algorithms are obtained from 50 independent Monte-Carlo trials and are compared with the results of the 2D-MODE algorithm and with the corresponding theoretical asymptotic statistical performances.

We consider the case of large $N$, which occurs in two-dimensional angle estimation by means of an $M \times N$ rectangular uniform linear array when the incident angles are related to each other [1]. The incident signals are assumed to be narrowband plane waves. In this application, $N$ denotes the number of snapshots taken at the output of the array. Consider the case shown in Figure 1 where a vertical 2-D rectangular uniform linear array with $M = 10$ and $N = 500$ is used to estimate the 2-D incident angles of a signal arriving from a low angle relative to a smooth reflecting surface such as the calm sea. Assume that the original signal arrives from $(\phi_1, \theta_1)$, where $\phi_1$ and $\theta_1$ denote the azimuth and elevation angles of the signal, as shown in Figure 1. Then its reflected signal arrives from $(\phi_2, \theta_2) = (\phi_1, 180^\circ - \theta_1)$. The 2-D incident angles can be calculated from the $\omega_k$ and $\mu_\tau$ in Equation (1) by:

$$\omega_k = \frac{2\pi k}{\lambda_0} \sin \theta_k \sin \phi_k, \quad k = 1, \ldots, N$$

and

$$\mu_\tau = \frac{2\pi \lambda_0}{\tau} \cos \theta_\tau, \quad \tau = 1, 2, \ldots$$

where $\lambda_0$ denotes the wavelength of the incident signals.

Figure 2 shows the root-mean-squared errors (RMSEs) of the angle estimates and the corresponding asymptotic statistical performances of the 1D-MODE and 2D-MODE algorithms as a function of SNR when the direct and reflected signals arrive from $(45^\circ, 85^\circ)$ and $(43^\circ, 95^\circ)$, respectively. The correlation coefficient between the direct and reflected signals is 0.99. Further, $M = 10$, $N = 16$, and $N = 500$. (a) The estimates of the azimuth angle $\phi$. (b) The estimates of the elevation angle $\theta$.

REFERENCES


Thus the estimates $\hat{\omega}$ and $\hat{\mu}$ obtained with the 1D-MODE asymptotically (for high SNR) achieve the CR-bounds in (25) and (26), respectively, which means that the 1D-MODE is an asymptotically (for SNR $\gg 1$) statistically efficient estimator for estimating the 2-D frequencies.

4.2. The Case of Large $N$

Similar to the case of high SNR, the asymptotic covariance matrices of $\hat{\omega}$ and $\hat{\mu}$ in the case of large $N$ are respectively given by

$$ E \left\{ (\hat{\omega} - \omega)(\hat{\omega} - \omega)^T \right\} = \frac{\sigma^2}{2N} \left[ \text{Re} \{ H \omega \circ V_\omega^T \} \right]^{-1}, $$

(29)

and

$$ E \left\{ (\hat{\mu} - \mu)(\hat{\mu} - \mu)^T \right\} = \frac{\sigma^2}{2N} \left[ \text{Re} \{ H \mu \circ V_\mu^T \} \right]^{-1}, $$

(30)

where

$$ V_\omega = P_\omega A^H R_\omega^{-1} H P_\omega, $$

(31)

and

$$ V_\mu = P_\mu B^H R_\mu^{-1} H P_\mu. $$

(32)

It has been shown in [1] that the large-sample covariance matrices of the estimates of $\omega$ and $\mu$ obtained with 2D-MODE are equal to the corresponding stochastic CRBs given by

$$ \left[ \text{CRB}_{\omega}^{-1} \right]_{ij} = (2N/\sigma^2) \text{Re} \left\{ \text{tr} \left\{ (A^i)^H P A^i \right\} \right\}, $$

(33)

$$ \left[ \text{CRB}_{\mu}^{-1} \right]_{ij} = (2N/\sigma^2) \text{Re} \left\{ \text{tr} \left\{ (B^j)^H P B^j \right\} \right\}, $$

(34)

respectively, where

$$ S = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} E \left\{ s(t_n) s^H(t_n) \right\}, $$

(35)

and

$$ R = E \left\{ \text{vec} \left[ Y^T(t_n) \right] \text{vec}^H \left[ Y^T(t_n) \right] \right\}. $$

(36)

In this case, the 1D-MODE is no longer an asymptotically statistically efficient estimator for estimating the 2-D frequencies. According to the general theory of the CR-bounds, we have

$$ E \left\{ (\hat{\omega} - \omega)(\hat{\omega} - \omega)^T \right\} \geq \text{CRB}_{\omega}, $$

