# Maximum Likelihood Localization of Sources in Noise Modeled as a Stable Process

Panagiotis Tsakalides, *Student Member, IEEE,* and Chrysostomos L. Nikias, *Fellow, IEEE*

***Abstract*—This paper introduces a new class of robust beamformers which perform optimally over a wide range of non-Gaussian additive noise environments. The maximum likelihood approach is used to estimate the bearing of multiple sources from a set of snapshots when the additive interference is impulsive in nature. The analysis is based on the assumption that the additive noise can be modeled as a *complex symmetric $\alpha$-stable* ($S\alpha S$) process. Transform-based approximations of the likelihood estimation are used for the general $S\alpha S$ class of distributions while the exact probability density function is used for the Cauchy case. It is shown that the Cauchy beamformer greatly outperforms the Gaussian beamformer in a wide variety of non-Gaussian noise environments, and performs comparably to the Gaussian beamformer when the additive noise is Gaussian. The Cramér-Rao bound for the estimation error variance is derived for the Cauchy case, and the robustness of the $S\alpha S$ beamformers in a wide range of impulsive interference environments is demonstrated via simulation experiments.**

## I. Introduction

THIS paper addresses the solution of the signal parameter estimation problem through the use of sensor array data retrieved in the presence of impulsive interference. One of the most interesting problems in this area is the estimation of the direction of arrival (DOA) of narrow-band source signals having the same known center frequency. In the past, the problem has been studied extensively under the assumption of Gaussian distributed signals and/or noise, and a variety of methods for its solution have been proposed. As a result of the Gaussianity assumption, most methods are based on the second-order statistics of the signals.

The maximum likelihood (ML) method was one of the first to be investigated [1]. When applying the ML technique to the source localization problem, two different assumptions for the signal waveforms result in two different methods. According to the *stochastic ML* (SML), the signals are modeled as Gaussian random processes. This is often motivated by the central limit theorem and results in mathematically convenient expressions. On the other hand, in the *deterministic ML* (DML), the signals are considered to be unknown, deterministic quantities that need to be estimated in conjunction with the direction of arrival. This is a natural model for digital communication applications, where the signals are far from being normal random variables and where estimation of the signal is of equal interest.

Manuscript received April 24, 1994; revised April 26, 1995. This work was supported by the Office of Naval Research under Contract N00014-92-J-1034.

The authors are with the Signal and Image Processing Institute Department of Electrical Engineering—Systems, University of Southern California, Los Angeles, CA 90089-2564 USA.

IEEE Log Number 9415086.

The importance of the ML technique comes from the mathematical property that under certain regularity conditions, the ML estimator is known to be asymptotically efficient, i.e., it achieves the Cramér-Rao bound (CRB) for the estimation error variance. In this sense, ML has the best possible asymptotic properties.

Many researchers have studied the ML technique as a means of approaching the source localization problem in the presence of Gaussian additive noise. Stoica and Nehorai examined the ML performance and its asymptotic properties, derived expressions for the CRB, and established some of its properties [2]. Additionally, they investigated the relationship between the ML and large sample approximation methods such as MUSIC. With high computational cost constituting the main shortcoming of the ML technique, Ziskind and Wax introduced a computationally attractive method for calculating the ML estimator [3]. Their method was based on an iterative technique called "alternating projection" (AP) that transformed the multivariate optimization problem into a sequence of simpler 1-D optimization problems.

However, due to the high computational load of the multivariate nonlinear maximization problem involved in the ML estimator, suboptimal methods have also been developed [4]–[13]. The better known ones are cited here: the minimum variance distortionless method of Capon [4] and the so-called eigenvector-based methods including the MUSIC [5], minimum norm [6], [7], and the ESPRIT method [8]. The performance of the aforementioned methods is inferior to that of the ML method, especially for low signal-to-noise ratio (SNR) values or when the number of observation snapshots is small.

The MUSIC method, which is a generalization of Pisarenko's harmonic retrieval method, has received the most attention and triggered the development of a large number of algorithms referred to as *eigenvector* or *subspace* techniques. Besides offering a new geometric interpretation of the array processing problem, MUSIC uses concepts from complex vector spaces and well-known tools from linear algebra, such as the singular value decomposition (SVD), in order to achieve high resolution while keeping the computational complexity relatively low compared with that of the ML methods.

Recently, special interest has been shown in relaxing some of the assumptions concerning the statistical nature of the noise in the bearing estimation problem [14]–[16]. One such method is the bispectrum beamformer introduced by Forster and Nikias [15]. It was demonstrated that for the case of spatially correlated Gaussian additive noise with an unknown

1053-587X/95$04.00 © 1995 IEEE

cross-spectral matrix (CSM), the bispectrum beamformer may provide asymptotically better bearing estimates than the stochastic ML method with known CSM.

Often, in actual applications, the Gaussian noise assumption proves inadequate, as systems designed under this assumption exhibit a significant performance degradation. There exist physical processes generating interferences that contain **noise components that are impulsive in nature**. These processes can be natural, as well as man-made, and include underwater acoustic signals, lightning in the atmosphere, and transients in power lines and car ignitions. In modeling these type of signals, the *stable distribution law* provides a very attractive theoretical tool. It was proven that under broad conditions, a general class of impulsive noise follows the stable law [17]. As a result, considerable research interest has been shown in designing robust signal processing algorithms for detection, direction finding, and equalization that can do well not only in the presence of Gaussian noise but also, more importantly, in the presence of impulsive noise environments [18]–[20].

This paper is devoted to the ML estimation of multiple sources in the presence of additive interference that is modeled as a complex isotropic $\alpha$-stable process. The paper is organized as follows: In Section II, the DOA problem is formulated. In Section III, the statistical model, based on the class of *bivariate symmetric $\alpha$-stable ($S\alpha S$) distributions*, is introduced. This model is well-suited for describing noise processes that are impulsive in nature and contains the Gaussian process as a special case. In Section IV, the ML estimator is developed, and the CRB is derived for the Cauchy case. The application of the ML method for the general case of $S\alpha S$ processes is also discussed. Finally, Monte-Carlo simulation results are presented in Section V, and conclusions are drawn in Section VI.

## II. Problem Formulation

Consider an array of $r$ sensors with arbitrary locations and arbitrary directional characteristics, which receive signals generated by $q$ narrowband sources with known center frequency $\omega$ and locations $\theta_1, \theta_2, \ldots, \theta_q$. Since the signals are narrowband, the propagation delay across the array is much smaller than the reciprocal of the signal bandwidth, and it follows that by using a complex envelop representation, the array output can be expressed as

$$\mathbf{x}(t) = \sum_{k=1}^{q} \mathbf{a}(\theta_k) s_k(t) + \mathbf{n}(t) \tag{1}$$

where we have the following:

- $\mathbf{x}(t) = [x_1(t), \ldots, x_r(t)]^T$ is the vector of the signals received by the array sensors.
- $s_k(t)$ is the signal emitted by the $k$th source as received at the reference sensor 1 of the array.
- $\mathbf{a}(\theta_k) = [1, e^{-j\omega\tau_2(\theta_k)}, \ldots, e^{-j\omega\tau_r(\theta_k)}]^T$ is the steering vector of the array toward direction $\theta_k$.
- $\tau_i(\theta_k)$ is the propagation delay between the first and the $i$th sensor for a waveform coming from direction $\theta_k$.
- $\mathbf{n}(t) = [n_1(t), \ldots, n_r(t)]^T$ is the noise vector.

Equation (1) can be expressed in a compact form as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t) \tag{2}$$

where $\mathbf{A}(\boldsymbol{\theta})$ is the $r \times q$ matrix of the array steering vectors

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_q)] \tag{3}$$

and $\mathbf{s}(t)$ is the $q \times 1$ vector of the signals

$$\mathbf{s}(t) = [s_1(t), \ldots, s_q(t)]^T. \tag{4}$$

Assuming that $M$ snapshots are taken at time instants $t_1, \ldots, t_M$, the data can be expressed as

$$\mathbf{X} = \mathbf{A}(\boldsymbol{\theta})\mathbf{S} + \mathbf{N} \tag{5}$$

where $\mathbf{X}$ and $\mathbf{N}$ are the $r \times M$ matrices

$$\mathbf{X} = [\mathbf{x}(t_1), \ldots, \mathbf{x}(t_M)] \tag{6}$$

$$\mathbf{N} = [\mathbf{n}(t_1), \ldots, \mathbf{n}(t_M)] \tag{7}$$

and $\mathbf{S}$ is the $q \times M$ matrix

$$\mathbf{S} = [\mathbf{s}(t_1), \ldots, \mathbf{s}(t_M)]. \tag{8}$$

Our objective is to estimate the DOA's $\theta_1, \ldots, \theta_q$ of the sources from the $M$ snapshots of the array $\mathbf{x}(t_1), \ldots, \mathbf{x}(t_M)$.

Toward this goal, we are going to make the following assumptions regarding the array, the signals, and the noise:

A.1) The number of signals is known and is smaller than the number of sensors, i.e., $q < r$.

A.2) The set of any $q$ steering vectors is linearly independent.

A.3) The noise samples $n_i(t_j)$; $i = 1, \ldots, r$; $j = 1, \ldots, M$ come from a complex (bivariate) *isotropic stable distribution.*

A.4) The noise samples $n_i(t_j)$ are statistically independent from one another both along the array sensors, namely, along index $i$ and along time, namely, along index $j$.

Assumptions A.1) and A.2) guarantee the uniqueness of the solution. Assumption A.3) draws a new element into our analysis as we deviate from the conventional assumption that the noise in sensor arrays is a complex-valued Gaussian process.

## III. Symmetric Alpha-Stable ($S\alpha S$) Distributions

In this section, we introduce the statistical model that will be used to describe the additive noise. The model is based on the class of *bivariate symmetric $\alpha$-stable ($S\alpha S$) distributions* and is well-suited for describing noise processes that are impulsive in nature. An extensive review of the $S\alpha S$ family can be found in [21]. Here, we focus on the elements useful for our analysis later on the paper.

### *A. The Class of SαS Distributions*

The symmetric $\alpha$-stable ($S\alpha S$) distribution is best defined by its characteristic function

$$\varphi(\omega) = \exp(\jmath\delta\omega - \gamma|\omega|^{\alpha}) \tag{9}$$

where $\alpha$ is the *characteristic exponent* restricted to the values $0 < \alpha \le 2$, $\delta$ ($-\infty < \delta < \infty$) is the *location parameter*, and $\gamma$ ($\gamma > 0$) is the *dispersion* of the distribution. For values of $\alpha$ in the interval $(1, 2]$, the location parameter $\delta$ corresponds to the mean of the $S\alpha S$ distribution, whereas for $0 < \alpha \le 1$, $\delta$ corresponds to its median. The dispersion parameter $\gamma$ determines the spread of the distribution around its location parameter $\delta$ similar to the variance of the Gaussian distribution. The characteristic exponent $\alpha$ is the most important parameter of the $S\alpha S$ distribution, and it determines the shape of the distribution.

A stable distribution is called *standard* if $\delta = 0$ and $\gamma = 1$. Clearly, if a random variable $X$ is stable with parameters $\alpha, \gamma, \delta$, then $(X - \delta)/\gamma^{1/\alpha}$ is standard with characteristic exponent $\alpha$.

