\left(\theta\right)$ is a certain random function that depends on the ML method. More spe cifically, this random function can be written as

$$\hat{\eta}_{CML}\left(\theta\right) = \frac{1}{M}\operatorname{tr}\left[\mathbf{P}_A^{\perp}(\theta)\hat{\mathbf{R}}\right] \qquad (1)$$

for the CML method and

$$\hat{\eta}_{UML}\left(\theta\right) = \frac{1}{M}\log\det\left[\hat{\eta}_{CML}\left(\theta\right)\mathbf{P}_A^{\perp}(\theta) + \mathbf{P}_A(\theta)\hat{\mathbf{R}}\mathbf{P}_A(\theta)\right] \qquad (2)$$

for the UML method [2, 3], where we have defined the orthogonal projection matrices

$$\mathbf{P}_A(\theta) = \mathbf{A}(\theta)\left(\mathbf{A}^H(\theta)\mathbf{A}(\theta)\right)^{-1}\mathbf{A}^H(\theta)$$

This work was partially supported by the Catalan and Spanish grants 2009 SGR 1046 and TEC2011-29006-C03-01, as well as the French Program ANR-07 MDCO-012-01 "SESAME".

and $\mathbf{P}_A^{\perp}(\theta) = \mathbf{I}_M - \mathbf{P}_A(\theta)$, and where $\hat{\mathbf{R}}$ is the sample covariance matrix, i.e. $\hat{\mathbf{R}} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{y}(n)\mathbf{y}^H(n)$.

In [1], Stoica and Nehorai provided a comparative study of the performance of the above two methods in the large sample size regime, showing that UML asymptotically outperforms CML as $N \to \infty$. However, these conclusions only hold for very large N in comparison with M, and do not carry over to the more practical situation where M, N are comparable in magnitude. It should be pointed out that it is precisely in this regime that the so-called breakdown effect (characterized by the systematic presence of outliers in the DoA estimates) can be effectively observed. Furthermore, it is indeed in this "threshold region" that ML methods become an interesting alternative to one-dimensional search approaches (such as subspace based techniques), since when $N \gg M$ all these methods perform very similarly.

Unfortunately, the behavior of these ML approaches in the regime where M, N are comparable in magnitude is still not fully understood, mainly due to the complicated dependence of $\hat{\eta}_{UML}(\theta)$ on the observations. Some interesting studies along these lines can be found in [4], where the outlier probability of the CML and an alternative UML method was studied for finite M, N; in [8], where the large signal to noise ratio situation is analyzed; or in [5], where the threshold effect is studied in terms of the eigenstructure of the asymptotic covariance matrix. It must be mentioned that [4] does not really analyze the UML method in (2) but a more manageable version based on the unconcentrated ML cost function (depending on all the parameters instead of only the DoAs), which will generally have a different resolution probability. Note that the use of (2) is more convenient for DoA detection applications, since the search space is constrained to the parameters of interest.

In this paper, we follow the steps in [4] and try to characterize the resolution probability of the two ML algorithms, although we concentrate on the more popular version of the UML cost function as in (2), instead of the unconcentrated version considered in [4]. We will assume that the received signals follow the unconditional model, so that they can be modeled as Gaussian random variables. To overcome the difficulty of the complicated dependence of the UML cost function on the received snapshots, we take an asymptotic approach and characterize the asymptotic behavior of the finite dimensional distributions of $\hat{\eta}_{UML}(\theta)$ when both the sample size and the array dimension increase without bound at the same rate, i.e. $M, N \to \infty$, $M/N \to c$, $0 < c < 1$. This asymptotic behavior will provide very good approximations of the behavior of $\hat{\eta}_{UML}(\theta)$, and hence the resolution probability, for finite values of M, N.

