MODE WITH EXTRA-ROOTS (MODEX): A NEW DOA ESTIMATION ALGORITHM WITH AN IMPROVED THRESHOLD PERFORMANCE

A.B. Gershman

Signal Theory Group
Ruhr University, Bochum, Germany
gsh@sth.ruhr-uni-bochum.de

P. Stoica

Systems and Control Group
Uppsala University, Uppsala, Sweden
ps@syscon.uu.se

ABSTRACT

We propose a new MODE-based direction of arrival (DOA) estimation algorithm with an improved SNR threshold as compared to the conventional MODE technique. Our algorithm preserves all good properties of MODE, such as asymptotic efficiency, excellent performance in scenarios with coherent sources, as well as a reasonable computational cost. Similarly to root-MODE, the proposed method does not require any global multidimensional optimization since it is based on a combination of polynomial rooting and a simple combinatorial search. Our technique is referred to as MODEX (MODE with EXtra roots) because it makes use of a certain polynomial with a larger degree than that of the conventional MODE-polynomial. The source DOA's are estimated via checking a certain (enlarged) number of candidate DOA's using either the stochastic or the deterministic Maximum Likelihood (ML) function. To reduce the computational cost of MODEX, *a priori* information about source localization sectors can be exploited.

1. INTRODUCTION

MODE is a recently developed DOA estimation technique which is known to be statistically efficient when either the SNR or the number of snapshots N is sufficiently large [5]. In the case of a Uniform Linear Array (ULA), MODE requires a very simple implementation based on the eigendecomposition of the array covariance matrix and polynomial rooting [3], [5], [6]. In contrast to the eigenstructure methods [3], asymptotic efficiency is achieved by MODE for both uncorrelated and coherent source scenarios [5].

These excellent properties make root-MODE a strong candidate for the best possible ULA processing algorithm with nearly optimal asymptotic performance/complexity tradeoff [3], [6]. However, MODE is known to be only a *large sample approximation* of the ML method. Therefore, the statistical efficiency property is achieved by MODE only *asymptotically* (for large N and SNR). In the low SNR and short sample cases, the performance of MODE may degrade severely [6]. This type of performance degradation is usually referred to as the *threshold breakdown effect* [1], [4]. Obviously, the threshold degradation may reduce the attractivity of MODE as the "best array processing method". Therefore, a search for new modifications of MODE improving its threshold estimation performance (while preserving its optimal asymptotic performance and moderate computational complexity) appears to be a very important task. Recently, several attempts have been made to improve the threshold performance of MODE using the Forward-Backward (FB) approach (see [6] and references therein). However, the results of this study were quite unexpected. It was shown that FB-MODE is outperformed by conventional MODE in both the threshold and asymptotic domains.

Supported in part by DFG under Grant Bo 568/22-1, INTAS under Grant INTAS-93-642-Ext, and the Senior Individual Grant Program of the Swedish Foundation for Strategic Research.

In this paper, we present a novel MODE-based method referred to as MODEX (MODE with EXtra-roots). The key idea of our technique is to remove the outliers caused by the large difference between the exact and sample signal-subspaces. This can be done increasing the degree of the MODE-polynomial to obtain the set of signal DOA estimates from a larger set of candidate DOA's based on either the deterministic or the stochastic ML criterion. It is worth noting that MODEX has the same asymptotic performance as conventional MODE. At the same time, our algorithm is demonstrated to provide dramatically better threshold performance than MODE. We stress that MODEX is computationally somewhat more expensive than MODE, mainly because it requires an additional combinatorial search. However, it should be noted that the computational cost of our technique may be reduced considerably if preliminary information about the source localization sectors is available [1]. The additional computational cost of MODEX can be viewed as a natural price for the significant threshold improvements achieved relative to MODE. Furthermore, it should be noted that MODEX can be implemented in a much simpler way than the stochastic and deterministic ML algorithms because it avoids a nonlinear optimization over a multidimensional parameter space.

2. THE CONVENTIONAL MODE ALGORITHM

Assume that a ULA of n sensors receives narrowband plane waves from q far-field sources. The $n \times 1$ vector of sensor outputs can be modeled as [3]

$$\boldsymbol{x}(i) = \boldsymbol{A}\boldsymbol{s}(i) + \boldsymbol{n}(i) \tag{1}$$

where $\boldsymbol{A} = [\boldsymbol{a}(\theta_1), \ldots, \boldsymbol{a}(\theta_q)]$ is the $n \times q$ matrix of the source steering vectors, $\boldsymbol{a}(\theta)$ is the $n \times 1$ steering vector toward the direction θ, $\boldsymbol{s}(i)$ is the $q \times 1$ vector of random source waveforms, $\boldsymbol{n}(i)$ is the $n \times 1$ vector of sensor noise, and $q < n$ is the number of sources. The parameter q is assumed to be known [1]-[6].