By letting $\alpha$ take the values 1 and 2, we get two important special cases of $S\alpha S$ distributions, namely, the **Cauchy** ($\boldsymbol{\alpha = 1}$), and the **Gaussian** ($\boldsymbol{\alpha = 2}$):

Cauchy

$$f_1(\gamma, \delta; x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x - \delta)^2} \tag{10}$$

Gaussian

$$f_2(\gamma, \delta; x) = \frac{1}{\sqrt{4\pi\gamma}} \exp\left[-\frac{(x - \delta)^2}{4\gamma}\right]. \tag{11}$$

Unfortunately, no closed-form expressions exist for general $S\alpha S$ distributions other than the Cauchy and the Gaussian. However, power series expansions can be derived for $f_\alpha(\gamma, \delta; x)$. In the following, we shall assume that all $S\alpha S$ distributions are centered at the origin, i.e., the location parameter $\delta = 0$. This is equivalent to the zero-mean assumption for Gaussian distributions. Then, the standard $S\alpha S$ density function is given by (12), which appears at the bottom of the page [21].

Although the $S\alpha S$ density behaves approximately like a Gaussian density near the origin, its tails decay at a lower rate than the Gaussian density tails. **The smaller the characteristic exponent $\boldsymbol{\alpha}$ is, the heavier the tails of the SαS density.** This implies that random variables following $S\alpha S$ distributions with small characteristic exponents are *highly impulsive*. It is this heavy-tail characteristic that makes the $S\alpha S$ densities appropriate for modeling signals and noise or interference which is impulsive in nature. Furthermore, $S\alpha S$ densities obey two important properties that further justify their role in data modeling:

- The *stability property*, which states that the random variables $X_1, \ldots, X_n$ are independent and symmetrically stable with the same characteristic exponent $\alpha$ if and only if for any constants $a_1, \ldots, a_n$, the linear combination $\sum_{i=1}^{n} a_i X_i$ is also $S\alpha S$
- the *generalized central limit theorem*, which states that the family of stable distributions contains all limiting distributions of sums of i.i.d. random variables.

An important difference between the Gaussian and the other distributions of the $S\alpha S$ family is that only moments of order less than $\alpha$ exist for the non-Gaussian $S\alpha S$ family members. The so-called *fractional lower order moments* (FLOM) of a $S\alpha S$ random variable with zero location parameter and dispersion $\gamma$ are given by

$$E|X|^p = C(p, \alpha)\gamma^{\frac{p}{\alpha}} \quad \text{for } 0 < p < \alpha \tag{13}$$

where

$$C(p, \alpha) = \frac{2^{p+1}\Gamma(\frac{p+1}{2})\Gamma(-\frac{p}{\alpha})}{\alpha\sqrt{\pi}\Gamma(-\frac{p}{2})}. \tag{14}$$

In modeling the signals and/or noise for the bearing estimation problem, we need a complex (bivariate) model for the noise samples. In the following subsection, we present an important family of multidimensional stable distributions suitable for our purposes.

### *B. Bivariate Isotropic Stable Distributions*

Multivariate stable distributions, much like the univariate stable distributions, are characterized by the stability property and the generalized central limit theorem. However, they are much more difficult to describe because they form a nonparametric set [21]. An exception is the family of multidimensional isotropic stable distributions. Here, we concentrate on the 2-D (bivariate) case that is appropriate for modeling signals and noise in the bearing estimation problem.

The characteristic function of a bivariate isotropic $\alpha$-stable distribution has the form

$$\varphi(\omega_1, \omega_2) = \exp(\jmath(\delta_1\omega_1 + \delta_2\omega_2) - \gamma|\boldsymbol{\omega}|^{\alpha}), \tag{15}$$

where $\boldsymbol{\omega} = (\omega_1, \omega_2)$, and $|\boldsymbol{\omega}| = \sqrt{\omega_1^2 + \omega_2^2}$.

Here again, the parameters $\alpha$ and $\gamma$ are the characteristic exponent and the dispersion, respectively. The parameters $\delta_1, \delta_2$ are the location parameters. The distribution is isotropic with respect to the point $(\delta_1, \delta_2)$. Note that the two marginal distributions of the isotropic stable distribution are $S\alpha S$ with parameters $(\delta_1, \gamma, \alpha)$ and $(\delta_2, \gamma, \alpha)$. In the following, we will assume that $(\delta_1, \delta_2) = (0, 0)$. The bivariate *isotropic Cauchy*

$$f_\alpha(x) = \begin{cases} \frac{1}{\pi x} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k!} \Gamma(\alpha k + 1) x^{-\alpha k} \sin(\frac{k\alpha\pi}{2}) & \text{for } 0 < \alpha < 1 \\ \frac{1}{\pi(x^2+1)} & \text{for } \alpha = 1 \\ \frac{1}{\pi\alpha} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k!} \Gamma(\frac{2k+1}{\alpha}) x^{2k} & \text{for } 1 < \alpha < 2 \\ \frac{1}{2\sqrt{\pi}} \exp\left[-\frac{x^2}{4}\right] & \text{for } \alpha = 2. \end{cases} \tag{12}$$

and *Gaussian* distributions are special cases for $\alpha = 1$ and $\alpha = 2$, respectively.

As in the case of the univariate $S\alpha S$ density function, when $\alpha \neq 1$ or $\alpha \neq 2$, no closed-form expressions exist for the density function of the bivariate stable random variable. By using the polar coordinate $\rho = |\mathbf{x}| = \sqrt{x_1^2 + x_2^2}$, the density function can be written as $f_{\alpha,\gamma}(x_1, x_2) = \chi_{\alpha,\gamma}(\rho)$ and is expressed as a power series expansion form in (16), which appears at the bottom of the page.The density function $\chi_{\alpha,\gamma}(\rho)$ described above is also a heavy-tailed function. An expression for the FLOM's, similar to the one for the single-dimensional case, can be found in [23]. If $X$ is a random vector in $\mathbf{R}^n$ having the isotropic stable distribution with dispersion $\gamma$, then

$$E|\mathbf{X}|^p = C_n(p, \alpha)\gamma^{\frac{p}{\alpha}} \qquad \text{for } 0 < p < \alpha \tag{17}$$

where

$$C_n(p, \alpha) = \frac{2^{p+1}\Gamma(\frac{p+n}{2})\Gamma(-\frac{p}{\alpha})}{\alpha\Gamma(\frac{n}{2})\Gamma(-\frac{p}{2})}. \tag{18}$$

Having established a statistical model for the additive noise at the array sensors, we proceed to the next section in which we estimate the direction of arrival.

## IV. Maximum Likelihood Estimation in Alpha-Stable Noise

In this section, we develop the ML estimator of the source locations in the presence of noise modeled as a **complex isotropic $\alpha$-stable process** with dispersion $\gamma$. In a similar approach, as in [3], we do not regard the source signals as sample functions of random processes but rather as *unknown deterministic sequences*. We will see that under the unknown signal assumption, the ML estimation of the source bearings cannot be decoupled from the ML estimation of the signal waveforms as it can be in the case of Gaussian additive noise. Hence, in order to reduce the dimensionality of the optimization procedure, we will use suboptimal estimates of the signals, and we will refer to the resulting processor as the pseudo ML processor.

Initially, we concentrate on the case where $\alpha = 1$, i.e., we consider the additive complex Cauchy noise case. There are two reasons for doing this: First, the Cauchy distribution has a closed-form expression for its density function. This results in a straightforward implementation of the ML estimation, with closed-form expressions for the CRB. Second, it is shown through simulations that the Cauchy beamformer is very robust in different impulsive noise environments, i.e., its performance does not change significantly when the parameter $\alpha$ of the $S\alpha S$ noise varies in the interval $[1, 2]$. When the noise follows the $S\alpha S$ distribution with $\alpha \neq 1$, the ML estimation of the DOA is approximated by exploring the simple expression for the characteristic function of the noise given by (15).

### A. Pseudo ML Bearing Estimation of Sources in Cauchy Noise

In this section, we assume that the noise present at the array sensors is modeled as a **complex isotropic Cauchy process** with pdf given by (16), shown at the bottom of the page, $\alpha = 1$). We are most interested in estimating the bearings of the sources, and therefore, we consider the signals themselves, as well as the noise dispersion, $\gamma$ as nuisance parameters in the estimation problem.

Under Assumption A.4), it follows from (1) and (16) that the joint density function of the sampled data is given by

$$f(\mathbf{X}) = \prod_{t=1}^{M}\prod_{i=1}^{r}\chi_{1,\gamma}\left(\left|x_i(t) - \sum_{k=1}^{q} a_i(\theta_k)s_k(t)\right|\right) \tag{19}$$

or

$$f(\mathbf{X}) = \prod_{t=1}^{M}\prod_{i=1}^{r}\frac{1}{2\pi}\frac{\gamma}{\left(\gamma^2 + |x_i(t) - \sum_{k=1}^{q} a_i(\theta_k)s_k(t)|^2\right)^{3/2}} \tag{20}$$

where $a_1(\theta_k) = 1$ and $a_i(\theta_k) = e^{-\jmath\omega\tau_i(\theta_k)}$; $i = 2, \ldots, r$. Hence, the log likelihood function $L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})$, ignoring constant terms, is expressed as

$$L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta}) = Mr\log(\gamma) - \frac{3}{2}\sum_{t=1}^{M}\sum_{i=1}^{r}\log \cdot \left(\gamma^2 + \left|x_i(t) - \sum_{k=1}^{q} a_i(\theta_k)s_k(t)\right|^2\right). \tag{21}$$

The ML estimator is obtained by maximizing $L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})$ with respect to $\gamma$, $\mathbf{S}$, and $\boldsymbol{\theta}$, i.e.

$$\max_{\gamma, \mathbf{S}, \boldsymbol{\theta}} L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta}). \tag{22}$$

To reduce the dimension of this optimization problem, we first fix $\gamma$ and $\boldsymbol{\theta}$, and minimize $L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})$ with respect to the signal $\mathbf{S}$. For fixed $t$, we take the derivative of $L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})$ with respect to $s_k(t)$:

$$\frac{\partial L}{\partial s_k(t)} = -3\sum_{i=1}^{r}\frac{a_i(\theta_k)[x_i(t) - a_i(\theta_k)s_k(t)]^*}{\gamma^2 + |x_i(t) - a_i(\theta_k)s_k(t)|^2}. \tag{23}$$

Unfortunately, no explicit solution of (23) is possible. In order to be able to obtain closed-form expressions for the signals, we resort to the application of the pseudo maximum likelihood (PML) estimation. PML estimation is an important

$$\chi_{\alpha,\gamma}(\rho) = \begin{cases} \frac{1}{\pi^2\gamma^{2/\alpha}}\sum_{k=1}^{\infty}\frac{2^{\alpha k}(-1)^{k-1}}{k!}(\Gamma(\alpha k/2+1))^2\sin(\frac{k\alpha\pi}{2})(\frac{\rho}{\gamma^{1/\alpha}})^{-\alpha k-2} & \text{for } 0 < \alpha < 1 \\ \frac{\gamma}{2\pi(\rho^2+\gamma^2)^{3/2}} & \text{for } \alpha = 1 \\ \frac{1}{\pi\alpha\gamma^{2/\alpha}}\sum_{k=0}^{\infty}\frac{(-1)^k}{2^{2k+1}(k!)^2}\Gamma(\frac{2k+2}{\alpha})(\frac{\rho}{\gamma^{1/\alpha}})^{2k} & \text{for } 1 < \alpha < 2 \\ \frac{1}{4\pi\gamma}\exp\left[-\frac{\rho^2}{4\gamma}\right] & \text{for } \alpha = 2. \end{cases} \tag{16}$$

method in applications where probability models abound for which the analytical derivation of the ML estimate for all the parameters is virtually impossible. The problem formulated in [24] can be stated as follows:

Let $X_1, \ldots, X_n$ be i.i.d. random variables with probability distribution $f(X; \boldsymbol{\theta}, \mathbf{S})$ indexed by two sets of parameters. Let $\hat{\mathbf{S}} = \hat{\mathbf{S}}(X_1, \ldots, X_n)$ be an estimate of $\mathbf{S}$ other than the maximum likelihood estimate, and let $\hat{\boldsymbol{\theta}}$ be the solution of the likelihood equation $\partial/\partial\boldsymbol{\theta} \; \log L(\mathbf{X}; \boldsymbol{\theta}, \hat{\mathbf{S}}) = 0$, which maximizes the likelihood. Then, $\hat{\boldsymbol{\theta}}$ is called a pseudo maximum likelihood estimate of $\boldsymbol{\theta}$, and under certain conditions, it is consistent and asymptotically normal.

