Efficient Mixed-Spectrum Estimation with Applications to Target Feature Extraction

Jian Li, Member, IEEE, and Petre Stoica, Fellow, IEEE

Abstract—In this paper, we present a decoupled parameter estimation (DPE) algorithm for estimating sinusoidal parameters from both 1-D and 2-D data sequences corrupted by autoregressive (AR) noise. In the first step of the DPE algorithm, we use a relaxation (RELAX) algorithm that requires simple fast Fourier transforms (FFT's) to obtain the estimates of the sinusoidal parameters. We describe how the RELAX algorithm may be used to extract radar target features from both 1-D and 2-D data sequences. In the second step of the DPE algorithm, a linear least-squares approach is used to estimate the AR noise parameters. The DPE algorithm is both conceptually and computationally simple. The algorithm not only provides excellent estimation performance under the model assumptions, in which case the estimates obtained with the DPE algorithm are asymptotically statistically efficient, but is also robust to mismodeling errors.

I. INTRODUCTION

Many algorithms have been devised to estimate signal parameters for sinusoids in additive white noise. The success of those algorithms is typically demonstrated by means of simulated numerical examples. When such algorithms are used in practical applications, however, the results are often not as encouraging. One of the main reasons for this to occur is that the aforementioned data model assumption is often violated in practical applications.

More robust sinusoidal parameter estimation algorithms may be derived by relaxing the additive white noise assumption. In general, the additive noise may be modeled as unknown autoregressive (AR) noise [1], [2]. Then, the estimation of the sinusoidal parameters and AR noise coefficients becomes the problem of mixed-spectrum estimation since the sinusoids have a discrete spectrum, and the AR noise has a continuous spectrum [3]. Other noise modeling approaches are also possible. For example, the noise covariance matrix may be modeled as an unknown Toeplitz matrix that has an embedding in a larger circulant matrix [4]. The price paid for the aforementioned more general data modeling approaches, however, is that the algorithms derived to jointly estimate the sinusoidal and noise parameters are computationally demanding [1], [2], [4].

In this paper, we present a decoupled parameter estimation (DPE) algorithm for estimating sinusoidal and AR noise parameters from both 1-D and 2-D data sequences. In the first step of the DPE algorithm, a nonlinear least-squares criterion is minimized to obtain the estimates of the sinusoidal parameters. These estimates are used in the second step of the DPE algorithm, which estimates the AR noise parameters by a linear least-squares approach.

To minimize the nonlinear least-squares criterion in the first step of DPE, we propose a relaxation (RELAX) algorithm that requires simple fast Fourier transforms (FFT’s). We shall explain how the RELAX algorithm is related to the CLEAN algorithm [5] and its improved version [6], which have been shown to provide good performance with both numerical and experimental data [5]-[7]. Since the alternating projection (AP) [8] and the related alternating notch-periodogram (ANPA) [9] algorithms can also be used to minimize the nonlinear least-squares criterion in the first step of DPE, we shall compare the performance of these algorithms with that of the RELAX algorithm. We also describe in detail how the RELAX algorithm may be used to extract radar target features from both 1-D and 2-D data sequences.

Regarding the “detection problem,” we propose the use of a generalized Akaike information criterion (GAIC) as a solution. Together with DPE and RELAX, GAIC can be used to determine the number of sinusoids and the AR noise model order.

Numerical examples are given in the paper to demonstrate the performance of the proposed parameter estimation algorithms and of the order detection criterion.

In Section II, we formulate the problem of interest for both 1-D and 2-D data sequences. In Section III, we present the DPE and RELAX algorithms for parameter estimation and the GAIC criterion for order detection. In Section IV, we present numerical examples showing the performance of the proposed parameter estimation algorithms and the order detection criterion. We also describe in detail how the RELAX algorithm may be used to extract target features with high-resolution radar and synthetic aperture radar (SAR). Finally, Section V contains our conclusions.

II. PROBLEM FORMULATION

In this section, we formulate the problems of both 1-D and 2-D parameter estimation of sinusoids in AR noise. The 1-D parameter estimation problem occurs, for example, in high-
resolution range signature estimation of a radar target. The 2-D parameter estimation problem occurs, for example, in target feature extraction from synthetic aperture radar (SAR) phase histories. High-resolution range signatures and SAR target features are useful for many applications such as the noncooperative target identification (NCTI). In the NCTI application, the target features are extracted from both training data sets and the data set under investigation. These features are used in pattern recognition algorithms for target identification.

Consider first the problem of 1-D parameter estimation of K complex sinusoids in complex AR noise. The data model for this problem can be written as

\[ y_n = \sum_{k=1}^{K} \alpha_k e^{2\pi f_k n} + e_n \quad (1) \]

where \( n = 0, 1, \ldots, N - 1 \), with \( N \) denoting the number of data samples; \( \alpha_k, k = 1, 2, \ldots, K \) denotes the unknown complex amplitude of the \( k \)th sinusoid; \( f_k, k = 1, 2, \ldots, K \) denotes the unknown frequency of the \( k \)th sinusoid; and \( e_n, n = 0, 1, \ldots, N - 1 \) denotes an \( M \)th-order AR noise, which can be written as

\[ e_n = -\sum_{m=1}^{M} a_m e_{n-m} + w_n \quad (2) \]

where \( a_m, m = 1, 2, \ldots, M \) are the complex AR coefficients of the noise and \( w_n \) is a complex Gaussian white random process with zero-mean and variance \( \sigma^2 \).

Next, consider the problem of 2-D parameter estimation of \( K \) complex sinusoids in complex AR noise. The data model for this problem can be written as

\[ y_{n,\bar{n}} = \sum_{k=1}^{K} \alpha_k e^{(2\pi f_k n + 2\pi \bar{f}_k \bar{n})} + e_{n,\bar{n}} \quad (3) \]

where \( n = 0, 1, \ldots, N - 1 \) and \( \bar{n} = 0, 1, \ldots, \bar{N} - 1 \) with \( N \) and \( \bar{N} \) denoting the numbers of available data samples; \( \alpha_k, k = 1, 2, \ldots, K \) denotes the unknown complex amplitude of the \( k \)th sinusoid; \( f_k \) and \( \bar{f}_k, k = 1, 2, \ldots, K \) denote the 2-D unknown frequencies of the \( k \)th sinusoid; and \( e_{n,\bar{n}}, n = 0, 1, \ldots, N - 1 \) and \( \bar{n} = 0, 1, \ldots, \bar{N} - 1 \) is the 2-D AR noise with orders \( M \) and \( \bar{M} \) whose region of support is on the quarter plane \( [0, \infty] \) in rectangular coordinates, i.e., \( e_n, n = 0, 1, \ldots, N - 1 \), can be written as

\[ e_{n,\bar{n}} = -\sum_{m=1}^{M} \sum_{\bar{m}=1}^{\bar{M}} a_{m,\bar{m}} e_{n-m,\bar{n}-\bar{m}} + w_{n,\bar{n}} \quad (4) \]

where \( a_{m,\bar{m}}, m = 1, 2, \ldots, M, \bar{m} = 1, 2, \ldots, \bar{M} \) are the complex AR coefficients and \( w_{n,\bar{n}} \) is a 2-D Gaussian white random process with zero-mean and variance \( \sigma^2 \).