2. MAIN RESULT

The following theorem establishes the asymptotic behavior of finite dimensional distributions of the two cost functions $\hat{\eta}_{CML}(\theta)$ and $\hat{\eta}_{UML}(\theta)$ when $M, N \to \infty$. Under these asymptotic conditions, the two ML cost functions become (pointwise) asymptotically close to two deterministic counterparts, which will be denoted by $\bar{\eta}_{CML}(\theta)$ and $\bar{\eta}_{UML}(\theta)$, in the sense that $|\hat{\eta}_{CML}(\theta) - \bar{\eta}_{CML}(\theta)| \to 0$ almost surely pointwise in θ as $M, N \to \infty$, and equivalently for the UML cost function. These deterministic functions take the form $\bar{\eta}_{CML}(\theta) = \frac{1}{M}\operatorname{tr}\left[\mathbf{P}_A^{\perp}(\theta)\mathbf{R}\right]$ and

$$\bar{\eta}_{UML}(\theta) = \frac{1}{M}\log\det\left[\sigma^2(\theta)\mathbf{P}_A^{\perp}(\theta) + \mathbf{P}_A(\theta)\mathbf{R}\mathbf{P}_A(\theta)\right] + \frac{N-K}{M}\log\left(\frac{N}{N-K}\right) - \frac{K}{M} \quad (3)$$

respectively, where $\mathbf{R} = \mathbf{A}(\bar{\theta})\mathbf{P}_s\mathbf{A}(\bar{\theta})^H + \sigma^2\mathbf{I}_M$ is the true covariance matrix of the observations and $\sigma^2(\theta) = \frac{1}{M-K}\operatorname{tr}\left[\mathbf{P}_A^{\perp}(\theta)\mathbf{R}\right]$.

The following theorem characterizes the nature of the fluctuations of these two cost functions in the asymptotic regime. Let $\theta_1, \ldots, \theta_L$ be a set of multidimensional points in Θ. We define

$$\hat{\eta}_{CML} = \left[\hat{\eta}_{CML}(\theta_1), \ldots, \hat{\eta}_{CML}(\theta_L)\right]^T$$
$$\bar{\eta}_{CML} = \left[\bar{\eta}_{CML}(\theta_1), \ldots, \bar{\eta}_{CML}(\theta_L)\right]^T$$

and take the equivalent definitions for the UML cost function. The next theorem shows that when $M, N \to \infty$ the random vectors $\hat{\eta}_{CML}, \hat{\eta}_{UML}$ fluctuate as Gaussian random variables around their asymptotic means $\bar{\eta}_{CML}, \bar{\eta}_{UML}$. The result is also valid if we allow the number of sources K to depend on M, as long as $\limsup K/M < 1$. However, to avoid unnecessary technical complications, we will generally stick to the simpler case of fixed K in the following discussion. The angle vectors θ_ℓ may depend on M, assuming that $\mathbf{P}_A(\theta_\ell)$ is well defined over all M.

Theorem 1 *Assume that $\{\mathbf{y}(n), n = 1, \ldots, N\}$ forms a sequence of i.i.d. circularly symmetric Gaussian random variables, whose real and imaginary parts are independent, have zero mean and covariance $0.5\mathbf{R}$. Assume that the spectral norm of $\mathbf{R}$ is bounded in M and that σ^2 remains constant regardless of M. Consider the quantity $\sigma_\ell^2 = \bar{\eta}_{CML}(\theta_\ell) = \frac{1}{M-K}\operatorname{tr}\left[\mathbf{P}_A^{\perp}(\theta_\ell)\mathbf{R}\right]$ and define the $M \times M$ matrix $\mathcal{P}_\ell$ as*

$$\mathcal{P}_\ell = \mathbf{P}_A(\theta_\ell)\left(\mathbf{P}_A(\theta_\ell)\mathbf{R}\mathbf{P}_A(\theta_\ell)\right)^{\#}\mathbf{P}_A(\theta_\ell)$$

where $(\cdot)^{\#}$ denotes the Moore-Penrose pseudo-inverse. Then, as $M, N \to \infty$, $M/N \to c$, $0 < c < 1$, the random vectors $M\Gamma_{CML}^{-1}(\hat{\eta}_{CML} - \bar{\eta}_{CML})$ and $M\Gamma_{UML}^{-1}(\hat{\eta}_{UML} - \bar{\eta}_{UML})$ converge in law to a multivariate standardized Gaussian distribution, where

$$\{\Gamma_{CML}\}_{\ell,m} = \frac{1}{N}\operatorname{tr}\left[\mathbf{R}\mathbf{P}_A^{\perp}(\theta_\ell)\mathbf{R}\mathbf{P}_{\mathbf{A}}^{\perp}(\theta_m)\right]$$