The PML estimator $\hat{\boldsymbol{\theta}}$ has good large sample properties when $\hat{\mathbf{S}}$ does. In general, the asymptotic analysis for $\hat{\boldsymbol{\theta}}$ will depend on the asymptotic characteristics of $\hat{\mathbf{S}}$.

Returning to the optimization problem described in (22), we observe that maximizing $L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})$ with respect to the signal $\mathbf{S}$ is equivalent to the following minimization problem:

$$\min_{\mathbf{S}} \mathcal{L}(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta}) = \min_{\mathbf{S}} \left\{ \sum_{t=1}^{M} \sum_{i=1}^{r} \log \left( 1 + \frac{|x_i(t) - \sum_{k=1}^{q} a_i(\theta_k) s_k(t)|^2}{\gamma^2} \right) \right\}. \tag{24}$$

As we can see, (24) involves minimizing a double sum expression of logarithmic functions of the form $\log(1+z)$. In the unit disc $B_1(0) = \{z \in \mathcal{C} : |z| < 1\}$, the function $\log(1+z)$ can be expressed as an infinite series:

$$\log(1+z) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} z^n = z - \frac{z^2}{2} + - \cdots; \quad |z| < 1. \tag{25}$$

Hence, for $|x_i(t) - \sum_{k=1}^{q} a_i(\theta_k) s_k(t)| < \gamma$, the functional $\mathcal{L}(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})$ can be written in the form

$$\mathcal{L}(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta}) = \sum_{t=1}^{M} \sum_{i=1}^{r} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n\gamma^{2n}} \left| x_i(t) - \sum_{k=1}^{q} a_i(\theta_k) s_k(t) \right|^{2n}. \tag{26}$$

A first-order approximation of the above expression results in the following $\mathcal{L}^{(1)}(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})$ functional:

$$\mathcal{L}^{(1)}(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta}) = \frac{1}{\gamma^2} \sum_{t=1}^{M} \sum_{i=1}^{r} |x_i(t) - \sum_{k=1}^{q} a_i(\theta_k) s_k(t)|^2 \tag{27}$$

which, by using (2), can be written in a more compact form as

$$\mathcal{L}^{(1)}(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta}) = \frac{1}{\gamma^2} \sum_{t=1}^{M} |\mathbf{x}(t) - \mathbf{A}(\boldsymbol{\theta}) \mathbf{s}(t)|^2. \tag{28}$$

Hence, the minimization of $\mathcal{L}^{(1)}(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})$ with respect to $\mathbf{S}$ is equivalent to the least-squares (LS) estimation of $\mathbf{S}$. This problem has a well-known solution:

$$\hat{\mathbf{s}}(t) = (\mathbf{A}^H(\boldsymbol{\theta}) \mathbf{A}(\boldsymbol{\theta}))^{-1} \mathbf{A}^H(\boldsymbol{\theta}) \mathbf{x}(t). \tag{29}$$

The dispersion $\gamma$ can be estimated by using the method of moments. Namely, in the expression for the FLOM of the noise given in (17), $E|X|^p$ can be approximated by an average sum:

$$\hat{\gamma} = \frac{\left[ \frac{1}{M} \sum_{t=1}^{M} \sum_{i=1}^{r} |x_i(t) - \sum_{k=1}^{q} a_i(\theta_k) \hat{s}_k(t)|^p \right]^{\frac{1}{p}}}{[C_2(p,1)]^{\frac{1}{p}}} \tag{30}$$

where $p < 1$, and $C_2(p,1)$ is given by

$$C_2(p,1) = p 2^p \frac{\Gamma(\frac{p}{2}) \Gamma(-p)}{\Gamma(-\frac{p}{2})}. \tag{31}$$

By using the above estimates for the signal $\mathbf{S}$ and the noise dispersion $\gamma$, we obtain the reduced optimization problem given in (32), which appears at the bottom of the page.

An iterative procedure based on the gradient descent principle can be applied in order to solve for $\boldsymbol{\theta}$. In general, the cost function described in (32) is nonconvex, and the optimization procedure has to be initialized sufficiently close to the global extremum. Suboptimal DOA estimators such as the MUSIC algorithm or the recently proposed ROC-MUSIC method [22] can be used to obtain initial bearing estimates. Concluding this section, we point out, once again, the two main assumptions made in order to obtain a closed-form expression for the signal estimate $\hat{\mathbf{S}}$:

B.1) Assumption $|x_i(t) - \sum_{k=1}^{q} a_i(\theta_k) s_k(t)| < \gamma$ (holds with probability $(\sqrt{2}-1)/\sqrt{2}$) enabled us to express the logarithmic function as an infinite series;

B.2) Assumption $|x_i(t) - \sum_{k=1}^{q} a_i(\theta_k) s_k(t)| \ll \gamma$ enabled us to take a first order approximation of the infinite series, and thus to obtain a least-squares closed form estimate for the signal.

Obviously, when assumptions B.1) and B.2) are not satisfied, the pseudo maximum likelihood processor will suffer from suboptimal signal estimation. For the general application case, when the processor does not possess enough information about the transmitted signals, the two assumptions B.1) and B.2) provide a closed-form expression for the signal estimates. The degree to which these assumptions affect the performance of the processor is studied in the simulated experiments, in Section V.

### B. The CRB for Cauchy Noise

Under the assumptions stated in Section II and for the case of complex isotropic Cauchy noise, the following theorem holds:

$$\max_{\boldsymbol{\theta}} L(\mathbf{X}; \hat{\gamma}, \hat{\mathbf{S}}, \boldsymbol{\theta}) = \max_{\boldsymbol{\theta}} \left\{ Mr \log(\hat{\gamma}) - \frac{3}{2} \sum_{t=1}^{M} \sum_{i=1}^{r} \log \left( \hat{\gamma}^2 + \left| x_i(t) - \sum_{k=1}^{q} a_i(\theta_k) \hat{s}_k(t) \right|^2 \right) \right\} \tag{32}$$

*Theorem 1:* The CRB for $\boldsymbol{\theta}$, and $\gamma$ is given by

$$\text{CRB}(\boldsymbol{\theta}) = \frac{5\gamma^2}{3}\left\{\sum_{t=1}^{M}\Re\left\{\mathbf{S}^H(t)\mathbf{D}^H\left[I-\mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H\right]\mathbf{D}\mathbf{S}(t)\right\}\right\}^{-1} \tag{33}$$

and

$$\text{CRB}(\gamma) = \frac{5}{4}\frac{\gamma^2}{Mr}, \tag{34}$$

where

$$\mathbf{S}(t) = \begin{bmatrix} s_1(t) & 0 & \cdots & 0 & 0 \\ 0 & s_2(t) & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & s_{q-1}(t) & 0 \\ 0 & 0 & \cdots & 0 & s_q(t) \end{bmatrix}; \; t = 1, \ldots, M \tag{35}$$

$$\mathbf{D} = [\mathbf{d}(\theta_1), \ldots, \mathbf{d}(\theta_q)], \tag{36}$$

$\mathbf{d}(\theta_i) = \partial\mathbf{a}(\theta_i)/\partial\theta_i$; $i = 1, \ldots, q$, and $\Re\{\cdot\}$ is the real part operator.

*Proof:* See Appendix A.

The above expression for the CRB is very similar to the one derived in [2] for the case of additive Gaussian noise. We should note that the above bound can be achieved only when there exist unbiased estimators of all the model parameters $\gamma$, $\mathbf{S}$, and $\boldsymbol{\theta}$. A useful insight on the CRB can be gained if we consider the case of a single source ($q = 1$) impinging from direction $\theta$ in a linear array whose sensors are spaced a half-wavelength apart. In this case

$$\mathbf{A} = [1, e^{-\jmath\pi\sin(\theta)}, \ldots, e^{-\jmath(r-1)\pi\sin(\theta)}]^T \tag{37}$$

and

$$\mathbf{D} = [0, -\jmath\pi\cos(\theta)e^{-\jmath\pi\sin(\theta)}, \ldots, -\jmath(r-1)\pi\cos(\theta)e^{-\jmath(r-1)\pi\sin(\theta)}]^T. \tag{38}$$

Therefore, it holds that

$$\mathbf{A}^H\mathbf{A} = r \tag{39}$$

$$\mathbf{D}^H\mathbf{D} = \pi^2\frac{r(r-1)(2r-1)}{6}\cos^2(\theta) \tag{40}$$

and

$$\mathbf{D}^H\mathbf{A} = \jmath\pi\frac{r(r-1)}{2}\cos(\theta). \tag{41}$$

Then

$$\text{CRB}(\theta) = \frac{20}{\pi^2}\cdot\frac{1}{r(r^2-1)\cos^2(\theta)}\cdot\frac{\gamma^2}{\sum_{t=1}^{M}|s(t)|^2}. \tag{42}$$

The term $\frac{\gamma^2}{\sum_{t=1}^{M}|s(t)|^2}$ in the above expression for the CRB can be viewed as the inverse of a quantity analogous to the signal-to-noise ratio (SNR) for the Gaussian case, i.e., a generalized SNR so to speak. The larger the dispersion $\gamma$ of the noise, the higher the CRB.

### C. ML Estimation for General $\alpha$-Stable Noise

In this section, we discuss the ML estimation problem in the presence of additive stable noise with characteristic exponent $\alpha$ in the interval $(1, 2)$. In this case, there are no closed-form expressions but only power series expansions for the noise distributions as shown in (16). Here, we present an approximate solution to the ML estimation problem based on the characteristic function of the bivariate $\alpha$-stable process given by (15).

The problem of approximating likelihood calculations has been studied in the past for applications where probability distributions are only conveniently represented by their transforms, like the characteristic function (see [25] and references therein). A natural approach for estimating density ($p(x;\theta)$) based functions, like the likelihood function $L(x;\theta) = \log p(x;\theta)$, the score function $S(x;\theta) = \frac{\partial \log p(x;\theta)}{\partial\theta}$, and the Fisher information $\mathcal{I}(\theta) = \text{var}\{S(x;\theta)\}$ is based on prediction theory: The best linear approximation of a function $h(x)$ of a random variable in terms of a kernel class of functions $G = \{g(\omega_i, x); \quad \omega_i \in (\omega_1, \ldots, \omega_k)\}$ is defined to be $\hat{h}(x) = \sum_{i=1}^{k} d_i g(\omega_i, x)$ minimizing the mean-square error $E\{(\hat{h}(x) - h(x))^2\}$. The well-known solution for this problem is given by

$$\hat{h}(x) = E\{h(x)\} + \boldsymbol{\lambda}^H\boldsymbol{\Sigma}^{-1}[\mathbf{g}(x) - \boldsymbol{\gamma}], \tag{43}$$

where $\mathbf{g}(x) = [g(\omega_1, x), \ldots, g(\omega_k, x)]^T$, $\boldsymbol{\lambda} = \text{cov}\{h(x), \mathbf{g}(x)\}$, $\boldsymbol{\Sigma} = \text{var}\{\mathbf{g}(x)\}$, and $\boldsymbol{\gamma} = E\{\mathbf{g}(x)\}$.