For the problem of 1-D parameter estimation of real sinusoids in real AR noise, the data model can still be written as in (1) but with the additional specification that \( K \) is an even number, and there are \( K/2 \) real sinusoids with \( \alpha_{2k-1} = \alpha_{2k} \) and \( f_{2k-1} = -f_{2k}, k = 1, 2, \ldots, K/2 \). Further, the real AR noise can also be written as in (2), where all quantities are real. For the problem of 2-D parameter estimation of real sinusoids in real AR noise, the data model in (3) can be modified similarly.

The problem of interest herein is to estimate the mixed-spectrum from the data samples \( y_n, n = 0, 1, \ldots, N - 1 \), or \( y_{n,\bar{n}}, n = 0, 1, \ldots, N - 1, \bar{n} = 0, 1, \ldots, \bar{N} - 1 \). Equivalently stated, we want to estimate the signal parameters (frequencies and amplitudes) and the AR coefficients along with the white noise variance, as well as the number of sinusoids and the AR noise model order. We are particularly interested in estimating the sinusoidal parameters. In NCTI applications, these parameters are radar target features useful for target recognition.

### III. Efficient Parameter Estimation and Order Detection Algorithms

The maximum likelihood (ML) estimator of the unknown parameters of the mixed-spectrum data model is asymptotically (for large \( N \) or \( N \) and \( \bar{N} \)) statistically efficient. Yet, the ML estimator requires a multidimensional search over the signal and noise parameter space \([1, 2, 4, 11]\). Since the search over a high-dimensional parameter space is computationally prohibitive, computationally more efficient methods are attractive. We describe below a conceptually and computationally simple method that, similarly to the ML estimator, provides asymptotically statistically efficient estimates.

We first describe our approach for the data model given in (1) for which we estimate the 1-D parameters of the complex sinusoids and of the AR noise by assuming that the number of sinusoids and the AR model order are known. We then propose a generalized information theoretic criterion to estimate the number of sinusoids and the AR model order. Finally, we extend our discussion to mixed-spectrum estimation from real data and for 2-D data sequences.

#### A. Decoupled Parameter Estimation for the Sinusoids and AR Noise

Consider the estimation of the parameters for the complex sinusoids and for the AR noise based on the data model given in (1). Let

\[ \hat{y}_n = c(n) + a_1 c(n-1) + \ldots + a_M c(n-M), \]

\[ n = 0, 1, \ldots, N - 1 \quad (5) \]

where

\[ e_n = y_n - \sum_{k=1}^{K} \alpha_k e^{2\pi f_k n}, \quad n = 0, 1, \ldots, N - 1 \quad (6) \]

with \( y_n \) set to zero for \( n < 0 \). Then, the asymptotic (for \( N \gg 1 \)) negative log-likelihood function of the observed data sequence \( \{y_n\} \) in (1) is given by

\[ C_1 = N \ln(\pi) + N \ln(\sigma^2) + \frac{1}{\sigma^2} \tilde{y}^H \tilde{y} \quad (7) \]

where \((\cdot)^H\) denotes the complex conjugate transpose, and \( \tilde{y} = [\tilde{y}_0 \quad \tilde{y}_1 \quad \ldots \quad \tilde{y}_{N-1}]^T \) with \((\cdot)^T\) denoting the transpose. The
achieve the Cramér-Rao bound (CR-bound) in spite of the
DPE, are also asymptotically statistically efficient. Therefore,
estimates of a maximum likelihood estimator. Hence, the least-squares
to estimate the sinusoidal parameters decouples the estimation
of the sinusoidal parameters from the estimation of the noise
model parameters. Thus, we can, in principle, consider more
general noise models, such as autoregressive moving average
(ARMA) models, for the noise without significant difficulties.

We now consider the implementations of the two steps of
the DPE algorithm. We first describe the implementation of
Step 2 since this step is straightforward. Later, we present the
implementation of Step 1 in detail.

1) The Least-Squares Estimator for AR Noise Parameters:
The estimates of \( \{a_m\} \) are determined as the least-squares
solution to the following overdetermined (for \( N > M \)) system
of linear equations:

\[
\begin{bmatrix}
\hat{e}_0 & 0 & \cdots & 0 \\
\hat{e}_1 & \hat{e}_0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
\hat{e}_{N-2} & \hat{e}_{N-3} & \cdots & \hat{e}_{N-M-1}
\end{bmatrix}
\begin{bmatrix}
a_1 \\
a_2 \\
\vdots \\
a_M
\end{bmatrix}
= \begin{bmatrix}
\hat{\epsilon}_1 \\
\hat{\epsilon}_2 \\
\vdots \\
\hat{\epsilon}_{N-1}
\end{bmatrix}
\tag{14}
\]
or (in matrix form)

\[
\hat{\Theta}_n \approx \hat{\Theta}
\tag{15}
\]
where \( \hat{\epsilon}_n, n = 0, 1, \ldots, N - 1 \) has been defined in (13).

2) The RELAX Algorithm for Sinusoidal Parameter Estimation:
To minimize the cost function in (12), we could concentrate
out the complex amplitudes \( \alpha = [\alpha_1 \alpha_2 \cdots \alpha_K]^T \)
and then minimize the concentrated function with respect to
\( \{f_k\} \). More specifically, let

\[
\hat{\Theta} = \left[ \Theta(f_1) \Theta(f_2) \cdots \Theta(f_K) \right]
\]

Then, we have

\[
\{\hat{f}_k, \hat{\alpha}_k\} = \arg \min_{\{f_k, \alpha_k\}} \| \hat{\Theta} - \hat{\Theta} \alpha \|^2.
\tag{17}
\]

Minimizing the right side of (17) with respect to \( \alpha \) gives

\[
\hat{\alpha} = \left( \Omega^H \Omega \right)^{-1} \Omega^H \hat{\Theta}.
\tag{18}
\]

Then, the estimates of the frequencies \( f = [f_1 f_2 \cdots f_K]^T \)
are obtained by minimizing the following concentrated function

\[
C_3(f_1, f_2, \cdots, f_K) = \left\| P_{\Omega} \hat{\Theta} \right\|^2
\tag{19}
\]

where \( P_{\Omega} \) stands for the orthogonal projector onto the null
space of \( \Omega^H \), i.e.

\[
P_{\Omega} = I - \Omega \left( \Omega^H \Omega \right)^{-1} \Omega^H
\tag{20}
\]
with \( I \) denoting the identity matrix. The frequency estimates
can be obtained by minimizing (19) with the AP algorithm pro-
posed in [8] or the related ANPA in [9]. The implementation
and the associated computational burden of these algorithms
are discussed in detail in [9].