The $n \times n$ spatial covariance matrix of array outputs is given by [3]

$$\boldsymbol{R} = \mathrm{E}[\boldsymbol{x}(i)\boldsymbol{x}^H(i)] = \boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^H + \sigma^2 \boldsymbol{I} \tag{2}$$

where $\boldsymbol{S} = \mathrm{E}[\boldsymbol{s}(i)\boldsymbol{s}^H(i)]$ is the $q \times q$ covariance matrix of signal waveforms, $\boldsymbol{I}$ is the identity matrix, $\mathrm{E}[\cdot]$ and $(\cdot)^H$ denote the expectation operator and the Hermitian transpose, respectively, and σ^2 is the noise power. The sample covariance matrix is obtained as

[3]

$$\hat{\boldsymbol{R}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}(i)\boldsymbol{x}^H(i) \tag{3}$$

The eigendecomposition of (3) is given by

$$\hat{\boldsymbol{R}} = \hat{\boldsymbol{E}}_S\hat{\boldsymbol{\Lambda}}_S\hat{\boldsymbol{E}}_S^H + \hat{\boldsymbol{E}}_N\hat{\boldsymbol{\Lambda}}_N\hat{\boldsymbol{E}}_N^H \tag{4}$$

where the $q \times q$ and $(n-q)\times(n-q)$ diagonal matrices $\hat{\boldsymbol{\Lambda}}_S$ and $\hat{\boldsymbol{\Lambda}}_N$ contain the q and $n-q$ signal and noise-subspace eigenvalues, whereas the columns of the $n \times q$ and $n \times (n-q)$ matrices

$$\hat{\boldsymbol{E}}_S = [\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_q] \qquad \hat{\boldsymbol{E}}_N = [\hat{e}_{q+1}, \hat{e}_{q+2}, \ldots, \hat{e}_n] \tag{5}$$

contain the signal and noise-subspace eigenvectors, respectively.

The conventional MODE algorithm estimates the source DOA's via the minimization of the following function [5]:

$$f_{\text{MODE}}(\boldsymbol{b}) = \text{tr}\{\boldsymbol{\Pi}(\boldsymbol{b})\hat{\boldsymbol{E}}_S\boldsymbol{W}\hat{\boldsymbol{E}}_S^H\} \tag{6}$$

where

$$\boldsymbol{\Pi}(\boldsymbol{b}) = \boldsymbol{B}(\boldsymbol{B}^H\boldsymbol{B})^{-1}\boldsymbol{B}^H \tag{7}$$

$$\boldsymbol{W} = (\hat{\boldsymbol{\Lambda}}_S - \hat{\sigma}^2\boldsymbol{I})^2\hat{\boldsymbol{\Lambda}}_S^{-1} \tag{8}$$

$$\hat{\sigma}^2 = \frac{1}{n-q}\,\text{tr}\{\hat{\boldsymbol{\Lambda}}_N\} \tag{9}$$

and $\boldsymbol{B}$ is a standard $n \times (n-q)$ Toeplitz matrix [5]

$$\boldsymbol{B}^H = \begin{bmatrix} b_0 & \ldots & b_q & & 0 \\ & \ddots & & \ddots & \\ 0 & & b_0 & \ldots & b_q \end{bmatrix} \tag{10}$$

defined by the $(q+1)\times 1$ vector of the complex-valued polynomial coefficients

$$\boldsymbol{b} = (b_0, b_1, \ldots, b_q)^T \tag{11}$$

It can be shown [5] that an asymptotically efficient estimator of the source DOA's can be formulated using the following two-step procedure:

- **Step 1**. Obtain an initial estimate $\hat{\boldsymbol{b}}$ of (11) by minimizing the following quadratic function:

$$f_1(\boldsymbol{b}) = \text{tr}\{\boldsymbol{B}^H\hat{\boldsymbol{E}}_S\boldsymbol{W}\hat{\boldsymbol{E}}_S^H\boldsymbol{B}\} \tag{12}$$

- **Step 2**. Obtain the refined estimate $\hat{\hat{\boldsymbol{b}}}$ by minimizing the quadratic function

$$f_2(\boldsymbol{b}) = \text{tr}\{(\hat{\boldsymbol{B}}^H\hat{\boldsymbol{B}})^{-1}\boldsymbol{B}^H\hat{\boldsymbol{E}}_S\boldsymbol{W}\hat{\boldsymbol{E}}_S^H\boldsymbol{B}\} \tag{13}$$

  where the matrix $\hat{\boldsymbol{B}}$ is made from the estimate $\hat{\boldsymbol{b}}$ obtained in step 1. Finally, find the estimates of the source DOA's by rooting the polynomial with coefficients $\hat{\hat{\boldsymbol{b}}}$.