Approximations of the score function $S(x;\theta)$ and the Fisher information $\mathcal{I}(\theta)$ based on the above expression can be obtained as follows [25]: Under mild regularity conditions

$$E_\theta\{S(x;\theta)g(x)\} = \frac{\partial E_\theta\{g(x)\}}{\partial\theta} \tag{44}$$

for any function $g(x)$. Furthermore, $E_\theta\{S(x;\theta)\} = 0$, and hence, $\boldsymbol{\lambda} = \text{cov}\{S(x;\theta), \mathbf{g}(x)\} = \frac{\partial\boldsymbol{\gamma}(\theta)}{\partial\theta} = \dot{\boldsymbol{\gamma}}(\theta)$, where $\boldsymbol{\gamma}(\theta) = E_\theta\{\mathbf{g}(x)\}$. By using (43), we find that for any set of functions $\mathbf{g}(x)$, an approximation to the score function is given by

$$\hat{S}(x;\theta) = \dot{\boldsymbol{\gamma}}(\theta)^H\boldsymbol{\Sigma}^{-1}(\theta)[\mathbf{g}(x) - \boldsymbol{\gamma}(\theta)]. \tag{45}$$

According to standard arguments, the approximate expression for the Fisher information is given by

$$\mathcal{I}(\theta) = \dot{\boldsymbol{\gamma}}(\theta)^H\boldsymbol{\Sigma}^{-1}(\theta)\dot{\boldsymbol{\gamma}}(\theta). \tag{46}$$

Returning to the bearing estimation problem in the presence of additive complex isotropic $\alpha$-stable noise, we consider the case of a single source impinging on the array from direction $\theta$. Then, the likelihood function is given by

$$L(\mathbf{X};\gamma,\mathbf{S},\theta) = \sum_{t=1}^{M}\sum_{i=1}^{r}\log\chi_{\alpha,\gamma}(|x_i(t) - a_i(\theta)s(t)|). \tag{47}$$

The approximation to the ML equations based on the data samples at the array sensors can be written as

$$\hat{S}(\mathbf{X};\gamma,\mathbf{S},\theta) = \sum_{t=1}^{M}\sum_{i=1}^{r}\hat{S}_{(t,i)}(x_i(t);\gamma,s(t),\theta) = 0. \tag{48}$$

By choosing the kernel functions to be $g(\omega, x) = \exp[\jmath\Re\{\omega x\}]$, the approximation of the score functions $\hat{S}_{(t,i)}(x_i(t); \gamma, s(t), \theta)$ will depend on the characteristic functions $\phi_{(t,i)}(\omega; \gamma, s(t), \theta)$. Using (15), we can write

$$\phi_{(t,i)}(\omega; \gamma, s(t), \theta) = \exp[\jmath(\omega_\Re \Re\{a_i(\theta)s(t)\} + \omega_\Im \Im\{a_i(\theta)s(t)\}) - \gamma|\omega|^\alpha] \tag{49}$$

where $\omega = \omega_\Re + \jmath\omega_\Im$, and $\Re\{\cdot\}$ and $\Im\{\cdot\}$ are the real and imaginary part operators. Evaluation in a complex grid $\mathbf{\Omega} = \boldsymbol{\omega}_\Re \times \boldsymbol{\omega}_\Im$, where $\boldsymbol{\omega}_\Re = \boldsymbol{\omega}_\Im = [-k\tau, \ldots, -\tau, \tau, \ldots, k\tau]^T$ and application of (45) give the following expression for $\hat{S}(\mathbf{X}; \theta)$:

$$\hat{S}(\mathbf{X}; \theta) = \sum_{t=1}^{M}\sum_{i=1}^{r} [\jmath\mathbf{D}_\Omega(\theta)\boldsymbol{\phi}_{(t,i)}(\mathbf{\Omega}, \theta)]^H \mathbf{\Sigma}_\Omega^{-1}(\theta)[\mathbf{g}(\mathbf{\Omega}, x_i(t)) - \boldsymbol{\phi}_{(t,i)}(\mathbf{\Omega}, \theta)]. \tag{50}$$

In addition, because of the independence of the observed data, and by using (46), the approximate Fisher information is given by

$$\mathcal{I}(\theta) = \sum_{t=1}^{M}\sum_{i=1}^{r} \boldsymbol{\phi}_{(t,i)}^H(\mathbf{\Omega},\theta)\, \mathbf{D}_\Omega^H(\theta)\, \mathbf{\Sigma}_\Omega^{-1}(\theta)\mathbf{D}_\Omega(\theta)\, \boldsymbol{\phi}_{(t,i)}(\mathbf{\Omega},\theta). \tag{51}$$

The vectors and matrices in (50) and (51) are defined as follows:

$$\mathbf{g}(\mathbf{\Omega}, x_i(t)) = \begin{bmatrix} \exp[\jmath(-k\tau x_{i,\Re}(t) - k\tau x_{i,\Im}(t))] \\ \exp[\jmath(-k\tau x_{i,\Re}(t) + (-k+1)\tau x_{i,\Im}(t))] \\ \vdots \\ \exp[\jmath(-k\tau x_{i,\Re}(t) + k\tau x_{i,\Im}(t))] \\ \exp[\jmath((-k+1)\tau x_{i,\Re}(t) - k\tau x_{i,\Im}(t))] \\ \vdots \\ \exp[\jmath(k\tau x_{i,\Re}(t) + k\tau x_{i,\Im}(t))] \end{bmatrix}_{4k^2\times 1} \tag{52}$$

where $x_{i,\Re}(t) = \Re\{x_i(t)\}$, and $x_{i,\Im}(t) = \Im\{x_i(t)\}$. In addition, we have (53), which appears at the bottom of the page. The matrix $\mathbf{D}_\Omega(\theta)$ is the $4k^2 \times 4k^2$ diagonal matrix

$$\begin{aligned}\mathbf{D}_\Omega(\theta) = \mathrm{diag}[&-k\tau\Re\{d_i(\theta)s(t)\} - k\tau\Im\{d_i(\theta)s(t)\}, \\ &- k\tau\Re\{d_i(\theta)s(t)\} + (-k+1)\tau\Im\{d_i(\theta)s(t)\}, \ldots, \\ &- k\tau\Re\{d_i(\theta)s(t)\} + k\tau\Im\{d_i(\theta)s(t)\}, \ldots, \\ &k\tau\Re\{d_i(\theta)s(t)\} + k\tau\Im\{d_i(\theta)s(t)\}].\end{aligned} \tag{54}$$

Finally, $\mathbf{\Sigma}_\Omega(\theta)$ is the $4k^2 \times 4k^2$ covariance matrix $\mathrm{cov}\{\mathbf{g}(\mathbf{\Omega}, x_i(t))\}$.

We can make the following remarks concerning the described approximation.

*Remark 1:* In this section, we considered only the case of a single source impinging on the array. The generalization to multiple sources of a known number is a conceptually straightforward problem, but it is one that involves a considerable computational load since it requires the solution of a nonlinear system of equations of the form (50).

*Remark 2:* In (50), we considered the approximation of the derivatives of the likelihood function only with respect to the parameter of interest $\theta$, i.e., we assumed that the noise dispersion $\gamma$ and the signal $S$ are known or can be estimated via some method other than ML (see also Section IV-A for the Cauchy case).

*Remark 3:* The kernel class $G = \{\exp[\jmath\Re\{\omega x\}];\ \omega \in \mathcal{C}\}$ is complete, and therefore, an approximation $\hat{h}(x)$ of any function can be made arbitrarily close to $h(x)$. In practice, we use a finite number of functions indexed by $\mathbf{\Omega}$. For the case of a linear array of sensors and a single impinging source from direction $\theta$, the likelihood function is periodic with respect to $\theta$, and its domain lies in the interval $[-\pi/2, \pi/2]$. If the domain of interest is the interval $[-F, F]$, an appropriate choice for $\tau$ is $\tau = 2\pi/F$. The approximation will depend on the values of the characteristic function at the points of the grid $\mathbf{\Omega}$. This approach is basically equivalent to sampling the probability density function at the nodes of an orthogonal grid with spacing $2\pi/\tau$. The computational complexity associated with the method is increasing with order $\mathcal{O}(k^2)$. We found that a reasonable choice for $k$ is between 5 and 10 for the general case of the $\alpha$-stable law ($1 \le \alpha < 2$).

## V. A Performance Study

To demonstrate the performance of the proposed method for the direction-of-arrival (DOA) problem, we conducted several simulation experiments where we compared the ML estimator based on the Cauchy noise assumption (MLC) with the ML estimator based on the Gaussian noise assumption (MLG) and with the MUSIC estimator.

In all the experiments, the array is linear with five sensors spaced a half-wavelength apart. A single signal impinges to the array from a source located at a direction of $5^o$. At first, the signal is assumed to be known at the receiver, and the ML method is applied to estimate the direction of arrival. In addition, for the unknown signal case, the pseudo ML method

$$\boldsymbol{\phi}_{(t,i)}(\mathbf{\Omega}, \theta) = \begin{bmatrix} \exp[\jmath(-k\tau\Re\{a_i(\theta)s(t)\} - k\tau\Im\{a_i(\theta)s(t)\}) - \gamma[(-k\tau)^2 + (-k\tau)^2]^{\frac{\alpha}{2}}] \\ \exp[\jmath(-k\tau\Re\{a_i(\theta)s(t)\} + (-k\tau+1)\Im\{a_i(\theta)s(t)\}) - \gamma[(-k\tau)^2 + ((-k+1)\tau)^2]^{\frac{\alpha}{2}}] \\ \vdots \\ \exp[\jmath(-k\tau\Re\{a_i(\theta)s(t)\} + k\tau\Im\{a_i(\theta)s(t)\}) - \gamma[(-k\tau)^2 + (k\tau)^2]^{\frac{\alpha}{2}}] \\ \exp[\jmath((-k+1)\tau\Re\{a_i(\theta)s(t)\} - k\tau\Im\{a_i(\theta)s(t)\}) - \gamma[((-k+1)\tau)^2 + (-k\tau)^2]^{\frac{\alpha}{2}}] \\ \vdots \\ \exp[\jmath(k\tau\Re\{a_i(\theta)s(t)\} + k\tau\Im\{a_i(\theta)s(t)\}) - \gamma[(k\tau)^2 + (k\tau)^2]^{\frac{\alpha}{2}}] \end{bmatrix}_{4k^2\times 1}. \tag{53}$$

TABLE I
GSNR AND AVERAGE PSNR FOR DIFFERENT VALUES OF $M$

| | *Number of snapshots, M* | | | | |
|---|---|---|---|---|---|
| | $M=5$ | $M=10$ | $M=20$ | $M=50$ | $M=100$ |
| *GSNR [dB]* | 22.6951 | 22.7323 | 22.7416 | 22.5003 | 22.6543 |
| *PSNR [dB]* | 7.3712 | 3.3938 | 0.5371 | -4.3158 | -7.149 |

is applied by using a least-squares estimate for the signal. The noise is assumed to follow the bivariate isotropic stable distribution. We chose $p$ to be $p = 0.5$ when we estimated the noise dispersion $\gamma$ via expression (30). In theory, (17) holds for any value of $p$ as long as $0 < p < \alpha$. In practice, the variance of the estimator is a function of $p$. It is shown in [22] that values of $p$ in the range $[1/2, \alpha/2]$ give estimators with the smallest variance.

In every experiment, we perform 500 Monte-Carlo runs and compute the mean square error (MSE) of the DOA estimates. The optimization for the MLC and MLG methods is performed by a steepest-descent algorithm with variable stepsize selected by means of Armijo's rule [26]. Since the alpha-stable family for $\alpha < 2$ determines processes with infinite variance, we define two alternative SNR's. Namely, we define the *generalized SNR (GSNR)* to be the ratio of the signal power over the noise dispersion $\gamma$:

$$GSNR = 10\log(\frac{1}{\gamma M}\sum_{t=1}^{M}|s(t)|^2). \quad (55)$$

In addition, for finite sample realizations, we define the *pseudo-SNR (PSNR)* as follows:

$$\text{PSNR} = 10\log(\frac{\sum_{t=1}^{M}|s(t)|^2}{\sum_{t=1}^{M}|n(t)|^2}). \quad (56)$$

In the following, we present a comparative simulation study for the estimation accuracy and the probability of convergence of the aforementioned algorithms.