Yet, we shall show below and with numerical examples in
Section IV that the concentration of the cost function in Step
1 above with respect to \( \alpha \), instead of leading to a simpler
problem, actually complicates the optimization of the cost
function. More exactly, we show that the relaxation-based
minimization of the cost function in (12) with respect to both
\( f \) and \( \alpha \) actually leads to a conceptually and numerically much
simpler algorithm.
Before we present our approach to minimize the cost function in Step 1 above, let us consider the following preparations. Let
\[
y_k = y - \sum_{i=1, i \neq k}^{K} \hat{\alpha}_i \omega(f_i)
\]
(21)
where \{\hat{f}_1, \hat{\alpha}_1\}_{i=1, i \neq k}^K are assumed to be given. Then, the minimization of the cost function in Step 1 above with respect to \(f_k\) and \(\alpha_k\), via using (18) and (19), gives
\[
\hat{\alpha}_k = \frac{\omega^H(f_k)y_k}{N} \bigg|_{f_k=f_k}
\]
and [12]
\[
\hat{f}_k = \arg \min f_k \left\| \left[ 1 - \frac{\omega(f_k)\omega^H(f_k)}{N} \right] y_k \right\|^2
= \arg \max f_k \left| \omega^H(f_k)y_k \right|^2.
\]
(22)
(23)
Hence, \(\hat{f}_k\) is obtained as the location of the dominant peak of the periodogram \(\left| \omega^H(f_k)y_k \right|^2/N\), which can be efficiently computed by using the FFT with the data sequence \(y_k\) padded with zeros. (Note that padding with zeros is necessary to determine \(\hat{f}_k\) with high accuracy.) Then, \(\hat{\alpha}_k\) is easily computed from the complex height of the peak of \(\omega^H(f_k)y_k/N\).

With the above simple preparations, we now proceed to present the relaxation algorithm for the minimization of the nonlinear least-squares cost function in Step 1 of DPE. The subject relaxation algorithm is referred to as RELAX below.

**Step (0):** Assume \(K = 0\). (Hence, only the AR noise exists.)

**Step (1):** Assume \(K = 1\). Obtain \(\hat{f}_1\) and \(\hat{\alpha}_1\) from \(y\) as described above.

**Step (2):** Assume \(K = 2\). Compute \(y_2\) with (21) by using \(\hat{f}_1\) and \(\hat{\alpha}_1\) obtained in Step (1). Obtain \(\hat{f}_2\) and \(\hat{\alpha}_2\) from \(y_2\) as described above.

Next, compute \(y_1\) with (21) by using \(\hat{f}_2\) and \(\hat{\alpha}_2\), and redetermine \(\hat{f}_1\) and \(\hat{\alpha}_1\) from \(y_1\) as in (22) and (23) above.

Iterate the previous two substeps until "practical convergence" is achieved (to be discussed later on).

**Step (3):** Assume \(K = 3\). Compute \(y_3\) with (21) by using \(\hat{f}_1, \hat{\alpha}_1\) obtained in Step (2). Obtain \(\hat{f}_3\) and \(\hat{\alpha}_3\) from \(y_3\) as described above.

Next, compute \(y_1\) with (21) by using \(\hat{f}_3, \hat{\alpha}_3\), and redetermine \(\hat{f}_1\) and \(\hat{\alpha}_1\) from \(y_1\). Then, compute \(y_2\) with (21) by using \(\hat{f}_3, \hat{\alpha}_3\) and redetermine \(\hat{f}_2\) and \(\hat{\alpha}_2\) from \(y_2\).

Iterate the previous three substeps until "practical convergence."

**Remaining Steps:** Continue similarly until \(K\) is equal to the desired or estimated number of sinusoids.

The "practical convergence" in the iterations of the above RELAX algorithm may be determined by checking the relative change of the cost function \(C_3(\{\hat{f}_1, \hat{\alpha}_1\}_{i=1}^K)\) in (12) between the \(j\)th and \((j+1)\)st iterations. In our numerical examples, we terminate the iterations when the relative change is less than or equal to \(\varepsilon = 10^{-3}\).

We note that the iterations performed in each step of the proposed RELAX algorithm, while time consuming, can have the beneficial effect of producing an excellent initial estimate for the next step, which may in fact lead to a decrease in the total amount of computations needed by the algorithm. The fact that RELAX operates with the periodogram and the corrected periodograms (i.e., the periodograms of \(y_k\)'s in (21)) makes it rather clear that RELAX determines good initial estimates of \(\{f_k, \alpha_k\}\) for each search step. A good initial estimate is an important advantage since it is well known that the cost function in (12) typically has many false local minima [13]. It appears that the specific sequence of the steps of RELAX pushes the search toward the global minimum of (12). Moreover, since the number of sinusoids is often unknown, the parameter estimates obtained in each step, after convergence, are useful to determining the number of sinusoids and the AR model order (see the following subsections). This gives another reason why the iterations are necessary to minimize the cost function in each step of the RELAX algorithm.

The RELAX algorithm can also be referred to as SUPER CLEAN. If we set the number of iterations in each step of the RELAX algorithm to zero, then the RELAX algorithm becomes the CLEAN algorithm, which was first proposed in radio astronomy [5] and later used in microwave imaging [7]. If we redetermine all of the sinusoidal parameters once in each step of the RELAX algorithm, then RELAX becomes the complex version of the algorithm proposed in [6] for real data sequences. We shall refer to the latter algorithm (both the real and complex versions) as MCLEAN (more clean). It follows from the previous discussion that CLEAN and MCLEAN are equivalent to RELAX only for \(K = 1\). Furthermore, both CLEAN and MCLEAN are only approximately a relaxation method for the minimization of the NLS criterion. Yet, it is precisely this interpretation that explains the good performance noticed in the numerical and experimental examples in [5]-[7]. Note that the relationship between the CLEAN and MCLEAN algorithms of [5]-[7] and the (approximate) relaxation minimization of the NLS criterion (12) was apparently overlooked in the cited references. Note also that both CLEAN and MCLEAN assume that the number of sinusoids is known, whereas the RELAX algorithm can be used to determine the number of sinusoids, as described in the next subsection.

We also remark that using RELAX in DPE makes DPE a computationally much more efficient approach than the methods of jointly estimating the sinusoidal and AR noise parameters [1], [2], which we refer to as the joint parameter estimation (JPE) methods. (Note that a JPE method for real sinusoids in real AR noise is given in Appendix B.) The JPE cost function (9) usually has many false local optimum points. Perhaps for this reason, a computationally prohibitive multidimensional search over the parameter space is proposed for JPE in [1] and [2]. DPE, however, only requires simple FFT's and, hence, is computationally much more efficient. This claim will be supported by means of numerical examples in Section IV, where quantitative comparisons between the computational burdens of the DPE and ANPA algorithms are made as well.