To exclude the trivial solution, the minimization in steps 1 and 2 should be performed under the norm constraint. Additionally, the conjugate-symmetry constraint [5] should be used. Note that step 2 can be iterated a few times, though even without repeating this step the asymptotic efficiency is guaranteed [5].

3. THE MODEX ALGORITHM

In the situation of low SNR or short sample size, MODE may suffer from the threshold breakdown effect. Qualitatively speaking, this effect may be caused by the fact that the sample signal-subspace eigenvectors have a significant component in the null-space of the matrix $\boldsymbol{A}$. In this case, MODE incorrectly tends to spend one or more roots to model this component as a signal source. Of course, this leads to a strong degradation of the performance of MODE, because in this case there are not enough roots left to model the signal sources (i.e., q sources will be modeled using less than q roots). The key idea of MODEX is to use p ($p < n-q$) extra-roots to avoid this undesirable phenomenon. With such extra-roots, there will be no problem with wrong signal interpretation because the additional roots will provide an adequate modeling of both the signal and noise components. In order to obtain p extra-roots, let us use "zero-padding" on the weighting matrix (8), i.e., let us consider a new weighting matrix with extended dimension $(q+p)\times(q+p)$

$$\overline{\boldsymbol{W}} = \text{diag}\{\boldsymbol{W}, 0, \ldots, 0\} = \begin{bmatrix} \boldsymbol{W} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \tag{14}$$

With (14), the MODE function (6) can be extended as

$$f_{\text{MODE}}(\overline{\boldsymbol{b}}) = \text{tr}\{\boldsymbol{\Pi}(\overline{\boldsymbol{b}})\hat{\boldsymbol{E}}\overline{\boldsymbol{W}}\hat{\boldsymbol{E}}^H\} \tag{15}$$

where the $(q+p+1)\times 1$ vector

$$\overline{\boldsymbol{b}} = (\bar{b}_0, \bar{b}_1, \ldots, \bar{b}_{q+p})^T \tag{16}$$

contains the coefficients of the polynomial of enlarged degree $q+p$, the $n \times (n-q-p)$ matrix $\overline{\boldsymbol{B}}$ is defined as

$$\overline{\boldsymbol{B}}^H = \begin{bmatrix} \bar{b}_0 & \ldots & \bar{b}_{q+p} & & 0 \\ & \ddots & & \ddots & \\ 0 & & \bar{b}_0 & \ldots & \bar{b}_{q+p} \end{bmatrix} \tag{17}$$

and the $n \times (q+p)$ matrix $\hat{\boldsymbol{E}}$ is given by

$$\hat{\boldsymbol{E}} = [\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_{q+p}] \tag{18}$$

From (14) and (18), we obtain that

$$\hat{\boldsymbol{E}}\overline{\boldsymbol{W}}\hat{\boldsymbol{E}}^H = \hat{\boldsymbol{E}}_S\boldsymbol{W}\hat{\boldsymbol{E}}_S^H \tag{19}$$

Hence, the extended MODE function (15) is given by

$$f_{\text{MODE}}(\overline{\boldsymbol{b}}) = \text{tr}\{\boldsymbol{\Pi}(\overline{\boldsymbol{b}})\hat{\boldsymbol{E}}_S\boldsymbol{W}\hat{\boldsymbol{E}}_S^H\} \tag{20}$$

From (20), we obtain a very simple interpretation of the extended MODE function (15): the only difference between the functions (6) and (20) is that the "enlarged" vector (16) is used in (20) instead of the conventional vector (11) used in (6). Though the the framework of (15) may be also useful for preventing subspace swaps [2], we see that the minimization of the function (15) with the "zero-padded" weighting matrix (14) involves the signal-subspace only.