(37)

and

$$ E \left\{ (\hat{\mu} - \mu)(\hat{\mu} - \mu)^T \right\} \geq \text{CRB}_{\mu}. $$

(38)

The numerical examples given in the following section show that the larger the $M$ ($\hat{M}$) or the higher the SNR, the smaller the difference between $E \left\{ (\hat{\omega} - \omega)(\hat{\omega} - \omega)^T \right\}$ and $\text{CRB}_{\omega}$, and $E \left\{ (\hat{\mu} - \mu)(\hat{\mu} - \mu)^T \right\}$ and $\text{CRB}_{\mu}$.

4.3. Further Comments

We remark that when $X$ in (7) is a diagonal matrix, $y_n = \sum_{j=1}^{\kappa} a_j(t_n) e^{(\omega_j + \mu_j) t_n} + \epsilon_n$, can be modeled with the following data model [1]:

$$ y_n = \sum_{j=1}^{\kappa} a_j(t_n) e^{(\omega_j + \mu_j) t_n} + \epsilon_n. $$

(39)

For this case, the 1D-MODE approach can again be used to estimate the 2-D frequencies. Yet it can be shown that serious performance degradation can occur when any of the 1-D approaches (including 1D-MODE) is used with (39), which makes our results even more unexpected. An intuitive explanation is that when $X$ is a diagonal matrix, 1-D processing does not exploit all of the information available and hence lacks the statistical efficiency. When $X$ is a full matrix, which is the case we assume, there is no structural information that is missed by 1-D processing and hence there is no performance degradation under mild conditions.

We also remark that the 1D-MODE approach can be applied to data with non-Gaussian noise without any modification. Its asymptotic covariance matrices will be the same. The CRB matrix for the non-Gaussian case will be different from the one for the Gaussian case, but the CRB matrix computed under the Gaussian assumption remains the lower bound for a large class of estimators whose asymptotic covariance matrices do not depend on the data distribution (for instance all estimators based on second-order statistics).

5. NUMERICAL RESULTS

![Figure 1. Direction-of-arrival estimation with a 2-D array.](image-url)
where $E\{\cdot\}$ denotes the expectation and
\[
P_{\omega} = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} E \left\{ X(t_n) B^T B'X^H(t_n) \right\},
\]  
(10)
with $(\cdot)^*$ denoting the complex conjugate. Note that we use (10) to accommodate both the deterministic and stochastic signal models. The 1D-MODE algorithm [3], or, in a related form, WSF [6], can be applied to $\hat{\mathbf{R}}_{\omega}$ to obtain the estimate of $\hat{\mathbf{K}}_{\omega}$, where
\[
\hat{\mathbf{K}}_{\omega} = \text{min}[\mathbb{M} N, \text{rank}(\mathbf{P}_{\omega})].
\]  
(11)
We assume that $\hat{\mathbf{K}}_{\omega}$ is known. (If $\hat{\mathbf{K}}_{\omega}$ is unknown, it can be estimated from the data as described, for example, in [7].) Further, let $\hat{\mathbf{A}}_{\omega}$ be a diagonal matrix with diagonal elements $\hat{\lambda}_i \geq \hat{\lambda}_i \geq \cdots \geq \hat{\lambda}_{\hat{\mathbf{K}}_{\omega}}$, which are the $\hat{\mathbf{K}}_{\omega}$ largest eigenvalues of $\hat{\mathbf{R}}_{\omega}$, and
\[
\hat{\mathbf{A}}_{\omega} = \hat{\mathbf{A}}_{\omega} - M \sigma^2 \mathbf{I},
\]  
(12)
with $\mathbf{I}$ denoting the identity matrix and
\[
\hat{\sigma}^2 = \frac{1}{(M - \hat{\mathbf{K}}_{\omega})^2} \sum_{i=1}^{M - \hat{\mathbf{K}}_{\omega}} \hat{\lambda}_i
= \frac{1}{(M - \hat{\mathbf{K}}_{\omega})^2} \text{tr}(\hat{\mathbf{R}}_{\omega} - \sum_{i=1}^{\hat{\mathbf{K}}_{\omega}} \hat{\lambda}_i).
\]  
(13)
It is worth noting that the involved computational burden to evaluate $\hat{\mathbf{E}}_{\omega}$, $\hat{\mathbf{A}}_{\omega}$, and $\hat{\mathbf{A}}_{\omega}$ is of the order $O(M^2)$, since usually $\hat{\mathbf{K}}_{\omega} \ll M$, and hence much reduced compared with what would be required for a full eigendecomposition.