Finally, we discuss the asymptotic (for large \(N\)) statistical performance of the RELAX algorithm. The asymptotic statistical distribution of the sinusoidal parameter estimates
obtained with RELAX is Gaussian with mean equal to the true parameters and covariance matrix equal to the corresponding CR-bound, which is derived in Appendix A.

B. Determining the Number of Sinusoids and the AR Model Order

To determine the number of sinusoids $K$ and the model order $M$ of the AR noise, we consider using the following generalized Akaike information criterion (GAIC) (see [14] and [15] and the references therein). We note that when the parameter estimates obtained with the DPE algorithm are used in the cost function in (9), we obtain

$$
V_{K,M} = N \ln \left( \left\| \hat{\theta}_0 + \hat{e} \right\|^2 \right)
$$

which is a function of the assumed number of sinusoids $\hat{K}$ and AR model order $\hat{M}$. Then, the estimates $\hat{K}$ and $\hat{M}$ of $K$ and $M$, respectively, are determined as the integers that minimize the following GAIC cost function:

$$
\text{GAIC}_{\hat{K}, \hat{M}} = V_{\hat{K}, \hat{M}} + \beta (3\hat{K} + 2\hat{M} + 1)
$$

where $(3\hat{K} + 2\hat{M} + 1)$ denotes the total number of unknown real parameters for the data model in (1) (of which $3\hat{K}$ for the sinusoids and $2\hat{M} + 1$ for the AR noise model), and $\beta$ is a user’s variable. In our numerical examples, we have used $\beta = 4\ln(\ln N)$ for 1-D complex data sequences. It is known that the iterated logarithm $\ln(\ln N)$ gives the lowest rate at which $\beta$ should increase with $N$ to guarantee the consistency of the order estimates obtained by minimizing (25) [16]. With this fact in mind, we have selected $\beta$ proportional to $\ln(\ln N)$ to achieve consistency on the one hand and to reduce the risk of underfitting on the other hand.

C. Modifications and Extensions

The DPE algorithm for estimating sinusoidal and AR noise parameters for complex data sequences can be modified to deal with real 1-D data sequences. For the latter sequences, the DPE algorithm becomes the following:

**Step 1:** Estimate the sinusoidal parameters by minimizing the following nonlinear least-squares criterion:

$$
C_5(f_1, \alpha_1, \ldots, f_K, \alpha_K) = \left\| \mathbf{y} - \sum_{k=1}^{K/2} \left[ \omega(f_{2k-1}) \alpha_{2k-1} + \omega(-f_{2k-1}) \alpha^*_{2k-1} \right] \right\|^2
$$

**Step 2:** Estimate the AR noise parameters from

$$
\hat{e}_n = y_n - \sum_{k=1}^{K/2} \left[ \hat{\alpha}_{2k-1} e^{2\pi f_{2k-1} n} + \hat{\alpha}^*_{2k-1} e^{-2\pi f_{2k-1} n} \right]
$$

where $\{\hat{\alpha}_{2k-1}, \hat{\alpha}^*_{2k-1}\}_{k=1}^{K/2}$ are obtained in Step 1.

The periodogram/FFT-based RELAX algorithm presented for complex data sequences can also be modified to minimize the cost function in Step 1 above. Since the modification is straightforward, we omit the details. We remark, however, that similarly to the method in [6], the FFT-based RELAX does not exploit the structure in (26) and, hence, is an approximate approach.

Let

$$
W(f_k) = 2 \begin{bmatrix} 1 & \cos(2\pi f_k) & \cdots & \cos(2\pi f_k (N - 1)) \\ 0 & \sin(2\pi f_k) & \cdots & \sin(2\pi f_k (N - 1)) \end{bmatrix}^T
$$

Then, (22) and (23) can be modified, respectively, as follows to exploit the structure in (26):

$$
\begin{bmatrix} \text{Re}(\hat{\alpha}_k) \\ \text{Im}(\hat{\alpha}_k) \end{bmatrix} = \left[ W^T(f_k) W(f_k) \right]^{-1} W^T(f_k) y_k |_{f_k = f_k},
$$

$$
\hat{f}_k = \arg \min_{f_k} \left\| P^*_{W(f_k)} y_k \right\|^2, \quad k = 1, 3, \ldots, K - 1.
$$

Note that more computations are needed when we exploit the structure in (26).

For real 1-D data sequences, the GAIC cost function becomes

$$
\text{GAIC}_{\hat{K}, \hat{M}} = V_{\hat{K}, \hat{M}} + \beta (3\hat{K}/2 + \hat{M} + 1)
$$

where $(3\hat{K}/2 + \hat{M} + 1)$ denotes the total number of unknown real parameters for the real data model (of which $3\hat{K}/2$ for the sinusoids and $\hat{M} + 1$ for the AR noise model). In our numerical examples, we have used $\beta = 8\ln(\ln N)$ for real data sequences. The motivation for doubling the coefficient of $\ln(\ln N)$ in the expression of $\beta$ is as follows. In the complex-valued data case, the minimum negative log-likelihood function is $V_{\hat{K}, \hat{M}}$. In the real-valued data case, however, the similar minimum value is $1/2 V_{\hat{K}, \hat{M}}$. Hence, from the detection performance standpoint, $\beta = 4\ln(\ln N)$ in the complex case should be equivalent to $\beta = 8\ln(\ln N)$ in the real one.

Both the DPE and the RELAX algorithms can also be extended straightforwardly to estimate 2-D complex sinusoidal and AR noise parameters. Let

$$
\mathbf{y} = [y_{0,0}, \ldots, y_{0,N-1}, \ldots, y_{N-1,0}, \ldots, y_{N-1,N-1}]^T.
$$

Then, for 2-D complex data sequences, the DPE algorithm becomes the following:

**Step 1:** Estimate the sinusoidal parameters $\{f_k, \tilde{f}_k, \alpha_k\}$ by minimizing the following nonlinear least-squares criterion:

$$
C_0(f_1, \alpha_1, \ldots, f_K, \tilde{f}_K, \alpha_K) = \sum_{n=0}^{N-1} \sum_{\tilde{n}=0}^{N-1} \left\| \mathbf{y} - \sum_{k=1}^{K} \alpha_k \omega(f_k + 2\pi \tilde{f}_k \tilde{n}) \right\|^2
$$

where $\omega(f_k), k = 1, 2, \ldots, K$ are $N \times 1$ vectors that are defined similarly to $\omega(f_k)$, and $\otimes$ denotes the Kronecker product [17].
Step 2: Estimate the AR noise parameters from
\[
\hat{\theta}_{n,n'} = y_{n,n'} - \sum_{k=1}^{K} \hat{\alpha}_k e^{(2\pi f_k n + 2\pi f_k n')} \tag{34}
\]
where \(\{f_k, \hat{\alpha}_k\}\) are obtained in Step 1.