The minimization of the function (20) can be performed using the two-step procedure outlined in Section 2. However, after the minimization of (20), the "signal" roots of $\overline{\boldsymbol{b}}$ cannot be sorted out based on their proximity to the unit circle. Moreover, though the extended MODE criterion (20) with a good sorting algorithm is expected to remove outliers and improve the threshold performance, the use of (20) alone cannot guarantee the same asymptotic performance as achieved by the conventional MODE algorithm.

To preserve the asymptotic performance of MODE and provide an adequate signal root sorting procedure, we use two different tricks. First, we propose to exploit all $2q+p$ roots of the polynomials with the coefficients (11) and (16) which are obtained after the minimization of the functions (6) and (20). Moreover, we will use alternatively either the deterministic or the stochastic negative ML functions

$$f_{\mathrm{DET-ML}}(\boldsymbol{\theta}) = \mathrm{tr}\,\{\boldsymbol{P}_{\boldsymbol{A}}^{\perp}\hat{\boldsymbol{R}}\} \tag{21}$$

$$f_{\mathrm{STO-ML}}(\boldsymbol{\theta}) = \log\det\,\{\boldsymbol{P}_{\boldsymbol{A}}\hat{\boldsymbol{R}}\boldsymbol{P}_{\boldsymbol{A}} + \frac{1}{n-q}\,\mathrm{tr}\,\{\boldsymbol{P}_{\boldsymbol{A}}^{\perp}\hat{\boldsymbol{R}}\}\,\boldsymbol{P}_{\boldsymbol{A}}^{\perp}\} \tag{22}$$

to check all possible combinations of q DOA's taken from the $2q+p$ candidate DOA's (candidate roots), where $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_q)^T$ is the $q \times 1$ vector of unknown signal DOA's. The purpose of this search over the candidate DOA's is to pick up the combination for which the negative ML function has a minimum. In (21) and (22),

$$\boldsymbol{P}_{\boldsymbol{A}} = \boldsymbol{A}\boldsymbol{A}^{\dagger} \quad \boldsymbol{P}_{\boldsymbol{A}}^{\perp} = \boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{A}} \quad \boldsymbol{A}^{\dagger} = (\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H \tag{23}$$

Let us define the $(2q+p) \times 1$ vector of candidate DOA's as

$$\boldsymbol{\theta}_{\mathrm{CAND}} = \left(\hat{\boldsymbol{\theta}}^T, \hat{\hat{\boldsymbol{\theta}}}^T\right)^T = \left(\hat{\theta}_1, \ldots, \hat{\theta}_q, \hat{\hat{\theta}}_1, \ldots, \hat{\hat{\theta}}_{q+p}\right)^T \tag{24}$$

where $\hat{\theta}_i$ $(i = 1, 2, \ldots, q)$ and $\hat{\hat{\theta}}_i$ $(i = 1, 2, \ldots, q+p)$ are the estimates of the ith DOA obtained by the minimization of the conventional and extended MODE functions (6) and (20), respectively.

The ML-based sorting procedure can be formulated as

$$\hat{\boldsymbol{\theta}}_{\mathrm{MODEX}} = \arg\min_{\boldsymbol{\theta}\subset\boldsymbol{\theta}_{\mathrm{CAND}}} \{f_{\mathrm{DET-ML}}(\boldsymbol{\theta})\} \tag{25}$$

$$\hat{\boldsymbol{\theta}}_{\mathrm{MODEX}} = \arg\min_{\boldsymbol{\theta}\subset\boldsymbol{\theta}_{\mathrm{CAND}}} \{f_{\mathrm{STO-ML}}(\boldsymbol{\theta})\} \tag{26}$$

alternatively, where $\boldsymbol{\theta}$ is a $q \times 1$ vector whose elements form a subset[1] of the vector $\boldsymbol{\theta}_{\mathrm{CAND}}$.

We are now ready to formulate our MODEX algorithm:

- **Step 1**. Obtain the vector of candidate DOA's (24) via independent minimization of the functions (6) and (20) and combining the so-obtained DOA's. For each minimization, exploit the two-step procedure outlined in Section 2.
- **Step 2**. Check all possible combinations of q DOA's taken from the candidate DOA's using either the deterministic ML function (21), or, alternatively, the stochastic ML function (22). Finally, obtain the estimates of the source DOA's as the combination which minimizes the ML function (according to either (25) or (26)).