The 1D-MODE (or WSF) estimate $\hat{\omega}$ of $\omega$ can be obtained by minimizing the following function:
\[
f(\omega) = \text{tr} \left[ \mathbf{P}_{\mathbf{A}}(\omega) \hat{\mathbf{E}}_{\omega} \left( \hat{\mathbf{A}}_{\omega}^2 \hat{\mathbf{A}}_{\omega}^{-1} \left( \hat{\mathbf{E}}_{\omega} \right)^H \right) \right],
\]  
(14)
where, for some matrix $\mathbf{Z}$, the symbol $\mathbf{P}_{\mathbf{Z}}^+$ stands for the orthogonal projector onto the null space of $\mathbf{Z}^H$. To compute the estimate of $\omega$ without searching over $\mathbf{Z}$, the projector $\mathbf{P}_{\mathbf{A}}$ above must be reparameterized in terms of the coefficients of the so-called "linear predictor" polynomial [1, 3, 5]. We remark that $\hat{\omega}$ is a consistent estimate of $\omega$ for either large $N$ or high SNR [5, 8].

Let
\[
\hat{\mathbf{R}}_{\mu} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{Y}^T(t_n)\mathbf{Y}'(t_n),
\]  
(15)
be the estimate of the following spatial covariance matrix:
\[
\mathbf{R}_{\mu} = E \{ \mathbf{Y}^T(t_n)\mathbf{Y}'(t_n) \} = \mathbf{B} \mathbf{P}_{\mu} \mathbf{B}^H + M \sigma^2 \mathbf{I},
\]  
(16)
where
\[
\mathbf{P}_{\mu} = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} E \{ \mathbf{X}^T(t_n)\mathbf{X}'(t_n) \},
\]  
(17)
Similarly, the 1D-MODE algorithm can be applied to $\hat{\mathbf{R}}_{\mu}$ to obtain the estimate $\hat{\mu}$ of $\mu$.

We remark that the amount of computations required by the 2D-MODE algorithm is $O(M^2\mathbb{M}^2 N)$ and that required by 1D-MODE is $O(M\mathbb{M}(M + \mathbb{M}/N))$. Since the 2D-MODE algorithm requires the computation and eigendecomposition of an $\mathbb{M}^2 \times M\mathbb{M}$ matrix $\hat{\mathbf{R}}$ [1], where
\[
\hat{\mathbf{R}} = \frac{1}{N} \sum_{n=1}^{N} \text{vec} \left[ \mathbf{Y}^T(t_n) \right] \text{vec}^H \left[ \mathbf{Y}^T(t_n) \right],
\]  
(18)
with vec$(\cdot)$ denoting stacking all columns of a matrix into a single column vector, while both $\hat{\mathbf{R}}_{\omega}$ and $\hat{\mathbf{R}}_{\mu}$ can be formed from only the diagonal blocks of $\hat{\mathbf{R}}$. Thus for large $M$ and $\mathbb{M}$, 1D-MODE requires much less computations than 2D-MODE.

4. STATISTICAL PERFORMANCE ANALYSIS

4.1. The Case of High SNR

In the case of high SNR, i.e., $\sigma^2 \ll 1$ (whereas $\mathbf{P}_{\omega}$ and $\mathbf{P}_{\mu}$ are $O(1)$), the asymptotic covariance matrices of the estimate $\hat{\omega}$ of $\omega$ and $\hat{\mu}$ of $\mu$ are respectively given by
\[
E \left\{ (\hat{\omega} - \omega)(\hat{\omega} - \omega)^T \right\} = \frac{\sigma^2}{2N} \text{Re} \left( \mathbf{H}_{\omega} \odot \hat{\mathbf{P}}_{\omega}^T \right)^{-1},
\]  
(19)
\[
E \left\{ (\hat{\mu} - \mu)(\hat{\mu} - \mu)^T \right\} = \frac{\sigma^2}{2N} \text{Re} \left( \mathbf{H}_{\mu} \odot \hat{\mathbf{P}}_{\mu}^T \right)^{-1},
\]  
(20)
where $\odot$ denotes the Hadamard-Schur matrix multiplication.