The RELAX algorithm presented for 1-D complex data sequences can readily be modified to estimate the AR coefficients of 2-D data sequences. The least-squares estimator proposed for estimating the AR coefficients of 2-D data sequences can also be modified easily to estimate the AR coefficients of 2-D data sequences in Step 2 above. Since the modifications are straightforward, the details are omitted.

Let
\[
\hat{w}_{n,n'} = \hat{\theta}_{n,n'} + \sum_{m=1}^{M} \sum_{n=1}^{N} \hat{\alpha}_{m,n'} \hat{\theta}_{n-m,n-n'} \tag{35}
\]
where \(\{\hat{\alpha}_{m,n'}\}\) are estimated with the least-squares approach similarly to the 1-D case. Let
\[
V_{K,M} = N\tilde{N} \ln \left( \sum_{n=1}^{N} \sum_{n=1}^{N} |\hat{w}_{n,n'}|^2 \right) \tag{36}
\]
Then, for 2-D complex data sequences, the GAIC cost function becomes
\[
\text{GAIC}_{K,M} = V_{K,M} + \beta(4\hat{K} + 2\hat{M} + 2\tilde{K} + 1) \tag{37}
\]
where \((4\hat{K} + 2\hat{M} + 2\tilde{K} + 1)\) is the total number of unknown real parameters (of which \(4\hat{K}\) for the sinusoids and \(2\hat{M} + 2\tilde{K} + 1\) for the AR noise model). In our numerical examples, we have used \(\beta = 4 \ln(\ln(N\tilde{N}))\) for the 2-D complex data sequences.

IV. NUMERICAL AND EXPERIMENTAL RESULTS

We present numerical examples showing the performance of the proposed algorithms for estimating sinusoidal and the AR noise parameters and for determining the number of sinusoids and the AR model order.

We first present numerical examples of mixed-spectrum estimation for both real and complex data sequences. The proposed parameter estimation algorithms are compared with the ANPA and JPE methods. Next, we present numerical examples of using the RELAX algorithm to extract radar target features (dominant scatterers) from both high-resolution radar data and synthetic aperture radar phase histories.

In the examples below, all sequences are zero padded to 1024 before used with FFT in the RELAX algorithm. We have used \(\epsilon = 0.001\) to test the convergence of the RELAX algorithm for all of the examples below.

We remark that in all of the examples below, each step of the RELAX algorithm converges quickly, usually in a few iterations. In these examples, the sinusoids are sometimes so closely spaced that they cannot be resolved by the periodogram. When the sinusoids become even more closely spaced, however, the speed of convergence decreases. Furthermore, to obtain accurate parameter estimates in such a case, the data sequences should be padded with more zeros, and a smaller \(\epsilon\) should be chosen. To understand this behavior, consider a 1-D sequence containing two very closely spaced complex sinusoids. Since the sinusoids are so closely spaced, the \(y_2\) obtained in the first iteration in Step (2) of RELAX will be close to zero, which causes Step (2) to converge very slowly. We found that to speed up the convergence, we may choose \(f_3\) and \(\hat{\alpha}_1\) in Step (1) to correspond to one half of the dominant peak of the periodogram rather than the dominant peak itself. Further studies are needed to further improve the convergence rate of RELAX when some of the sinusoids are very closely spaced.

A. Mixed Spectrum Estimation

We first consider estimating the mixed-spectrum for a real data sequence whose true normalized power spectral density (PSD) is shown in Fig. 1(a) and whose data length is \(N = 64\). The data sequence is the test sequence used in [18] and consists of three real sinusoids and a colored noise process obtained by filtering a white Gaussian process. The three sinusoids are at frequencies 0.10, 0.20, and 0.21 Hz and have signal-to-noise ratios (SNR’s) of 10, 30, and 30 dB, respectively, where each SNR is defined as the ratio of the sinusoid power to the total power of the noise process [18]. Fig. 1(b) shows the estimated spectrum of the data sequence by using the periodogram. Note that the two dominant sinusoids are too closely spaced to be resolved by the periodogram. The colored noise model is unknown, and we assume that the noise is an AR noise process. Since the number of sinusoids \(K\) and the noise model order \(M\) are both assumed unknown, we use the GAIC criterion to estimate both \(K\) and \(M\). The estimates for \(K\) and \(M\) are, respectively, \(\hat{K} = 2\) and \(\hat{M} = 2\).

Using the \(\hat{K} = 2\) and \(\hat{M} = 2\) with the FFT-based DPE algorithm, we obtain the mixed-spectrum estimate as shown in Fig. 1(c). Fig. 1(d) shows the mixed-spectrum estimate obtained by using the JPE method described in Appendix B with \(\hat{K} = 2\) and \(\hat{M} = 2\). Note that due to the underestimation of \(K\), the sinusoid at 0.1 Hz is missing in Fig. 1(c) and (d). Note also that both DPE and JPE can not only resolve the two dominant sinusoids but also accurately estimate their parameters, unlike the periodogram. We note that the DPE and JPE algorithms give similar estimates for the mixed-spectrum for \(\hat{K} = 2\) and \(\hat{M} = 2\). However, the DPE algorithm is computationally much more efficient than the JPE algorithm.

To determine the sinusoidal parameters, the FFT-based DPE uses RELAX, which requires simple FFT’s, while JPE requires a computationally prohibitive multidimensional search over the parameter space [2]. For \(\hat{K} = 2\), JPE requires a 2-D search over \(f_1\) in \([0, 0.5]\) and \(f_3\) in \([0, 0.5]\). The step size chosen for the search for JPE was 1/512 Hz. (Recall that the data sequences are zero padded to 1024 before used with FFT in RELAX.) With these choices, the amount of computations required by JPE is \(5.4 \times 10^4\) times that required by DPE. As the number of sinusoids increases, JPE will be computationally even more costly than DPE.