This algorithm avoids multidimensional optimization since it is based on a combination of polynomial rooting and combinatorial search. The number of combinations to be checked depends on the number of sources and the parameter p (which is chosen by the user), and is given by

$$M = \binom{2q+p}{q} = \frac{(2q+p)!}{q!\,(q+p)!} \tag{27}$$

It is worth noting that M is independent of the number of sensors and is reasonably small for a moderate number of sources. For example, for $q = 4$ and $p = 4$, the combinatorial search in our algorithm requires testing $M = 495$ points, which is comparable with the complexity of the spectral search in the standard MUSIC algorithm [3].

However, the number of points (27) may be unacceptably high if $q \gg 1$. In such a situation, a sector information can be used for lowering this number and decreasing the computational cost of MODEX. Assume that the source localization sectors (clusters) are pre-estimated as [1]

$$\hat{\Theta}_S = [\theta_{l,1}, \theta_{r,1}] \cup [\theta_{l,2}, \theta_{r,2}] \cup \cdots \cup [\theta_{l,m}, \theta_{r,m}] \tag{28}$$

where $\theta_{l,i}$ and $\theta_{r,i}$ are the left and the right bounds of each sector, respectively, and m is the number of sectors. Note that such preliminary sector information is used in a variety of popular array processing methods and can be easily obtained via conventional

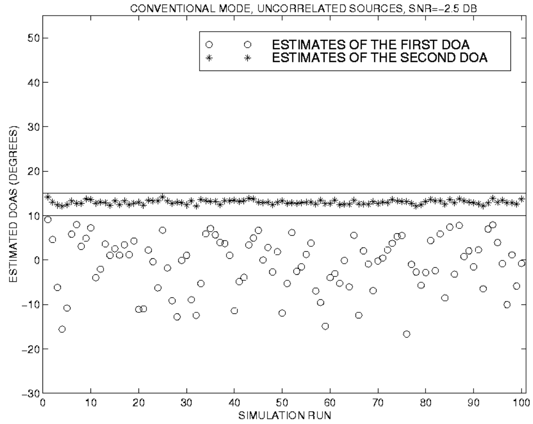

Figure 1: MODE estimates in the first example.

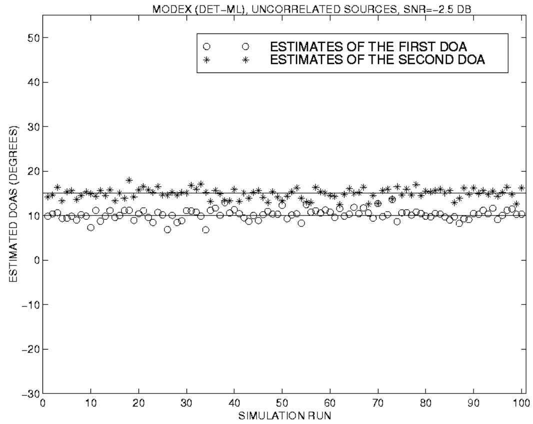

Figure 2: MODEX estimates in the first example.

[1] For the sake of simplicity, we use similar notation for sets and vectors meaning that any $d \times 1$ vector corresponds to a set with d elements.

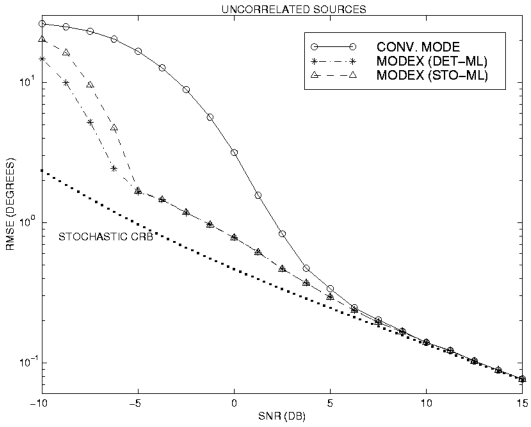

Figure 3: RMSE's vs. the SNR for the second example.

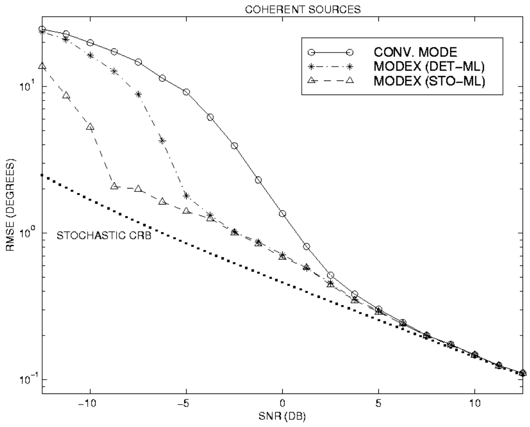

Figure 4: RMSE's vs. the SNR for the third example.