and

$$\{\Gamma_{UML}\}_{\ell,m} = \frac{\{\Gamma_{CML}\}_{\ell,m}}{\sigma_\ell^2 \sigma_m^2} + \frac{1}{\sigma_\ell^2}\frac{1}{N}\operatorname{tr}\left[\mathbf{P}_A^\perp(\theta_\ell)\mathbf{R}\mathcal{P}_m\mathbf{R}\right]$$
$$+\frac{1}{\sigma_m^2}\frac{1}{N}\operatorname{tr}\left[\mathbf{P}_A^\perp(\theta_m)\mathbf{R}\mathcal{P}_\ell\mathbf{R}\right] - \log\left(1 - \frac{1}{N}\operatorname{tr}\left[\mathbf{R}\mathcal{P}_\ell\mathbf{R}\mathcal{P}_m\right]\right).$$

Before turning to the proof of the above result, let us draw some conclusions that can be derived from it. First of all, it is interesting to observe that when $M, N \to \infty$, the two cost functions are asymptotically close to the two deterministic counterparts $\bar{\eta}_{CML}(\theta)$ and $\bar{\eta}_{UML}(\theta)$. Curiously enough, these two cost functions are equal (up to a constant, for the UML case) to the cost functions that one would obtain by letting $N \to \infty$ for fixed M. Hence, if the problem is well posed and the DoAs are identifiable, the two deterministic equivalents $\bar{\eta}_{CML}(\theta)$ and $\bar{\eta}_{UML}(\theta)$ will have a single global minimum located at the true DoAs, namely $\bar{\theta} = \arg\min_{\theta\in\Theta} \bar{\eta}_{CML}(\theta) = \arg\min_{\theta\in\Theta} \bar{\eta}_{UML}(\theta)$. This corroborates the good asymptotic properties of both ML algorithms in situations where M, N are comparable in magnitude.

On the other hand, it is also true that, even if the two asymptotic equivalents present a single global minimum at the true value of the DoAs, these functions are in practice highly multimodal, i.e. they present several local minima. Since the original random cost functions fluctuate around these deterministic equivalents, there is a certain probability that one of these local minima will be lower than the one in $\bar{\theta}$, thus generating an outlier in the DoA estimation process. When this happens, we declare loss of resolution. Therefore, we can define the resolution probability as the complementary of the outlier probability, which was was well studied in [4]. More specifically, if $\hat{\eta}(\theta)$ is a generic cost function that fluctuates around a deterministic $\bar{\eta}(\theta)$, which has $L+1$ local minima at the values $\bar{\theta}, \theta_1, \ldots, \theta_L$, the probability of resolution can be defined as

$$P_{res} = \mathbb{P}\left[\bigcap_{\ell=1}^{L}\left\{\hat{\eta}(\theta_\ell) > \hat{\eta}(\bar{\theta})\right\}\right]. \quad (4)$$

It was shown in [4] that this definition of P_{res} provides a very accurate description of both the breakdown effect and the expected mean squared error (MSE) of the DoA estimation process. Unfortunately, in our ML setting, (4) is difficult to analyze for finite values of M, N due to the complicated structure of the cost functions (1)-(2). For this reason, [4] focused instead on the union bound of (4) –obtained when assuming independent events– and considered a simpler form for the UML cost function, assuming independent source signals. Theorem 1 provides a very simple way of approximating (4), by simply using the asymptotic distributions (as $M, N \to \infty$) instead of the actual ones. It will be shown below via simulations that the result provides a very accurate description of the actual probability, even for very low M, N.

It is important to remark that Theorem 1 does not formally imply that the use of the asymptotic finite-dimensional distributions in (4) would give the asymptotic resolution probability of the CML and UML methods. This is because the number of local minima L of $\bar{\eta}(\theta)$ may in practice increase with M, and this substantially complicates the asymptotic behavior of (4). We conjecture that this will be the case for reasonably well behaved $\mathbf{A}(\theta)$, but a more rigorous study of this problem is left for future research.