We have also tried other choices of \(\hat{K}\) and \(\hat{M}\). Using \(\hat{K} = 2\) and \(\hat{M} = 10\) with the FFT-based DPE algorithm and the JPE algorithm, we obtain the mixed-spectrum estimates as shown
Fig. 1. True and estimated spectra for a real data sequence with \( N = 64 \): (a) True spectrum; (b) estimated with the periodogram; (c) estimated with FFT-based DPE and \( K = M = 2 \); (d) estimated with JPE and \( K = M = 2 \); (e) estimated with FFT-based DPE, \( K = 2 \), and \( M = 10 \); (f) estimated with JPE, \( K = 2 \), and \( M = 10 \); (g) estimated with FFT-based DPE, \( K = 3 \), and \( M = 8 \); (h) estimated with FFT-based RELAX, \( K = 10 \), and \( M = 0 \); (i) estimated with real-RELAX, \( K = 10 \), and \( M = 0 \).
in Fig. 1(e) and (f), respectively. We note that using a larger \( M \) with DPE yields a better spectrum estimate. However, the spectrum estimate obtained with JPE is worse, and the two lines obtained with JPE are around the sinusoid at frequency 0.1 Hz, which has the lowest SNR among the three sinusoids. Hence, JPE is not as robust as DPE when model errors exist. JPE is also sensitive to the choice of \( M \), whereas DPE is not.

Fig. 1(g) shows the mixed-spectrum estimate obtained by using \( K = 3 \) and \( M = 8 \) with the FFT-based DPE algorithm. We note that since the sinusoid at frequency 0.1 Hz is below the peak noise level, the parameters of the sinusoid cannot be estimated accurately by using DPE with \( K = K = 3 \). Since the brute-force search over the 3-D parameter space is computationally burdensome, we did not use JPE with \( K = 3 \) and \( M = 8 \).

Fig. 1(h) and (i) show the line spectrum estimates obtained by using \( K = 10 \) with, respectively, the FFT-based RELAX and with the RELAX algorithm that exploits the structure of the real 1-D data sequence. We shall refer to the latter as the real-RELAX. We note that although the real-RELAX yields slightly more accurate line estimates than the FFT-based RELAX, the two RELAX approaches give rather similar spectrum estimates. The amount of computations required by the real-RELAX, however, is about 20 times more than that required by the FFT-based RELAX. Thus, the FFT-based RELAX algorithm is in general preferred over the real-RELAX algorithm. We also note that the shapes of the line spectrum estimates obtained with RELAX resemble the true mixed spectrum. The parameters of the three sinusoids are accurately estimated by using RELAX with \( K = 10 \).

We next consider estimating the mixed spectrum for the following complex data sequence:

\[
y_n = 2 \cos(2\pi f_1 n) + 2 \cos(2\pi f_3 n) + 2 \cos(2\pi f_5 n) + e_n,
\quad n = 0, 1, \cdots, N - 1
\]

where \( N = 32, f_1 = 0.05 \text{ Hz}, f_3 = 0.40 \text{ Hz}, \) and \( f_5 = 0.42 \text{ Hz} \).
Hz; $e_n$ is a complex AR process of order 1

$$e_n = -a_1 e_{n-1} + w_n$$  \(39\)

with $a_1 = -0.850848$, and $w_n$ is a zero-mean complex Gaussian white random variable with variance $\sigma^2 = 0.101043$. This data model is used to generate a test data sequence in [10]. Since the data sequence is complex, we assume that the sinusoids are complex. The local SNR's of the complex sinusoids at frequencies $\pm 0.40$ and $\pm 0.42$ Hz are about 30 dB, and the SNR's of the complex sinusoids at frequencies $\pm 0.05$ Hz are about 15 dB [10].

Since the number of sinusoids $K$ and the noise model order $M$ for the above data model are assumed unknown, we apply the GAIC criterion to the test sequence in [10] to estimate both $K$ and $M$. The estimates for $K$ and $M$ are, respectively, $\hat{K} = 7$ and $\hat{M} = 0$. We note that the GAIC criterion considers the peaky AR(1) noise process above as a complex sinusoid.

Fig. 2(a) shows the true normalized spectrum of the test sequence in [10]. Fig. 2(b) shows the estimated spectrum by using the periodogram. Note that the two closely spaced sinusoids are not resolved by the periodogram. Using the GAIC order estimates $\hat{K} = 7$ and $\hat{M} = 0$ with the DPE algorithm, we obtain the spectrum estimate as shown in Fig. 2(c). Using the true model orders, i.e., $K = 6$ and $M = 1$, with the DPE algorithm, we obtain the spectrum estimate shown in Fig. 2(d). We note that the two closely spaced sinusoids are resolved in both Fig. 2(c) and (d). Note also that the sinusoidal parameter estimates obtained with the DPE algorithm are not sensitive to the model orders $K$ and $M$ used to obtain the estimates.

We now compare the estimation performance of ANPA, CLEAN, MCLEAN, and RELAX when using $\hat{K} = 7$ and $\hat{M} = 0$ with the data model in (38). Fig. 3 shows the mean-squared error (MSE) of the frequency and complex amplitude
We next note that the RELAX algorithm provides significantly better frequency estimates than the MCLEAN algorithm. For some sinusoids, the amplitude estimates obtained with RELAX are better than those obtained with MCLEAN, but for other sinusoids, the RELAX amplitude estimates are slightly worse. The overall amplitude estimation performance of RELAX, however, is better than that of MCLEAN. The amount of computations required by RELAX is approximately 3.7 times more than that required by MCLEAN.

We also note that although RELAX and ANPA attempt to minimize the same nonlinear least-squares cost function, the performance of both frequency and amplitude estimation of RELAX is much better than that of ANPA. Yet, the amount of computations required by ANPA is approximately 3.9 times of that required by RELAX. The significantly poorer performance of ANPA occurs because the initial conditions obtained with the ANPA algorithm are poorer than those obtained with the RELAX algorithm. The poor initial conditions may cause the ANPA algorithm to converge to local instead of global optimum points. Note that the amplitude estimation performance of ANPA is significantly poorer even when compared with that of the CLEAN and MCLEAN algorithms. This occurs because ANPA obtains the frequency estimates first, and those estimates may be very closely spaced, as may be seen from Fig. 4(a). When very closely spaced frequency estimates are used in (18) to estimate the sinusoidal amplitudes, the amplitude estimates may be very poor due to the ill conditioning of the matrix to be inverted in (18). This behavior may be seen in Fig. 4(b).

Finally, we use the results derived in Appendix A to calculate the CRB’s for the sinusoidal parameter estimates. These CRB’s are also shown in Fig. 3. (Note that the CRB for each complex amplitude estimate is equal to the sum of the CRB’s for the real and imaginary parts of the complex amplitude estimate.) We note that the RELAX algorithm is an asymptotically (for large $N$) statistically efficient estimator. Yet, even for the small value of $N = 32$ in this example, the variances of the sinusoidal parameter estimates obtained with the RELAX algorithm (see Fig. 3) are close to the corresponding CRB’s (which are the lowest possible variances in the class of asymptotically unbiased estimators).

estimates obtained with 100 independent Monte-Carlo trials. The frequency estimates are sorted from small to large. The smallest frequency estimate is considered to be the estimate for the smallest frequency, and so on. To test the convergence of the ANPA algorithm, we have used 0.001 as the maximum relative change threshold for the frequency estimates obtained from consecutive iterations.