### A. Estimation Accuracy

In this example, we study the estimation accuracy of MLC, MLG, and MUSIC as a function of three parameters, namely, the number of snapshots $M$, the noise dispersion $\gamma$, and the noise characteristic exponent $\alpha$.

*Number of Snapshots* M*:* In the first experiment, we study the influence of the number of snapshots $M$ to the performance of the algorithms. The noise follows the complex isotropic Cauchy distribution with dispersion $\gamma = 1$ (cf. (16)). For this experiment, the GSNR is kept almost constant at 22.5 dB as shown in Table I. The PSNR is different in every Monte-Carlo run; therefore, we calculate the average PSNR over the 500 Monte-Carlo runs (cf. Table I). As the number of snapshots $M$ increases, the PSNR decreases because increasingly impulsive noise samples are incorporated into the data.

Fig. 1(a) shows the resulting MSE of the estimated DOA as a function of the number of snapshots when the signal is known. The CRB is also plotted. As expected, the MLC estimator has the best performance since it is the optimal estimator for this type of noise. In addition, the complete failure of the MLC and MUSIC processors for this type of impulsive noise is apparent. Fig. 1(b) shows similar plots for the case of the pseudo ML estimators where we use a LS estimate for the signal. The MLC estimator again has the least MSE. Comparing these curves with the analogous curves obtained assuming exact signal knowledge, we observe a larger MSE for the pseudo ML estimates, as expected.

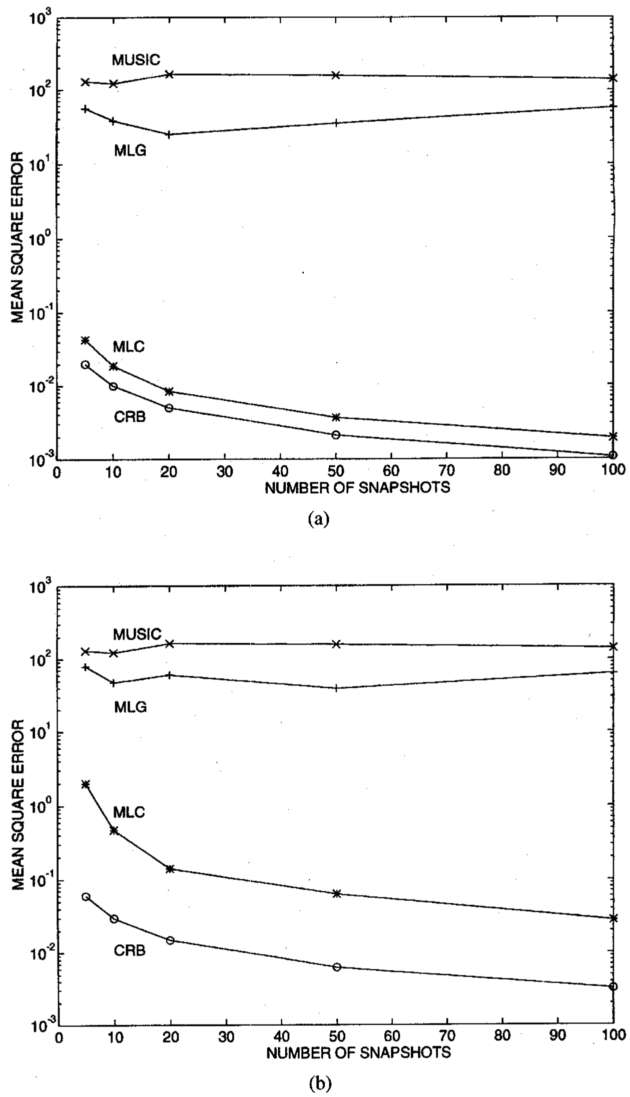



Fig. 1. MSE of the estimated DOA and CRB as functions of the number of snapshots $M$: (a) Exact signal knowledge; (b) least-squares estimate of the signal.

*Noise Dispersion* $\gamma$*:* In the second experiment, we study the influence of the noise dispersion $\gamma$, i.e., the influence of the GSNR to the performance of the methods. Here, the noise follows the bivariate isotropic Cauchy distribution with dispersion $\gamma$. The number of snapshots available to the algorithms is $M = 20$. The GSNR and average PSNR for this experiment are shown in Table II.

Fig. 2 shows the resulting MSE of the estimated DOA as a function of the GSNR. Again, the MLC estimate has the best performance. As evident in Fig. 2(b), the performance of the MLC using a LS estimate for the signal degrades more rapidly

TABLE II
GSNR AND AVERAGE PSNR FOR DIFFERENT VALUES OF $\gamma$

| | Noise Dispersion, $\gamma$ | | | | | | |
|---|---|---|---|---|---|---|---|
| | $\gamma = 0.5$ | $\gamma = 1$ | $\gamma = 2$ | $\gamma = 4$ | $\gamma = 6$ | $\gamma = 8$ | $\gamma = 10$ |
| *GSNR [dB]* | 25.7519 | 22.7416 | 19.7313 | 16.7210 | 14.9601 | 13.7107 | 12.7416 |
| *PSNR [dB]* | 6.5577 | 0.5371 | -5.4835 | -11.5041 | -15.0259 | -17.5247 | -19.4629 |

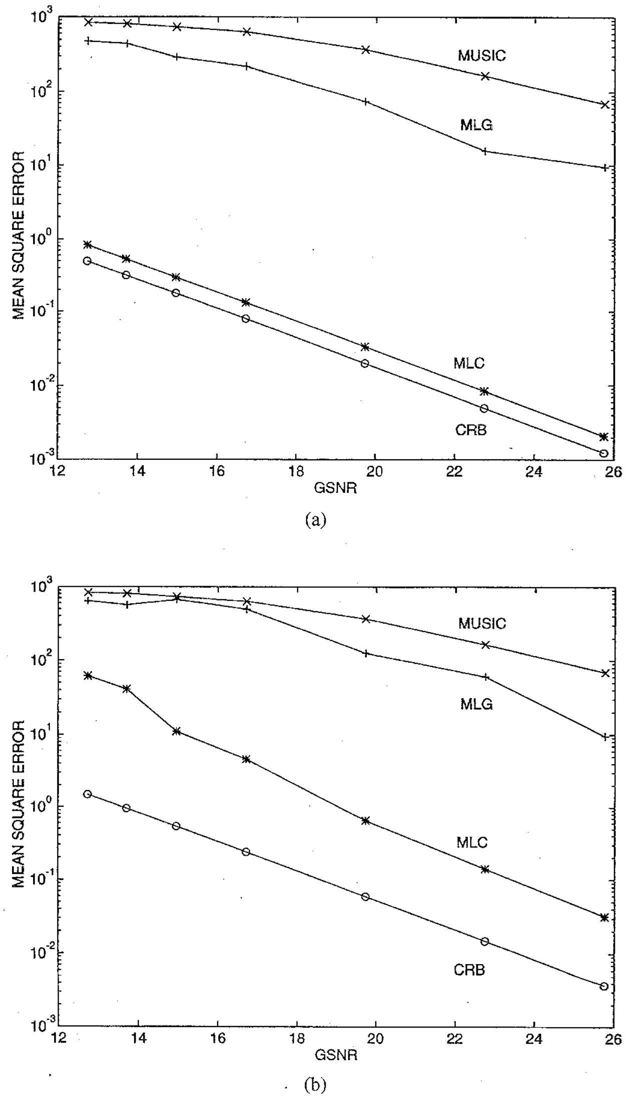



Fig. 2. MSE of the estimated DOA and CRB as functions of the GSNR: (a) Exact signal knowledge; (b) least-squares estimate of the signal.

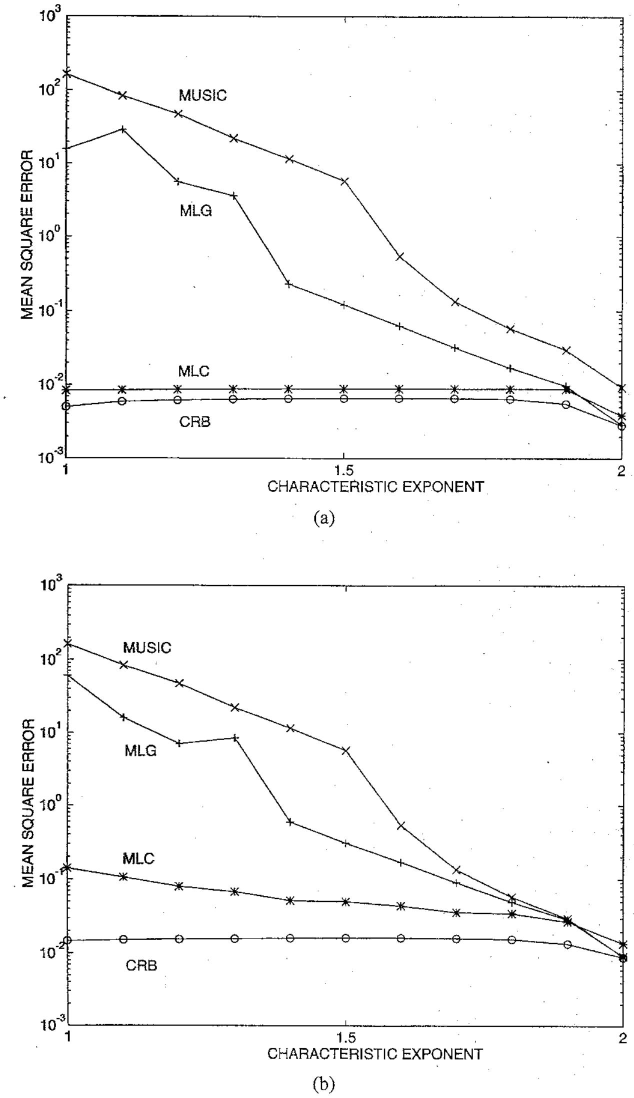



Fig. 3. MSE of the estimated DOA and CRB as functions of the characteristic exponent $\alpha$: (a) Exact signal knowledge; (b) least-squares estimate of the signal.

for large values of the noise dispersion $\gamma$ (low GSNR values). The reason is that for large values of the noise dispersion, there is a small probability that assumption B.2) made in Section IV-A holds. Thus, the MLC estimate suffers from suboptimal signal estimation. It is noted, however, that the MLC estimate still has the best performance.

*Characteristic Exponent $\alpha$:* The importance of this experiment rests in its study of the robustness of the algorithms in different noise environments. Of course, by design, the MLG estimator is optimal for additive Gaussian noise ($\alpha = 2$), and the introduced MLC estimator is optimal for additive Cauchy noise ($\alpha = 1$). An important property of any processor is to be able to perform reasonably well in a wide range of noise environments ($1 < \alpha < 2$). Here, we test the performance of the estimators when the characteristic exponent $\alpha$ of the noise stable law is changing.

Fig. 3 shows the resulting MSE curves as functions of the characteristic exponent $\alpha$. The number of snapshots available to the algorithms is $M = 20$. The GSNR is 22.7416 dB ($\gamma = 1$) and is shown together with the average PSNR on Table III. The CRB as a function of $\alpha$ was computed by means of (51), where we used the values $k = 10$, and $\tau = 0.157$.