3. SKETCH OF THE PROOF OF THEOREM 1

Due to the strict space constraints, we only provide an outline of the proof corresponding to the UML cost function (the proof for the CML cost follows the same steps). We begin by pointing that, using simple algebra, the expression of $\hat{\eta}_{UML}(\theta_\ell)$ in (2) can be re-written as

$$\hat{\eta}_{UML}(\theta_\ell) = \frac{M-K}{M}\log\hat{\sigma}_\ell^2 + \frac{1}{M}\log\det\left[\mathbf{Q}^H(\theta_\ell)\hat{\mathbf{R}}\mathbf{Q}(\theta_\ell)\right] \quad (5)$$

where $\mathbf{Q}(\theta_\ell)$ is an $M \times K$ matrix of orthogonal columns such that $\mathbf{P}_A(\theta_\ell) = \mathbf{Q}(\theta_\ell)\mathbf{Q}^H(\theta_\ell)$ and where $\hat{\sigma}_\ell^2 = \frac{1}{M-K}\operatorname{tr}\left[\mathbf{P}_A^\perp(\theta_\ell)\hat{\mathbf{R}}\right]$. It is apparent from (5) that the only source of randomness in $\hat{\eta}_{UML}(\theta_\ell)$ is through the spectrum (eigenvalues) of $\hat{\mathbf{R}}$ and $\mathbf{Q}^H(\theta_\ell)\hat{\mathbf{R}}\mathbf{Q}(\theta_\ell)$. We now introduce a novel result that describes the asymptotic fluctuations of linear spectral statistics associated with multiple matrices of this type.

3.1. An intermediate result

Let $\{\mathbf{R}_1, \ldots, \mathbf{R}_L\}$ denote a collection of L Hermitian positive semidefinite matrices of dimension $M \times M$, and assume that their spectral norm is bounded in M. Let $\hat{\mathbf{R}}_\ell = N^{-1}\mathbf{R}_\ell^{1/2}\mathbf{X}\mathbf{X}^H\left(\mathbf{R}_\ell^{1/2}\right)^H$, where $\mathbf{R}_\ell^{1/2}$ is a non-necessarily Hermitian $M \times M$ matrix such that $\mathbf{R}_\ell = \mathbf{R}_\ell^{1/2}\left(\mathbf{R}_\ell^{1/2}\right)^H$, and $\mathbf{X}$ is an $M \times N$ matrix of i.i.d. Gaussian random variables with law $\mathcal{CN}(0,1)$. Taking $z \in \mathbb{C}^+$ (upper complex semi-plane), we define the random complex functions

$$\hat{m}_\ell(z) = \frac{1}{M}\operatorname{tr}\left[\left(\hat{\mathbf{R}}_\ell - z\mathbf{I}_M\right)^{-1}\right]$$

where $\ell = 1, \ldots, L$. Similarly, for each $z \in \mathbb{C}^+$, we define $\omega_\ell(z)$ as the unique solution in $\mathbb{C}^+$ of the following equation

$$z = \omega_\ell(z)\left(1 - \frac{1}{N}\operatorname{tr}\left[\mathbf{R}_\ell\left(\mathbf{R}_\ell - \omega_\ell(z)\right)^{-1}\right]\right)$$

and we take $\bar{m}_\ell(z)$ as

$$\bar{m}_\ell(z) = \frac{\omega_\ell(z)}{z}\frac{1}{M}\operatorname{tr}\left[\left(\hat{\mathbf{R}}_\ell - \omega_\ell(z)\mathbf{I}_M\right)^{-1}\right].$$

Consider the limit of $\omega_\ell(z)$ and $\bar{m}_\ell(z)$ as z goes the real axis and let $\mathcal{S}_\ell$ be the positive support[1] of $\operatorname{Im}\bar{m}_\ell(z)$ on $\mathbb{R}$.

Finally, let $\{f_1(z),\ldots,f_L(z)\}$ be real-valued functions that are holomorphic on the positive real axis, $\mathbb{R}^+$. Consider the $L\times 1$ random vector $\hat{\eta}=\left[\hat{\eta}^{(1)}\ldots\hat{\eta}^{(L)}\right]^T$, each $\hat{\eta}^{(\ell)}$ being defined as

$$\hat{\eta}^{(\ell)}=\frac{1}{2\pi\,\mathrm{j}}\oint_{C_\ell} f_\ell(z)\hat{m}_\ell(z)dz \tag{6}$$

where C_ℓ is a clockwise oriented contour enclosing all or part of $\mathcal{S}_\ell$, $C_\ell\cap\mathcal{S}_\ell=\varnothing$, and such that it does not enclose $z=0$. Let $\bar{\eta}=\left[\bar{\eta}^{(1)}\ldots\bar{\eta}^{(L)}\right]^T$, where $\bar{\eta}^{(\ell)}$ is defined as $\eta^{(\ell)}$ replacing $\hat{m}_\ell(z)$ with $\bar{m}_\ell(z)$. The following theorem provides a central limit theorem on the quantity $M(\hat{\eta}-\bar{\eta})$. The proof can be obtained using the approach in [6]; we omit the details due to space constraints.