We first note that both the frequency and amplitude estimates of the sinusoids obtained with MCLEAN are better than those obtained with CLEAN. The better performance achieved by MCLEAN is at the cost of more computations. The total amount of computations required by MCLEAN for the 100 Monte Carlo trials is approximately 4.9 times of that needed by CLEAN.

**B. Radar Target Feature Extraction**

We illustrate below the performance of using RELAX with both 1-D and 2-D radar data sequences to extract target features of interest, such as the dominant scatterers.

1) **High-Resolution Range Signature Estimation with RELAX:** We first briefly describe the high-resolution radar range signature estimation problem. We then provide numerical examples showing the performance of using the RELAX algorithm for range signature estimation and target feature extraction.

The range resolution of a radar is determined by the radar bandwidth. To achieve high resolution in range, the radar must transmit wideband pulses, which are often linear frequency modulated (FM) chirp pulses. A normalized chirp pulse can...
where $c$ denotes the propagation speed of the transmitted electromagnetic wave. Mixing $r(t)$ with the reference chirp signal $s^*(t)$ gives

$$r_1(t) = \delta \exp \left[ j \left( 2\pi f_0 t - \gamma^2 \right) \right] \exp \left( j 2\gamma t \right).$$

(43)

Note that $r_1(t)$ is a sinusoidal signal with frequency $2\gamma$. Let $\tau_{\text{max}}$ and $\tau_{\text{min}}$ correspond to the maximum and minimum, respectively, of the round-trip time delays between the scatterers of a radar target and the radar. We assume that $(\tau_{\text{max}} - \tau_{\text{min}}) \ll T_0$. Then, for $-T_0/2 + \tau_{\text{max}} \leq t \leq T_0/2 + \tau_{\text{min}}$, the scatterers of the radar target at different frequencies correspond to different frequencies of the mixed signal.

The target range signature is the RCS’s of the target scatterers as a function of the range. The most important radar target features, which will be of interest herein, are the locations and RCS’s of the dominant scatterers. The less dominant scatterers of the radar target along with radar clutter and thermal noise also contribute to the mixed signal described above and are considered to be noise. This noise is generally nonwhite and non-Gaussian. Since we are most interested in extracting dominant scatterers of a radar target, we shall use below only the RELAX step of the DPE. In other words, we do not estimate any noise model. It is worth noting that for a prespecified value $K$ of $K$, the DPE estimates of the signal parameters do not depend, in any way, on the estimation of a noise model. However, if $K$ is determined from the data, then the noise modeling is important in that it affects the estimated number of sinusoids. Most likely, if we set $M$ to zero, then the estimated value of $K$ is larger than otherwise. Hence, the effect of estimating no noise model will be an inflated number of estimated scattering centers, which is a situation we can cope with as the estimated dominant scatterers are robust to the overestimation of $K$ (see below).

We first consider a numerical example to illustrate the performance of the RELAX algorithm. Consider a case where the magnitude of the true range signature of a radar target is as shown in Fig. 5(a). The simulated target has four dominant scatterers and 32 weak scatterers. The RCS’s of the four dominant scatterers are 1, 2, 2, and 2. The 32 weak scatterers are uniformly distributed between relative range 0 and 0.5. The RCS’s of the weak scatterers are statistically independent complex Gaussian random variables with zero-mean and variance 1/32. The length of the data sequence is $N = 64$, and no noise is added to the data sequence. Fig. 5(b) shows the modulus of the range signature estimate obtained with the FFT method. We note that two of the dominant scatterers are too closely spaced to be resolved by the FFT method. We now estimate the number of dominant scatterers with the GAIC criterion by setting the AR noise model order to zero. The estimate for $K$ obtained is $K = 4$. Using this $K = 4$ with the RELAX algorithm, we obtain the range signature estimate whose modulus is shown in Fig. 5(c). We note that using $K = 4$ and the estimate whose modulus is shown in Fig. 5(c). The simulated target has four dominant scatterers and 32 weak scatterers. The RCS’s of the four dominant scatterers are 1, 2, 2, and 2. The 32 weak scatterers are uniformly distributed between relative range 0 and 0.5. The RCS’s of the weak scatterers are statistically independent complex Gaussian random variables with zero-mean and variance 1/32. The length of the data sequence is $N = 64$, and no noise is added to the data sequence. Fig. 5(b) shows the modulus of the range signature estimate obtained with the FFT method. We note that two of the dominant scatterers are too closely spaced to be resolved by the FFT method. We now estimate the number of dominant scatterers with the GAIC criterion by setting the AR noise model order to zero. The estimate for $K$ obtained is $K = 4$. Using this $K = 4$ with the RELAX algorithm, we obtain the range signature estimate whose modulus is shown in Fig. 5(c). We note that using $K = 4$ and the estimate whose modulus is shown in Fig. 5(c).
Fig. 5. Magnitude of true and estimated range signatures for a radar target with $N = 64$: (a) True range signature; (b) estimated (solid line) with periodogram; (c) estimated (solid lines) with RELAX, $K = 4$, and $M = 0$; (d) estimated (solid lines) with RELAX, $K = 10$, and $M = 0$.

$K = 10$ with the RELAX algorithm yields similar estimates for the four dominant scatterers. Thus, the RELAX algorithm for estimating the dominant scatterers of a radar target is not sensitive to the estimated or desired number of sinusoids $K$ for estimating the dominant scatterers of a radar target.

We have also applied the RELAX algorithm to the experimental data measured by a ground-to-air radar and have obtained similar results [19].

2) Target Feature Extraction from SAR Phase Histories with RELAX: In SAR or inverse SAR (ISAR) imaging [20], [21], the radar also usually transmits linear FM chirp pulses. Upon receiving each pulse returned by an object being imaged, the radar demodulates the pulse by mixing the pulse with a reference chirp signal as described above for high-resolution range signature estimation. As a result, the scattering centers of the object at different ranges correspond to different frequencies of the mixed signal. Since either the radar or the object is moving or rotating, the pulses received at different angles between the radar and the object are used to form a synthetic aperture. After polar-to-Cartesian frequency interpolation, the scattering centers of the object at the same range but different cross-ranges correspond to different (Doppler) frequencies over the synthetic aperture. The 2-D SAR data sequence is in the form of (3), where for each scatterer, one frequency corresponds to the range, as in (43), the other frequency corresponds to the cross-range or the Doppler frequency, and the complex amplitude is proportional to the RCS of the scatterer. The conventional SAR imaging method is the FFT method, which is known for its poor resolution and high sidelobes.