As we can clearly see, the Cauchy beamformer is practically insensitive to the changes of $\alpha$, and for exact signal knowledge, it almost achieves the CRB for the whole range of values $\alpha$. On the other hand, both the MLG and the MUSIC algorithms exhibit very large mean-square estimation errors for non-Gaussian noise environments. Note that when $\alpha = 2$, i.e., for the Gaussian noise case, the MLG method has the least MSE, as expected.

The experiment demonstrates that for values of $\alpha$ in [1,2], the ML method based on the Cauchy noise assumption exhibits performance very close to the optimum. On the other hand, the optimum ML estimator for the general $S\alpha S$ case with $\alpha \in (1,2)$ involves the computationally intensive solution of $\hat{S}(\mathbf{X};\theta) = 0$ with $\hat{S}(\mathbf{X};\theta)$ given by (50). These observations, combined with the fact that the ML method based on the Cauchy assumption has computational complexity similar to the ML method based on the Gaussian assumption, justify the importance of the Cauchy beamformer for the DOA estimation problem in practice.

### B. Initialization and Convergence

In this example, we study the convergence of the MLC and MLG algorithms to the true DOA value as a function of the characteristic exponent $\alpha$ of the noise. The initialization of the optimization procedure is done by means of the MUSIC bearing estimates. Again, the number of snapshots available to the algorithms is $M = 20$, and the GSNR and average PSNR are as shown in Table III. In Fig. 4, we plot the probability that the processor will converge within $1^{o}$ from the true DOA as a function of the characteristic exponent of the noise. It is shown that for the known signal case, MLC exhibits higher probability of convergence than MLG in the range $1 \leq \alpha \leq 1.6$. The range of the superior performance of MLC is $1 \leq \alpha \leq 1.8$ when a LS estimate for the signal waveforms is used.

## VI. Concluding Remarks

We have presented a novel approach to the DOA estimation problem in the presence of impulsive interference. The method is based on the maximum likelihood estimation technique where the noise is modeled as a complex isotropic $S\alpha S$ process. The Cauchy beamformer has been shown to give better bearing estimates than the Gaussian beamformer in a wide range of impulsive noise environments and for very low SNR values. The technique inherits the computational complexity of the ML family of methods. The simulations show that for very large values of the noise dispersion, the Cauchy beamformer suffers when the LS estimate for the signal is used. Still, its performance is better than the performance of existing methods. Better estimates for the signal can be obtained by considering special application scenarios, as is the subject of current research. In addition, we will extend the subspace-based techniques for bearing estimation to processes with finite moments of order $p$ ($p < 2$) and to complex isotropic $S\alpha S$ processes. This is particularly useful for direction finding applications in the presence of impulsive noise when reduced computational cost is a crucial design requirement. Furthermore, future research includes the development of methods for detection of the number of signals in the presence of impulsive noise under an additive model. Finally, we will extend our present work to the area of matched field processing (MFP) [27] in ocean acoustics by incorporating the complexities of the ocean environment and by exploiting the field structure of the signals propagating in an ocean waveguide.

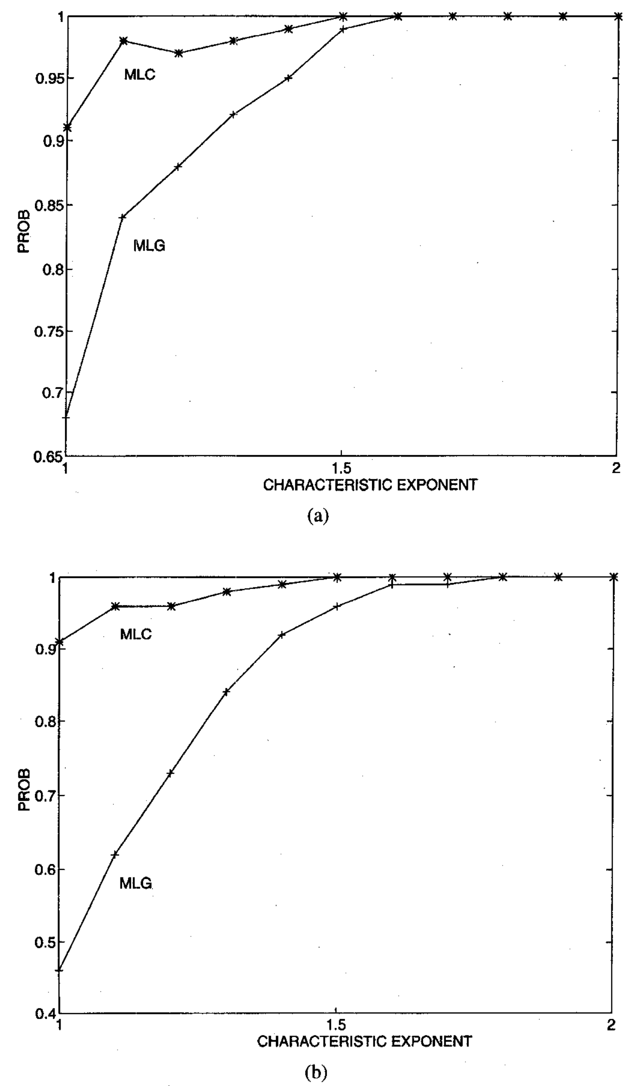



Fig. 4. Probability of convergence within $1^{o}$ of the true DOA as a function of the characteristic exponent $\alpha$: (a) Exact signal knowledge; (b) least-squares estimate of the signal.

TABLE III
GSNR and Average PSNR for Different Values of $\alpha$

| | *Noise Characteristic Exponent,* $\alpha$ | | | | | |
|---|---|---|---|---|---|---|
| | $\alpha = 1.0$ | $\alpha = 1.2$ | $\alpha = 1.4$ | $\alpha = 1.6$ | $\alpha = 1.8$ | $\alpha = 2.0$ |
| *GSNR [dB]* | 22.7416 ($\gamma = 1$) | | | | | |
| *PSNR [dB]* | 0.5371 | 6.2171 | 10.1869 | 13.1035 | 15.3269 | 20.0383 |

## Appendix A
## Derivation of CRB For Complex Isotropic Cauchy Noise

Here, we derive the CRB for the most general case of multiple signal sources in the presence of complex isotropic Cauchy noise of unknown dispersion. As a first step, we derive a useful proposition about the magnitude $\rho$ and phase $\phi$ of the noise:

*Proposition 1:* The amplitude $\rho$ and phase $\phi$ of the noise following the complex isotropic Cauchy distribution are independent, and the magnitude $\rho$ satisfies

$$E\left\{\frac{\rho^2}{(\gamma^2+\rho^2)^2}\right\} = \frac{2}{15\gamma^2}. \qquad (57)$$

*Proof:* The complex noise samples $n_i(t) = |n_i(t)|e^{\jmath\phi_i(t)}$; $i = 1,\ldots,r$; $t = 1,\ldots,M$ follow the bivariate isotropic Cauchy distribution with dispersion $\gamma$. In other words, the real and imaginary parts of the noise $n = n_{\Re} + \jmath n_{\Im}$ are jointly Cauchy with probability density $f_1(n_{\Re}, n_{\Im})$ given by

$$f_1(n_{\Re}, n_{\Im}) = \frac{1}{2\pi}\frac{\gamma}{(n_{\Re}^2 + n_{\Im}^2 + \gamma^2)^{3/2}} \quad -\infty < n_{\Re}, n_{\Im} < \infty. \qquad (58)$$

As we can see, the probability density $f_1(n_{\Re}, n_{\Im})$ can be expressed as a function of the noise magnitude $\rho = |n| = \sqrt{n_{\Re}^2 + n_{\Im}^2}$,

$$f_1(n_{\Re}, n_{\Im}) = \chi_1(\rho) = \frac{1}{2\pi}\frac{\gamma}{(\rho^2+\gamma^2)^{3/2}} \quad \rho \geq 0. \qquad (59)$$

The joint density of the noise magnitude $\rho$ and phase $\phi = \arctan(n_{\Im}/n_{\Re})$ is given by [28]

$$f(\rho,\phi) = \rho f_1(\rho\cos\phi, \rho\sin\phi) \quad \rho \geq 0, \ \ \phi \in [0, 2\pi]. \qquad (60)$$

From (59), it follows that

$$f(\rho,\phi) = \frac{1}{2\pi}\frac{\gamma\rho}{(\rho^2+\gamma^2)^{3/2}} \quad \rho \geq 0, \ \ \phi \in [0, 2\pi]. \qquad (61)$$

Since the density $f(\rho,\phi)$ is independent of $\phi$, the noise phase $\phi$ is uniformly distributed in $[0, 2\pi]$ and is independent of the magnitude $\rho$. Hence

$$f(\rho) = \frac{\gamma\rho}{(\rho^2+\gamma^2)^{3/2}} \quad \rho \geq 0 \qquad (62)$$

and

$$f(\phi) = \frac{1}{2\pi} \quad \phi \in [0, 2\pi]. \qquad (63)$$

Now, by using (62), we get

$$E\left\{\frac{\rho^2}{(\gamma^2+\rho^2)^2}\right\} = \int_0^{+\infty}\frac{\rho^2}{(\gamma^2+\rho^2)^2}\frac{\gamma\rho}{(\gamma^2+\rho^2)^{3/2}}d\rho = \gamma\int_0^{+\infty}\frac{\rho^3}{(\gamma^2+\rho^2)^{7/2}}d\rho. \qquad (64)$$

A general form of the above integral can be found in [29]

$$\int_0^{+\infty}\frac{\rho^3}{(\gamma^2+\rho^2)^{7/2}}d\rho = \frac{2}{15\gamma^3} \qquad (65)$$

and (57) follows. ∎

Now, we proceed by proving the theorem. Let $\boldsymbol{\eta}$ represent the $(2Mq+q+1)$-dimensional vector of unknown parameters in the stochastic model:

$$\begin{aligned}\boldsymbol{\eta} =& [\gamma, s_{1,\Re}(1),\ldots,s_{q,\Re}(1), s_{1,\Im}(1),\ldots,s_{q,\Im}(1),\ldots \\ & s_{1,\Re}(M),\ldots,s_{q,\Re}(M), s_{1,\Im}(M),\ldots,s_{q,\Im}(M), \boldsymbol{\theta}^T]^T \\ =& [\gamma, \mathbf{s}_{\Re}^T(1), \mathbf{s}_{\Im}^T(1),\ldots,\mathbf{s}_{\Re}^T(M), \mathbf{s}_{\Im}^T(M), \boldsymbol{\theta}^T]^T \end{aligned} \qquad (66)$$

where $\mathbf{s}_{\Re}(t) = \Re\{\mathbf{s}(t)\} = [s_{1,\Re}(t),\ldots,s_{q,\Re}(t)]^T$, $\mathbf{s}_{\Im}(t) = \Im\{\mathbf{s}(t)\} = [s_{1,\Im}(t),\ldots,s_{q,\Im}(t)]^T$; $t = 1,\ldots,M$, and $\boldsymbol{\theta} = [\theta_1,\ldots,\theta_q]^T$ is the vector of the unknown directions of arrival.