\[
\hat{\mathbf{P}}_{\omega} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{X}(t_n) \mathbf{B}^T \mathbf{B}' \mathbf{X}'(t_n),
\]  
(21)
\[
\hat{\mathbf{P}}_{\mu} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{X}^T(t_n) \mathbf{A}^T \mathbf{A}' \mathbf{X}'(t_n),
\]  
(22)
\[
\mathbf{H}_{\omega} = \mathbf{D}_{\omega}^T \mathbf{P}_{\mathbf{A}} \mathbf{D}_{\omega},
\]  
(23)
\[
\mathbf{H}_{\mu} = \mathbf{D}_{\mu}^T \mathbf{P}_{\mathbf{A}} \mathbf{D}_{\mu},
\]  
(24)
with the $k$th column of $\mathbf{D}_{\omega}$ being $\partial \omega_k / \partial \omega_k$, and $\mathbf{H}_{\mu}$ defined analogously.

It has also been shown in [1] that the asymptotic (for SNR $\gg 1$) covariance matrices of the estimate $\omega$ and $\mu$ obtained with 2D-MODE are equal to the corresponding deterministic Cramer-Rao bound (CRB) given by
\[
\left( \text{CRB}_{\omega}^{-1} \right)_{ij} = \frac{2N}{\sigma^2} \text{Re} \left\{ \text{tr} \left[ \left( \mathbf{A}_{\omega}^T \mathbf{P}_{\mathbf{A}} \mathbf{A}_{\omega} \right) \odot \left( \mathbf{B}^T \mathbf{B} \right) \right] \hat{S} \right\},
\]  
(25)
\[
\left( \text{CRB}_{\mu}^{-1} \right)_{ij} = \frac{2N}{\sigma^2} \text{Re} \left\{ \text{tr} \left[ \left( \mathbf{A}_{\mu}^T \mathbf{P}_{\mathbf{A}} \mathbf{A}_{\mu} \right) \odot \left( \mathbf{B}_{\mu}^T \mathbf{P}_{\mathbf{A}} \mathbf{B}_{\mu} \right) \right] \hat{S} \right\},
\]  
(26)
where $\odot$ denotes the Kronecker product, $\mathbf{A}_{\omega} = \partial \mathbf{A}(\omega_k) / \partial \omega_k$, $\mathbf{B}_{\mu} = \partial \mathbf{B}(\mu_k) / \partial \mu_k$, and $\hat{S} = \sum_{n=1}^{N} s(t_n) s^H(t_n)$ with $s(t_n) = \text{vec}[\mathbf{X}'(t_n)]$.

From a straightforward computation of Equations (19) and (20), we have
\[
E \left\{ (\hat{\omega} - \omega)(\hat{\omega} - \omega)^T \right\} = \text{CRB}_{\omega},
\]  
(27)
\[
E \left\{ (\hat{\mu} - \mu)(\hat{\mu} - \mu)^T \right\} = \text{CRB}_{\mu},
\]  
(28)
ONE-DIMENSIONAL MODE ALGORITHM FOR TWO-DIMENSIONAL FREQUENCY ESTIMATION

Dunmin Zheng 1  Jian Li 1  Petre Stoica 2

1 Department of Electrical & Computer Engineering, University of Florida, Gainesville, FL 32611-2044, USA.
2 Department of Technology, Uppsala University, P. O. Box 27, S-751 03 Uppsala, Sweden.

Abstract
This paper describes how the computationally efficient one-dimensional MODE (1D-MODE) algorithm can be used to estimate the frequencies of two-dimensional complex sinusoids. We show that the 1D-MODE algorithm is computationally more efficient than the asymptotically statistically efficient 2D-MODE algorithm, especially when the numbers of spatial measurements are large. We find that the 1D-MODE algorithm is asymptotically statistically efficient for high signal-to-noise ratio. We also show that although 1D-MODE is no longer statistically efficient when the number of temporal snapshots is large, the performance of 1D-MODE can still be very close to that of the 2D-MODE under mild conditions. Numerical examples comparing the performances of the 1D-MODE and 2D-MODE algorithms are also presented.

1. INTRODUCTION

In [1], we presented a two-dimensional MODE (2D-MODE) algorithm for estimating 2-D frequencies. There are many applications for 2-D frequency estimation, which include angle-of-arrival estimation with a 2-D sensor array and synthetic aperture radar imaging [1]. Compared with the exact maximum likelihood estimator, the 2D-MODE algorithm avoids the multidimensional search over the parameter space [2]. Yet 2D-MODE has been shown to be statistically efficient under the assumption that the number of temporal snapshots is large or the signal-to-noise ratio (SNR) is high.