Theorem 2 *According to the above definitions, consider an* $L\times L$ *matrix* Γ *with entries*

$$\{\Gamma\}_{\ell,m}=\frac{1}{2\pi\,\mathrm{i}}\frac{1}{2\pi\,\mathrm{i}}\oint_{C_{\omega_\ell}}\oint_{C_{\omega_m}} g_\ell(\omega_1)\,g_m(\omega_2)\times$$

$$\times\left(\frac{\Psi_{2,1}\Psi_{1,2}}{(1-\Psi_{1,1})^2}+\frac{\Psi_{2,2}}{1-\Psi_{1,1}}\right)d\omega_1 d\omega_2 \tag{7}$$

where $C_{\omega_\ell}=\omega_\ell(C_\ell)$,

$$g_\ell(\omega)=f_\ell\left(\omega\left(1-\frac{1}{N}\operatorname{tr}\left[\mathbf{R}_\ell\left(\mathbf{R}_\ell-\omega\right)^{-1}\right]\right)\right)$$

and where $\Psi_{k_1,k_2}=\Psi_{k_1,k_2}(\omega_1,\omega_2)$ *is a complex function of two variables* (ω_1,ω_2) *defined as*

$$\Psi_{k_1,k_2}=\frac{1}{N}\operatorname{tr}\left[\left(\mathbf{R}_\ell^{1/2}\right)^H\left(\mathbf{R}_\ell-\omega_1\mathbf{I}_M\right)^{-k_1}\mathbf{R}_\ell^{1/2}\times\right.$$

$$\left.\times\left(\mathbf{R}_m^{1/2}\right)^H\left(\mathbf{R}_m-\omega_2\mathbf{I}_M\right)^{-k_2}\mathbf{R}_m^{1/2}\right].$$

Assume that Γ *is invertible and that the spectral norm of* Γ^{-1} *is bounded in* M. *Then, for a fixed L and when* $M,N\to\infty$, $M/N\to c$, $0<c<1$, $M\Gamma^{-1}(\eta-\bar{\eta})$ *converges in law to a multivariate standardized Gaussian random vector.*

3.2. Finishing the proof of Theorem 1

Theorem 2 has a direct application to our problem, if we simply consider the matrices $\hat{\mathbf{R}}_\ell=\mathbf{P}_A(\theta_\ell)\hat{\mathbf{R}}\mathbf{P}_A(\theta_\ell)$ and $\mathbf{R}_\ell=\mathbf{P}_A(\theta_\ell)\mathbf{R}\mathbf{P}_A(\theta_\ell)$, together with $\hat{\mathbf{R}}_0=\hat{\mathbf{R}}$ and $\mathbf{R}_0=\mathbf{R}$. Using these definitions, we can express

$$\hat{m}_0(z)=\frac{1}{M}\operatorname{tr}\left[\left(\hat{\mathbf{R}}-z\mathbf{I}_M\right)^{-1}\right]$$

$$\hat{m}_\ell(z)=\frac{1}{M}\operatorname{tr}\left[\left(\mathbf{P}_A(\theta_\ell)\hat{\mathbf{R}}\mathbf{P}_A(\theta_\ell)-z\mathbf{I}_M\right)^{-1}\right]$$

for $\ell=1\ldots L$, which allows us to rewrite (5) as

$$\hat{\eta}_{UML}(\theta_\ell)=\frac{1}{2\pi\,\mathrm{j}}\oint_{C_\ell}\hat{m}_\ell(z)\log z dz+\left(1-\frac{K}{M}\right)\times$$