The Fisher information matrix $\mathbf{J}(\boldsymbol{\eta})$ is defined as

$$\mathbf{J}(\boldsymbol{\eta}) = E\left\{\left(\frac{\partial L(\mathbf{X};\gamma,\mathbf{S},\boldsymbol{\theta})}{\partial\boldsymbol{\eta}}\right)\left(\frac{\partial L(\mathbf{X};\gamma,\mathbf{S},\boldsymbol{\theta})}{\partial\boldsymbol{\eta}}\right)^T\right\}. \qquad (67)$$

First, we calculate the derivatives of the log likelihood function given in (21) with respect to the components of $\boldsymbol{\eta}$. We have

$$\frac{\partial L}{\partial\theta_k} = 3\sum_{t=1}^{M}\sum_{i=1}^{r}\frac{\Re\{s_k^*(t)d_i^*(\theta_k)n_i(t)\}}{\gamma^2+|n_i(t)|^2}; \quad k = 1,\ldots,q \qquad (68)$$

where $d_i(\theta_k) = \frac{\partial a_i(\theta_k)}{\partial\theta_k}$. In addition

$$\frac{\partial L}{\partial\gamma} = \frac{Mr}{\gamma} - 3\gamma\sum_{t=1}^{M}\sum_{i=1}^{r}\frac{1}{\gamma^2+|n_i(t)|^2} \qquad (69)$$

$$\frac{\partial L}{\partial s_{k,\Re}(t)} = 3\sum_{i=1}^{r}\frac{\Re\{a_i^*(\theta_k)n_i(t)\}}{\gamma^2+|n_i(t)|^2}; \quad k = 1,\ldots,q; \ \ t = 1,\ldots,M \qquad (70)$$

and

$$\frac{\partial L}{\partial s_{k,\Im}(t)} = 3\sum_{i=1}^{r}\frac{\Im\{a_i^*(\theta_k)n_i(t)\}}{\gamma^2+|n_i(t)|^2}; \quad k = 1,\ldots,q; \ \ t = 1,\ldots,M. \qquad (71)$$

In the following derivations, we will extensively use assumption A.4), which states that the noise samples are spatially and temporally independent and Proposition 1, which states that the noise phase is uniformly distributed in $[0, 2\pi]$ in order to get simplifications in the expressions. First, we have

$$\begin{aligned} E\left(\frac{\partial L}{\partial\theta_k}\frac{\partial L}{\partial\theta_l}\right) =& 9E \\ &\left\{\sum_{t=1}^{M}\sum_{t'=1}^{M}\sum_{i=1}^{r}\sum_{j=1}^{r}\frac{\Re\{s_k^*(t)d_i^*(\theta_k)n_i(t)\}}{\gamma^2+|n_i(t)|^2}\frac{\Re\{s_l^*(t')d_j^*(\theta_l)n_j(t')\}}{\gamma^2+|n_j(t')|^2}\right\} \\ =& 9E\left\{\sum_{t=1}^{M}\sum_{i=1}^{r}\frac{\Re\{s_k^*(t)d_i^*(\theta_k)n_i(t)\}}{\gamma^2+|n_i(t)|^2}\frac{\Re\{s_l^*(t)d_i^*(\theta_l)n_i(t)\}}{\gamma^2+|n_i(t)|^2}\right\}. \end{aligned} \qquad (72)$$

Setting $n_i(t) = |n_i(t)|e^{\jmath\phi_i(t)}$, $s_k(t) = |s_k(t)|e^{\jmath\beta_k(t)}$, and $a_i^*(\theta_k) = e^{\jmath\omega_i(\theta_k)}$, (72) can be written as

$$\begin{aligned} E\left(\frac{\partial L}{\partial\theta_k}\frac{\partial L}{\partial\theta_l}\right) =& 9E \\ &\left\{\sum_{t=1}^{M}\sum_{i=1}^{r}|s_k(t)d_i(\theta_k)|\frac{|n_i(t)|}{\gamma^2+|n_i(t)|^2}\right. \\ &\cos(-\beta_k(t)+\omega_i(\theta_k)+\frac{\pi}{2}+\phi_i(t)) \\ &\left.|s_l(t)d_i(\theta_l)|\frac{|n_i(t)|}{\gamma^2+|n_i(t)|^2}\cos(-\beta_l(t)+\omega_i(\vartheta_l)+\frac{\pi}{2}+\phi_i(t))\right\} \end{aligned}$$

$$= 9\sum_{t=1}^{M}\sum_{i=1}^{r}|s_k(t)d_i(\theta_k)||s_l(t)d_i(\theta_l)|E\left\{\frac{|n_i(t)|^2}{(\gamma^2+|n_i(t)|^2)^2}\right\}$$
$$\frac{1}{2}\cos(\beta_k(t)-\beta_l(t)-\omega_i(\theta_k)+\omega_i(\theta_l))$$
$$= \frac{3}{5\gamma^2}\sum_{t=1}^{M}\sum_{i=1}^{r}\Re\{s_k(t)s_l^*(t)d_i(\theta_k)d_i^*(\theta_l)\};\ k,l=1,\ldots,q. \quad (73)$$

Equation (73) can be written compactly as

$$E\left(\frac{\partial L}{\partial \boldsymbol{\theta}}\right)\left(\frac{\partial L}{\partial \boldsymbol{\theta}}\right)^T = \frac{3}{5\gamma^2}\sum_{t=1}^{M}\Re\{\mathbf{S}^H(t)\mathbf{D}^H\mathbf{D}\mathbf{S}(t)\}. \quad (74)$$

Now

$$E\left(\frac{\partial L}{\partial \gamma}\right)^2$$
$$= \frac{M^2r^2}{\gamma^2} - 6Mr\sum_{t=1}^{M}\sum_{i=1}^{r}E\left\{\frac{1}{\gamma^2+|n_i(t)|^2}\right\}$$
$$+ 9\gamma^2\sum_{t=1}^{M}\sum_{i=1}^{r}E\left\{\frac{1}{(\gamma^2+|n_i(t)|^2)^2}\right\}$$
$$+ 9\gamma^2\underbrace{\sum_{t=1}^{M}\sum_{t'=1}^{M}\sum_{i=1}^{r}\sum_{j=1}^{r}}_{t'\neq t \text{ or } j\neq i}E\left\{\frac{1}{\gamma^2+|n_i(t)|^2}\right\}$$
$$E\left\{\frac{1}{\gamma^2+|n_j(t')|^2}\right\}$$
$$= \frac{M^2r^2}{\gamma^2} - 6M^2r^2E\left\{\frac{1}{\gamma^2+|n_i(t)|^2}\right\}$$
$$+ 9\gamma^2MrE\left\{\frac{1}{(\gamma^2+|n_i(t)|^2)^2}\right\}$$
$$+ 9\gamma^2Mr(Mr-1)\left(E\left\{\frac{1}{\gamma^2+|n_i(t)|^2}\right\}\right)^2$$
$$= \frac{M^2r^2}{\gamma^2} - 6M^2r^2\frac{1}{3\gamma^2}$$
$$+ 9\gamma^2Mr\frac{1}{5\gamma^4} + 9\gamma^2Mr(Mr-1)\frac{1}{9\gamma^4}. \quad (75)$$

Therefore

$$E\left(\frac{\partial L}{\partial \gamma}\right)^2 = \frac{4}{5}\frac{Mr}{\gamma^2}. \quad (76)$$

In addition

$$E\left(\frac{\partial L}{\partial s_{k,\Re}(t)}\frac{\partial L}{\partial s_{l,\Re}(t')}\right)$$
$$= 9E\left\{\sum_{i=1}^{r}\sum_{j=1}^{r}\frac{\Re\{a_i^*(\theta_k)n_i(t)\}}{\gamma^2+|n_i(t)|^2}\frac{\Re\{a_j^*(\theta_l)n_j(t')\}}{\gamma^2+|n_j(t')|^2}\right\}$$
$$= 9E\left\{\sum_{i=1}^{r}\frac{\Re\{a_i^*(\theta_k)n_i(t)\}}{\gamma^2+|n_i(t)|^2}\frac{\Re\{a_i^*(\theta_l)n_i(t')\}}{\gamma^2+|n_i(t')|^2}\right\}$$
$$+ 9E\left\{\sum_{i=1}^{r}\sum_{j=1,j\neq i}^{r}\frac{\Re\{a_i^*(\theta_k)n_i(t)\}}{\gamma^2+|n_i(t)|^2}\frac{\Re\{a_i^*(\theta_l)n_j(t')\}}{\gamma^2+|n_j(t')|^2}\right\}$$
$$= 9\sum_{i=1}^{r}E\left\{\frac{|n_i(t)|^2}{(\gamma^2+|n_i(t)|^2)^2}\right\}$$
$$E\{\cos(\omega_i(\theta_k)+\phi_i(t))\cos(\omega_i(\theta_l)+\phi_i(t))\}\delta_{t,t'}$$
$$+ 9\sum_{i=1}^{r}\sum_{j=1,j\neq i}^{r}E\left\{\frac{|n_i(t)|}{\gamma^2+|n_i(t)|^2}\right\}E\{\cos(\omega_i(\theta_k)+\phi_i(t))\}$$
$$E\left\{\frac{|n_j(t')|}{\gamma^2+|n_j(t')|^2}\right\}E\{\cos(\omega_j(\theta_l+\phi_j(t'))\}$$
$$= 9\frac{2}{15\gamma^2}\frac{1}{2}\sum_{i=1}^{r}\cos(\omega_i(\theta_k)-\omega_i(\theta_l))\delta_{t,t'}$$
$$= \frac{3}{5\gamma^2}\sum_{i=1}^{r}\Re\{a_i(\theta_k)a_i^*(\theta_l)\}\delta_{t,t'}. \quad (77)$$

Equation (77) can be written compactly as

$$E\left(\frac{\partial L}{\partial \mathbf{s}_{\Re}(t)}\right)\left(\frac{\partial L}{\partial \mathbf{s}_{\Re}(t')}\right)^T = \frac{3}{5\gamma^2}\Re\{\mathbf{A}^H\mathbf{A}\}\delta_{t,t'}. \quad (78)$$

Similarly

$$E\left(\frac{\partial L}{\partial \mathbf{s}_{\Re}(t)}\right)\left(\frac{\partial L}{\partial \mathbf{s}_{\Im}(t')}\right)^T = -\frac{3}{5\gamma^2}\Im\{\mathbf{A}^H\mathbf{A}\}\delta_{t,t'}, \quad (79)$$

and

$$E\left(\frac{\partial L}{\partial \mathbf{s}_{\Im}(t)}\right)\left(\frac{\partial L}{\partial \mathbf{s}_{\Im}(t')}\right)^T = \frac{3}{5\gamma^2}\Re\{\mathbf{A}^H\mathbf{A}\}\delta_{t,t'}. \quad (80)$$

In addition, we have (81), which appears at the top of the next page, which in compact form can be written as

$$E\left(\frac{\partial L}{\partial \mathbf{s}_{\Re}(t)}\right)\left(\frac{\partial L}{\partial \boldsymbol{\theta}}\right)^T = \frac{3}{5\gamma^2}\Re\{\mathbf{A}^H\mathbf{D}\mathbf{S}(t)\}. \quad (82)$$

Similarly

$$E\left(\frac{\partial L}{\partial \mathbf{s}_{\Im}(t)}\right)\left(\frac{\partial L}{\partial \boldsymbol{\theta}}\right)^T = \frac{3}{5\gamma^2}\Im\{\mathbf{A}^H\mathbf{D}\mathbf{S}(t)\}. \quad (83)$$

Finally,

$$E\left(\frac{\partial L}{\partial \gamma}\right)\left(\frac{\partial L}{\partial \boldsymbol{\theta}}\right)^T = 0 \quad (84)$$

$$E\left(\frac{\partial L}{\partial \gamma}\right)\left(\frac{\partial L}{\partial \mathbf{s}_{\Re}(t)}\right)^T = 0 \quad (85)$$

and

$$E\left(\frac{\partial L}{\partial \gamma}\right)\left(\frac{\partial L}{\partial \mathbf{s}_{\Im}(t)}\right)^T = 0. \quad (86)$$