beamforming [1]. The central idea of exploiting the sector information (28) is to reduce the number of candidate DOA's prior to the MODEX algorithm by removing from the candidate DOA set (24) the DOA's which lie outside the intervals (28). Introduce a new vector $\tilde{\boldsymbol{\theta}}_{\mathbf{CAND}} = \mathcal{G}_{\hat{\Theta}_S} \boldsymbol{\theta}_{\mathbf{CAND}}$, where $\mathcal{G}_{\hat{\Theta}_S}$ is the operator which selects only the elements of the vector $\boldsymbol{\theta}_{\mathbf{CAND}}$ that belong to the sectors (28). This operator is expected to compress the original vector $\boldsymbol{\theta}_{\mathbf{CAND}}$ of dimension $\dim\{\boldsymbol{\theta}_{\mathbf{CAND}}\} = 2q + p$ to the vector $\tilde{\boldsymbol{\theta}}_{\mathbf{CAND}}$ of smaller dimension, say $\dim\{\tilde{\boldsymbol{\theta}}_{\mathbf{CAND}}\} = d$, $d < 2q + p$. The compression ratio $(2q + p)/d$ depends on the accuracy of the sector information (28) and on the chosen value of p.

4. SIMULATION RESULTS

In our simulations, we assumed a ULA of ten omnidirectional sensors with half-wavelength spacing, $N = 100$ snapshots, and two equally powered narrowband sources with DOA's $\theta_1 = 10°$ and $\theta_2 = 15°$ relative to broadside. In all examples, the parameter $p = 4$ was assumed and no sector information was exploited in MODEX.

In the first example, we assumed uncorrelated sources with SNR $= -2.5$ dB. In Figs. 1 and 2, 100 estimation trials are shown for conventional MODE and MODEX, respectively. In the last algorithm, the deterministic ML function was used [eqn. (25)]. This figure clearly demonstrates an improved performance of our technique relative to MODE: MODE is not able to resolve the closely spaced sources, whereas MODEX resolves them in almost all trials.

In the second example, we assumed uncorrelated sources once again. Fig. 3 displays the DOA estimation Root-Mean-Square Errors (RMSE's) versus the SNR. A total of 1000 independent simulation runs were performed to obtain each simulated point. The stochastic Cramér–Rao Bound (CRB) is also shown.

In the last example, we used the same scenario as in the second one but assumed coherent sources. Fig. 4 displays the DOA estimation RMSE's versus the SNR.

From Figs. 3 and 4, we see that MODEX has much better threshold performance than conventional MODE. It is also worth noting from these figures that the asymptotic performance of MODEX is similar to that of MODE (i.e., MODEX preserves the asymptotic efficiency of MODE). In the case of uncorrelated sources, the performance of MODEX does not depend significantly on what kind of the ML function is employed. However, in the case of coherent sources the stochastic ML function seems to be the best choice for MODEX.

Interestingly, MODEX can be interpreted as performing a local search of the ML criterion. We stress that (21) and (22) tend to have many false minima located in the vicinity of the global minimum, so that any standard local search algorithm is bound to fail, even if initialized close to the global minimum. It is therefore worth noting that among the fairly limited number of points in the parameter space tested by MODEX there is almost always one that falls near the global minimum of the negative ML criterion.

5. CONCLUSIONS

We proposed the MODEX algorithm, a new MODE-based technique having drastically improved SNR threshold relative to conventional MODE. Our technique preserves all good properties of MODE, such as asymptotic efficiency, reasonable computational cost, and excellent performance in scenarios with highly correlated sources.

6. REFERENCES

[1] A.B. Gershman, "Pseudo-randomly generated estimator banks: a new tool for improving the threshold performance of direction finding," *IEEE Trans. SP*, pp. 1351-1364, May 1998.

[2] A.B. Gershman and P. Stoica, "New MODE-based techniques for direction finding with an improved threshold performance," submitted to *Signal Processing*.

[3] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE SP Mag.*, pp. 67-94, July 1996.

[4] H. Lee and F. Li, "Quantification of the difference between detection and resolution thresholds for multiple closely spaced emitters," *IEEE Trans. SP*, pp. 2274-2277, June 1993.

[5] P. Stoica and K.C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. ASSP*, pp. 1132-1143, July 1990.

[6] P. Stoica and M. Jansson, "On forward-backward MODE for array signal processing," *DSP – A Review J.*, pp. 239-252, Oct. 1997.