$$\times\log\left[\frac{M}{M-K}\frac{1}{2\pi\,\mathrm{j}}\left(\oint_{C_0} z\hat{m}_0(z)dz-\oint_{C_\ell} z\hat{m}_\ell(z)dz\right)\right] \tag{8}$$

where C_ℓ (resp. C_0) is a clockwise oriented contour not intersecting with any eigenvalue of $\mathbf{P}_A(\theta_\ell)\hat{\mathbf{R}}\mathbf{P}_A(\theta_\ell)$ (resp. $\hat{\mathbf{R}}$), enclosing only its non-zero eigenvalues. According to the above identity, we are able to express the column vector $\hat{\eta}_{UML}$ as a function of quantities of the form in (6) for $\ell=0,\ldots,L$, where $f_0(z)=z/\sigma_\ell^2$ and $f_\ell(z)=\log z-z/\sigma_\ell^2$ are complex functions that are analytical on $\mathbb{R}^+$.

With some abuse of notation, let $\bar{\eta}_{UML}(\theta_\ell)$ denote the quantity in (8) after replacing $\hat{m}_\ell(z)$ with $\bar{m}_\ell(z)$ according to the definitions in Theorem 2. Using complex integration techniques one can show that, with this definition, $\bar{\eta}_{UML}(\theta_\ell)$ coincides with the quantity in (3). Therefore, we can write

$$\hat{\eta}_{UML}(\theta_\ell)-\bar{\eta}_{UML}(\theta_\ell)=\frac{1}{2\pi\,\mathrm{j}}\oint_{C_\ell}\left[\hat{m}_\ell(z)-\bar{m}_\ell(z)\right]f_\ell(z)dz$$

$$+\frac{1}{2\pi\,\mathrm{j}}\oint_{C_0} z\left[\hat{m}_0(z)-\bar{m}_0(z)\right]f_0(z)dz+\chi_\ell$$

where we have introduced the random variable

$$\chi_\ell=\left(1-\frac{K}{M}\right)\left[\frac{\hat{\sigma}_\ell^2-\sigma_\ell^2}{\sigma_\ell^2}-\log(1+\frac{\hat{\sigma}_\ell^2-\sigma_\ell^2}{\sigma_\ell^2})\right].$$

It can readily be seen that $|M\chi_\ell|\to 0$ in probability as $M,N\to\infty$, and therefore these terms do not contribute to the asymptotic distribution. From this point, Theorem 1 follows from carrying out the integration in (7) with the present definitions of $f_0(z)$ and $f_\ell(z)$ and then applying the linear transformation formula for multivariate Gaussian random variables. The fact that all covariance matrices are positive definite will follow from the fact that $\mathbf{R}>\sigma^2\mathbf{I}_M$ and that σ^2 does not vary with M.

4. NUMERICAL EVALUATION AND ANALYSIS OF THE RESULTS

We consider a uniform linear array with $M=5$ elements located half a wavelength apart. We first investigate the accuracy of the asymptotic laws in order to describe the resolution probability in (4). To that effect, we assume that the array receives the signals from two zero-mean Gaussian uncorrelated sources received with equal power from the angles of arrival $[16^\circ,18^\circ]$. In Figure 1 we compare the resolution probability in (4) obtained from averaging 10^4 realizations of the

[1] It can be shown that these limits exist, and that $\mathcal{S}_\ell$ is compact. Furthermore, both $\omega_\ell(z)$ and $\bar{m}_\ell(z)$ can be analytically extended to $\mathbb{C}\backslash\{0\}\cup\mathcal{S}_\ell$.

CML and UML cost functions (dotted lines), together with the predicted value using the asymptotic statistics in Theorem 1 (solid lines). Results are given as a function of the SNR observed at each array element, for different values of the sample volume (N). The Gaussian cdf was numerically evaluated using the approach in [7]. According to these results, a good match is observed between the simulated and predicted values of the resolution probability, even for relatively low M, N.