The Fisher information matrix $\mathbf{J}(\boldsymbol{\eta})$ can now be written as

$$\mathbf{J}(\boldsymbol{\eta}) = \begin{bmatrix} \Gamma & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \boldsymbol{\Sigma}_{\Re} & -\boldsymbol{\Sigma}_{\Im} & \cdots & 0 & 0 & \boldsymbol{\Delta}_{\Re}(1) \\ 0 & \boldsymbol{\Sigma}_{\Im} & \boldsymbol{\Sigma}_{\Re} & \cdots & 0 & 0 & \boldsymbol{\Delta}_{\Im}(1) \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \boldsymbol{\Sigma}_{\Re} & -\boldsymbol{\Sigma}_{\Im} & \boldsymbol{\Delta}_{\Re}(M) \\ 0 & 0 & 0 & \cdots & \boldsymbol{\Sigma}_{\Im} & \boldsymbol{\Sigma}_{\Re} & \boldsymbol{\Delta}_{\Im}(M) \\ 0 & \boldsymbol{\Delta}_{\Re}^T(1) & \boldsymbol{\Delta}_{\Im}^T(1) & \cdots & \boldsymbol{\Delta}_{\Re}^T(M) & \boldsymbol{\Delta}_{\Im}^T(M) & \boldsymbol{\Theta} \end{bmatrix} \quad (87)$$

$$E\left(\frac{\partial L}{\partial\theta_l}\frac{\partial L}{\partial s_{k,\Re}(t)}\right) = 9E\left\{\sum_{t'=1}^{M}\sum_{i=1}^{r}\sum_{j=1}^{r}\frac{\Re\{s_l^*(t')d_i^*(\theta_l)n_i(t')\}}{\gamma^2+|n_i(t')|^2}\frac{\Re\{a_j^*(\theta_k)n_j(t)\}}{\gamma^2+|n_j(t)|^2}\right\}$$

$$= 9E\left\{\sum_{i=1}^{r}\frac{\Re\{s_l^*(t)d_i^*(\theta_l)n_i(t)\}}{\gamma^2+|n_i(t)|^2}\frac{\Re\{a_i^*(\theta_k)n_i(t)\}}{\gamma^2+|n_i(t)|^2}\right\} = 9\sum_{i=1}^{r}|s_l(t)||d_i(\theta_l)|E\left\{\frac{|n_i(t)|^2}{(\gamma^2+|n_i(t)|^2)^2}\right\}$$

$$E\left\{\cos(-\beta_l(t)+\omega_i(\theta_l)+\frac{\pi}{2}+\phi_i(t))\cos(\omega_i(\theta_k)+\phi_i(t))\right\} = 9\sum_{i=1}^{r}|s_l(t)||d_i(\theta_l)|\frac{2}{15\gamma^2}$$

$$\frac{1}{2}E\left\{\cos(-\beta_l(t)+\omega_i(\theta_l)+\frac{\pi}{2}+\omega_i(\theta_k)+2\phi_i(t))+\cos(-\beta_l(t)+\omega_i(\theta_l)+\frac{\pi}{2}-\omega_i(\theta_k))\right\}$$

$$= \frac{3}{5\gamma^2}\sum_{i=1}^{r}\Re\{s_l^*(t)d_i^*(\theta_l)a_i(\theta_k)\}, \tag{81}$$

where

$$\Gamma = \frac{4}{5}\frac{Mr}{\gamma^2}, \tag{88}$$

$$\boldsymbol{\Sigma}_{\Re} = \Re\{\boldsymbol{\Sigma}\} = \frac{3}{5\gamma^2}\Re\{\mathbf{A}^H\mathbf{A}\}, \tag{89}$$

$$\boldsymbol{\Sigma}_{\Im} = \Im\{\boldsymbol{\Sigma}\} = \frac{3}{5\gamma^2}\Im\{\mathbf{A}^H\mathbf{A}\}, \tag{90}$$

$$\boldsymbol{\Delta}_{\Re}(t) = \Re\{\boldsymbol{\Delta}\} = \frac{3}{5\gamma^2}\Re\{\mathbf{A}^H\mathbf{D}\mathbf{S}(t)\}, \tag{91}$$

$$\boldsymbol{\Delta}_{\Im}(t) = \Im\{\boldsymbol{\Delta}\} = \frac{3}{5\gamma^2}\Im\{\mathbf{A}^H\mathbf{D}\mathbf{S}(t)\} \tag{92}$$

and

$$\boldsymbol{\Theta} = \frac{3}{5\gamma^2}\sum_{t=1}^{M}\Re\{\mathbf{S}^H(t)\mathbf{D}^H\mathbf{D}\mathbf{S}(t)\}. \tag{93}$$

Clearly

$$\mathrm{CRB}(\gamma) = \frac{1}{\Gamma} = \frac{5}{4}\frac{\gamma^2}{Mr}. \tag{94}$$

Setting $\boldsymbol{\Xi} \stackrel{\text{def}}{=} \boldsymbol{\Sigma}^{-1}$ and using a standard result on the inverse of a partitioned matrix [2], the CRB for the DOA $\boldsymbol{\theta}$ can be expressed as

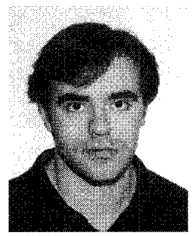
$$\mathrm{CRB}^{-1}(\boldsymbol{\theta}) =$$
$$\boldsymbol{\Theta} - \left[\boldsymbol{\Delta}_{\Re}^T(1), \boldsymbol{\Delta}_{\Im}^T(1), \ldots, \boldsymbol{\Delta}_{\Re}^T(M), \boldsymbol{\Delta}_{\Im}^T(M)\right]$$
$$\begin{bmatrix}\boldsymbol{\Xi}_{\Re} & -\boldsymbol{\Xi}_{\Im} & \cdots & 0 & 0\\ \boldsymbol{\Xi}_{\Im} & \boldsymbol{\Xi}_{\Re} & \cdots & 0 & 0\\ \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & \cdots & \boldsymbol{\Xi}_{\Re} & -\boldsymbol{\Xi}_{\Im}\\ 0 & 0 & \cdots & \boldsymbol{\Xi}_{\Im} & \boldsymbol{\Xi}_{\Re}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Delta}_{\Re}(1)\\ \boldsymbol{\Delta}_{\Im}(1)\\ \vdots\\ \boldsymbol{\Delta}_{\Re}(M)\\ \boldsymbol{\Delta}_{\Im}(M)\end{bmatrix}$$
$$= \boldsymbol{\Theta} - \sum_{t=1}^{M}\Re\{\boldsymbol{\Delta}^H(t)\boldsymbol{\Xi}\boldsymbol{\Delta}(t)\}$$
$$= \frac{3}{5\gamma^2}\sum_{t=1}^{M}\Re\{\mathbf{S}^H(t)\mathbf{D}^H\mathbf{D}\mathbf{S}(t)$$
$$-\mathbf{S}^H(t)\mathbf{D}^H\mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H\mathbf{D}\mathbf{S}(t)\Big\}$$
$$= \frac{3}{5\gamma^2}\sum_{t=1}^{M}\Re\left\{\mathbf{S}^H(t)\mathbf{D}^H\left[I-\mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H\right]\mathbf{D}\mathbf{S}(t)\right\} \tag{95}$$

and the proof is completed. ■

## REFERENCES


[1] R. O. Schweppe, "Sensor array data processing for multiple signal sources," *IEEE Trans. Inform. Theory*, vol. 14, pp. 294–305, 1968.
[2] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramér-Rao bound," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 720–741, 1989.
[3] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1553–1560, 1988.
[4] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, pp. 1408–1418, 1969.
[5] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagat.*, vol. AP-34, pp. 276–280, 1986.
[6] S. S. Reddi, "Multiple source location—A digital approach," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-15, pp. 95–105, 1979.
[7] R. Kumaresan and D. W. Tufts, "Estimating the angles of arrival of multiple plane waves," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-19, pp. 134–139, 1983.
[8] 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.
[9] H. Wang and M. Kaveh, "Estimation of angles of arrival for wide-band sources," in *Proc. ICASSP*, San Diego, CA, 1984, pp. 7.5.1–7.5.4.
[10] ——, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 823–831, 1985.
[11] H. Hung and M. Kaveh, "Focusing matrices for coherent signal-subspace processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1272–1281, 1988.
[12] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag.*, pp. 4–24, Apr. 1988.
[13] B. D. Van Veen, "An analysis of several partially adaptive beamformer designs," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 192–203, 1989.
[14] P. Forster and G. Vezzosi, "Application of spheroidal sequences to array processing," in *Proc. ICASSP* Dallas, TX, 1987, pp. 2268–2271.
[15] P. Forster and C. L. Nikias, "Bearing estimation in the bispectrum domain," *IEEE Trans. Signal Processing*, vol. 39, pp. 1994–2006, 1991.

[16] D. Kletter and H. Messer, "The role of third-order spectrum in maximum likelihood time delay estimation of a random multitone signal in noise," in *Proc. ICASSP*, Glasgow, Scotland, May 1989, pp. 2310–2313.
[17] M. Shao and C. L. Nikias, "On symmetric stable models for impulsive noise," Tech. Rep. USC-SIPI-231, Univ. of Southern Calif., Feb. 1993.
[18] G. A. Tsihrintzis and C. L. Nikias, "Performance of optimum and suboptimum receivers in the presence of impulsive noise modeled as an alpha-stable process," *IEEE Trans. Commun.*, vol. 43, pp. 904–914, 1995.
[19] ——, "On the detection of impulsive stochastic transients," *Signal Processing*, vol. 41, pp. 1–16, 1994.
[20] P. Tsakalides and C. L. Nikias, "A new robust array processing method based on the Cauchy beamformer," in *1994 Adaptive Sensor Array Processing (ASAP) Workshop*, Lexington, MA), MIT Lincoln Lab., Mar. 1994.
[21] M. Shao and C. L. Nikias, "Signal processing with fractional lower order moments: Stable processes and their applications," *Proc. IEEE*, vol. 81, pp. 986–1010, 1993.
[22] P. Tsakalides and C. L. Nikias, "The robust covariation-based MUSIC (ROC-MUSIC) algorithm for bearing estimation in impulsive noise environments," Tech. Rep. USC-SIPI-278, Univ. of Southern Calif., Jan. 1995.
[23] V. M. Zolotarev, "Integral transformations of distributions and estimates of parameters of multidimensional spherically symmetric stable laws," in *Contributions to Probability*, J. Gani and V. K. Rohatgi, Eds. New York: Academic, 1981, pp. 283–305.
[24] G. Gong and F. J. Samaniego, "Pseudo maximum likelihood estimation: Theory and applications," *Ann. Statist.*, vol. 9, pp. 861–869, 1981.
[25] R. Brant, "Approximate likelihood and probability calculations based on transforms," *Ann. Statist.*, vol. 12, pp. 989–1005, 1984.
[26] D. G. Luenberger, *Linear and Nonlinear Programming*. Menlo Park, CA: Addison-Wesley, 1984, 2nd ed.
[27] A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky, "An overview of matched field methods in ocean acoustics," *IEEE J. Ocean. Eng.*, vol. 18, pp. 401–424, 1993.
[28] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1987, 2nd ed.
[29] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*. New York: Academic, 1980.



**Panagiotis Tsakalides** was born in Serres, Greece, on December 15, 1967. He received the diploma in electrical engineering from the Aristotle University of Thessaloniki, Greece, in 1990. The same year he came to Los Angeles for graduate studies with scholarships from the Fulbright and Bodossakis Foundations. He received the M.S. degree in computer engineering from the University of Southern California (USC) in 1991, and he is currently working toward the doctoral degree in electrical engineering at the same university.

Since 1991, he has been a research assistant at the Signal and Image Processing Institute at USC. His general research interests lie in the area of statistical signal processing with emphasis in estimation and detection theory, and applications in array processing, sonar, and radar problems.

**Chrysostomos L. Nikias**, for a photograph and biography, see p. 2664 of this issue of this TRANSACTIONS.