The purpose of this paper is to describe how the computationally efficient one-dimensional MODE (1D-MODE) algorithm [3] can be used to estimate the frequencies of 2-D complex sinusoids. We show that the 1D-MODE algorithm is computationally more efficient than the 2D-MODE, especially when the numbers of spatial measurements are large. We also find the 1D-MODE algorithm is statistically efficient for high signal-to-noise ratio (SNR). Even though the 1D-MODE algorithm is no longer statistically efficient when the number of temporal snapshots is large, its performance can still be very close to that of the 2D-MODE under mild conditions. Numerical examples comparing the performances of the 1D-MODE and 2D-MODE algorithms are included in this paper.

2. PROBLEM FORMULATION

Consider the following model of 2-D complex sinusoids in additive noise:

\[ y_{m,\overline{m}}(t_n) = \sum_{k=1}^{K} \sum_{\overline{k}=1}^{\overline{M}} a_{k,\overline{k}}(t_n) e^{j(\omega_k m + \mu_{k} \overline{m})} + \epsilon_{m,\overline{m}}(t_n), \]

where \( m = 1, 2, \ldots, M \), \( \overline{m} = 1, 2, \ldots, \overline{M} \), and \( n = 1, 2, \ldots, N \). We refer to \( \overline{M} \) (\( M > K \)) and \( M \) (\( \overline{M} > \overline{K} \)) as the numbers of spatial measurements, and to \( N \) as the number of temporal snapshots. The additive noise \( \epsilon_{m,\overline{m}}(t_n) \) is assumed to be a complex Gaussian random process with zero-mean and

\[ E\{\epsilon_{m,\overline{m}}(t_n)\epsilon_{m,\overline{m}}^*(t_n)\} = \sigma^2 \delta_{m,\overline{m}}, \]

where \( \cdot^* \) denotes the complex conjugate and \( \delta_{m,\overline{m}} \) denotes the Kronecker delta. The \( \epsilon_{m,\overline{m}}(t_n) \) are independent of each other and the complex sinusoids. The complex amplitudes \( a_{k,\overline{k}}(t_n) \), \( k = 1, 2, \ldots, K \), \( \overline{k} = 1, 2, \ldots, \overline{K} \), may be modeled either as the stochastic (or unconditional) signal model or as the deterministic (or conditional) signal model [4, 5].

Let \( Y(t_n) \) and \( E(t_n) \) be \( M \times \overline{M} \) matrices whose \( (m, \overline{m}) \)th elements, respectively, are \( y_{m,\overline{m}}(t_n) \) and \( \epsilon_{m,\overline{m}}(t_n) \). Define \( X(t_n) \) to be a \( K \times \overline{K} \) matrix whose \( (k, \overline{k}) \)th element is \( a_{k,\overline{k}}(t_n) \). Let

\[ A = \begin{bmatrix} a(\omega_1) & \cdots & a(\omega_K) \end{bmatrix}, \]

\[ a(\omega_k) = \begin{bmatrix} e^{j\omega_k} & \cdots & e^{jM\omega_k} \end{bmatrix}^T, \]

\[ B = \begin{bmatrix} b(\mu_1) & \cdots & b(\mu_{\overline{M}}) \end{bmatrix}, \]

\[ b(\mu_k) = \begin{bmatrix} e^{j\mu_k} & \cdots & e^{j\overline{M}\mu_k} \end{bmatrix}^T, \]

where \( k = 1, 2, \ldots, K \), \( \overline{k} = 1, 2, \ldots, \overline{K} \), and \( (\cdot)^T \) denotes the transpose. Then \( Y(t_n) \) can be written as

\[ Y(t_n) = AX(t_n)B^T + E(t_n). \]

The problem of interest herein is to estimate \( \omega_1, \omega_2, \ldots, \omega_K \) and \( \mu_1, \mu_2, \ldots, \mu_{\overline{M}} \) from \( Y(t_n) \), \( n = 1, 2, \ldots, N \).

3. 2-D FREQUENCIES ESTIMATES WITH 1D-MODE

First consider using 1D-MODE to estimate \( \omega = [\omega_1, \omega_2, \ldots, \omega_K]^T \). Let

\[ \hat{R}_\omega = \frac{1}{N} \sum_{n=1}^{N} Y(t_n)Y^H(t_n), \]

where \( (\cdot)^H \) denotes the complex conjugate transpose and \( \hat{R}_\omega \) is the estimate of the following spatial covariance matrix:

\[ R_\omega = E\{Y(t_n)Y^H(t_n)\} = AP \omega A^H + M \sigma^2 I, \]