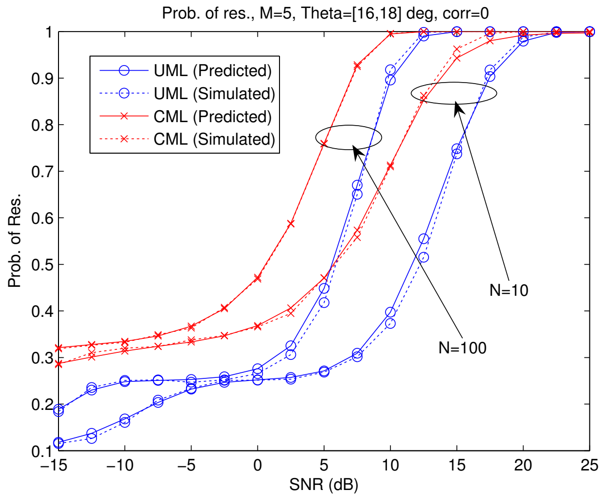

Fig. 1. Resolution probability (uncorrelated sources).

It might be concluded from Figure 1 that in general terms the UML techniques provide a much lower resolution probability compared to CML. However, there exist certain specific scenarios where UML is a clear alternative to CML. This is the case of situations where the source correlation is extremely high. To show this, in Figure 2 we plot the predicted resolution probability as a function of the SNR in the situation where the sources are highly correlated (correlation coefficient equal to 0.95). We observe that, when the sample size is sufficiently large, UML is clearly more robust against high source correlation, as already pointed out in [5].

5. CONCLUSIONS

This paper has investigated the asymptotic behavior of the conditional and unconditional ML DoA estimation procedures in the threshold region, where the generation of outliers causes a total performance breakdown of both methods. Assuming that the number of antennas and the sample size go to infinity at the same rate, the finite dimensional distributions of the CML and UML cost functions have been shown to be asymptotically Gaussian. This result has been used to approximate the resolution probability of these two methods under different scenarios. The resulting analysis corroborates the general idea that CML provides a better resolution probability than UML, except for situations with highly correlated source signals and relatively large sample size.

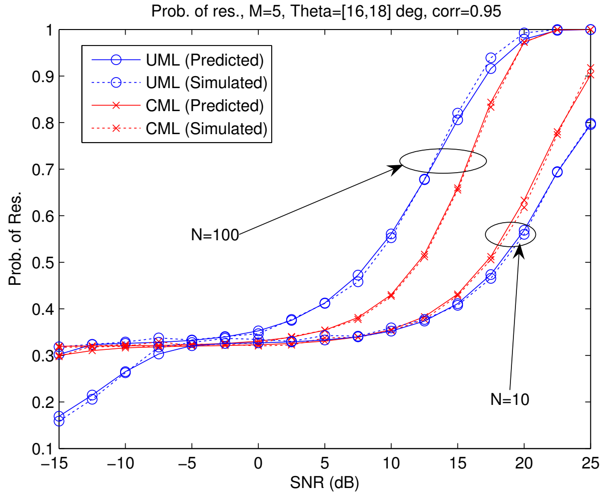

Fig. 2. Resolution probability (highly correlated sources).

6. REFERENCES

[1] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. ASSP*, vol. 38, pp. 1783–1795, Oct. 1990.

[2] J.F. Böhme, "Estimation of spectral parameters of correlated signals in wavefields," *Signal Processing (EURASIP)*, vol. 11, pp. 329–337, 1986.

[3] P. Stoica and A. Nehorai, "On the concentrated likelihood function in array signal processing," *Circuits, Syst. and Sig. Proc.*, vol. 14, no. 5, pp. 669–674, 1995.

[4] F. Athley, "Threshold region performance of maximum likelihood direction of arrival estimators," *IEEE Trans. on Sig. Proc.*, vol. 53, no. 4, pp. 1359–1373, Apr. 2005.

[5] Y.I. Abramovich and B.A. Johnson, "Comparative threshold performance study for conditional and unconditional direction-of-arrival estimation," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Sig. Proc.*, 2011, pp. 4176–4179.

[6] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, and L. Pastur, "A new approach for mutual information analysis of large dimensional multi-antenna channels," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3987–4004, Sept. 2008.

[7] A. Genz, "Numerical computation of multivar. normal prob.," *J. Comp. Graph Stat.*, vol. 1, pp. 141–149, 1992.

[8] A. Renaux, P. Forster, and P. Larzabal, "Unconditional maximum likelihood performance at finite number of samples and high signal-to-noise ratio," *IEEE Trans. on Sig. Proc.*, vol. 55, no. 5, pp. 2358–2364, May 2007.