The most important radar target features to be extracted from SAR phase histories are once more the locations (now in 2-D) and the RCS's of the dominant scatterers of a radar target, which are also the features of interest herein.
We now present an example showing the performance of the RELAX algorithm for SAR imaging and target feature extraction. Fig. 6(a) shows the true SAR image of a numerically simulated tank. Fig. 6(b) shows the SAR image obtained with the 2-D FFT method when \( N = \hat{N} = 32 \), and the data is corrupted by additive zero-mean white Gaussian noise with variance 20. Fig. 6(c) shows the SAR image obtained from the noisy data by using 2-D FFT with a circularly symmetric Kaiser window with shape parameter 3. The GAIC estimate for \( \hat{K} \) obtained from the noisy data is \( \hat{K} = K = 9 \). Fig. 6(d) shows the SAR image obtained from the noisy data by using the RELAX algorithm when the AR noise model order is set to zero and \( \hat{K} = 9 \) is used. We note that the RELAX algorithm can resolve closely spaced dominant scatterers that cannot be resolved by the 2-D FFT (with and without windowing) methods.

V. CONCLUSIONS

We have presented a decoupled parameter estimation (DPE) algorithm for estimating sinusoidal and AR noise parameters from both 1-D and 2-D data sequences. In the first step of the DPE, we use a relaxation (RELAX) algorithm that requires simple FFT’s to obtain the estimates of the sinusoidal parameters. We have described in detail how the RELAX algorithm can be used to extract radar target features from
both 1-D and 2-D data sequences. In the second step of the DPE algorithm, a linear least-squares approach is used to estimate the AR noise parameters. The DPE algorithm is both conceptually and computationally simple. We have shown with numerical examples that the algorithm not only provides excellent estimation performance under the model assumptions but also is robust to the violation of those assumptions. Moreover, we have shown that under the model assumptions, the estimates obtained with the DPE algorithm are asymptotically statistically efficient.

APPENDIX A

CRAMÉR-RAO BOUND FOR THE SINUSOIDAL PARAMETERS

When the additive noise is a zero-mean white Gaussian random process, the derivation of the CR bound for the sinusoidal parameters can be found in [22] and [23]. We sketch below the derivation of the CR bound for the sinusoidal parameters when the additive noise is a zero-mean colored Gaussian random process with an unknown covariance matrix $Q$. We only consider the CR bound for complex data. The derivation of the CR bound for real-valued data is similar [23].

The observed data vector can be written as

$$y = \Omega \alpha + e,$$

where for 1-D sinusoids, $y$ is defined in (10) and $\Omega$ is defined in (16), for 2-D sinusoids, $y$ is defined in (32), $\Omega = [\omega(f_1) \otimes \omega(f_1) \omega(f_2) \otimes \omega(f_2) \cdots \omega(f_K) \otimes \omega(f_K)]$, and $e$ is the noise vector defined similarly to $y$. Let $Q = E\{ee^H\}$, where $E\{\cdot\}$ denotes the expectation, be the covariance matrix of $e$.

The variables of the likelihood function of $y$ are the unknown elements of $Q$, the sinusoidal frequencies, and the real and imaginary parts of the sinusoidal amplitudes. The extended Bangs' formula for the $ij$th element of the Fisher information matrix has the form [24]

$$\{\text{FIM}\}_{ij} = \text{tr}\left( Q^{-1} Q_i Q^{-1} Q_j \right) + 2 \text{Re}\left( \alpha^H \Omega^H \right) \left( Q^{-1} (\Omega \alpha) \right)_i \left( Q^{-1} (\Omega \alpha) \right)_j$$

(45)

where $X_i$ denotes the gradient of $X$ with respect to the $i$th unknown of the likelihood function, $\text{tr}(X)$ denotes the trace of $X$, and $\text{Re}(X)$ denotes the real part of $X$. Note that $\text{FIM}$ is block diagonal since $Q$ does not depend on the parameters in $(\Omega \alpha)$, and $(\Omega \alpha)$ does not depend on the elements of $Q$. This simple observation implies at once that the CR bound of the sinusoidal parameters for the colored noise case can be obtained from the CR bound for the white noise case by replacing $\Omega$ with $Q^{-1/2} \Omega$ and $\Omega_i$ with $Q^{-1/2} \Omega_i$.

Let

$$G = 2 \Omega^H Q^{-1} \Omega$$

(46)

and let

$$\Delta = 2 \Omega^H Q^{-1} DP$$

(47)

where for 1-D sinusoids, $P = \text{diag}(\alpha)$ and for the $k$th column, $k = 1, 2, \cdots, K$ of $D$ is $\partial \omega(f_k) / \partial f_k$. For 2-D sinusoids, $P = \text{diag}(\alpha) \otimes I_2$, where $I_2$ is the identity matrix of dimension 2, and the $(2k-1)$th and $(2k)$th columns, $k = 1, 2, \cdots, K$, of $D$ are $\partial \omega(f_k) \otimes \omega(f_k) / \partial f_k$ and $\partial \omega(f_k) \otimes \omega(f_k) / \partial f_k$, respectively. Let

$$\Gamma = 2 \text{Re}[P^H D^H Q^{-1} DP].$$

(48)

Define

$$\eta = \begin{bmatrix} \text{Re}(\alpha) & \text{Im}(\alpha) & f^T \end{bmatrix}^T$$

(49)

where $\text{Im}(X)$ denotes the imaginary part of $X$. For 1-D sinusoids, $f = [f_1 \ f_2 \ \cdots \ f_K]^T$, and for 2-D sinusoids, $f = [f_1 \ f_1 \ f_2 \ f_2 \ \cdots \ f_K \ f_K]^T$. Then, similar to the results found in [22] and [23], the CR bound for $\eta$ is

$$\text{CRB}(\eta) = \begin{bmatrix} \text{Re}(G) & -\text{Im}(G) & \text{Re}(\Delta) \\ \text{Im}(G) & \text{Re}(G) & \text{Im}(\Delta) \\ \text{Re}^T(\Delta) & \text{Im}^T(\Delta) & \Gamma \end{bmatrix}^{-1}.$$  

(50)

We remark that when $Q = \sigma^2 I$, the results above are identical to those found in [22] and [23].

Note that when the additive noise is an AR or ARMA random process, the CR bound for the sinusoidal parameters can be easily calculated by using the above expression since the corresponding $Q$ can be calculated from the AR or ARMA noise parameters.

APPENDIX B

JPE FOR REAL DATA SEQUENCES

The JPE approaches presented in [2] for complex data sequences can easily be modified for real data sequences. One of the approaches is modified as follows. Let $\bar{y}_\text{JPE} = \text{Re}(\bar{y})$.

$$G = 2 \begin{bmatrix} \cos(2\pi f_1 M) & -\sin(2\pi f_1 M) & \cos(2\pi f_3 M) & -\sin(2\pi f_3 M) & \cdots \\ \cos(2\pi f_1 (M+1)) & -\sin(2\pi f_1 (M+1)) & \cos(2\pi f_3 (M+1)) & -\sin(2\pi f_3 (M+1)) & \cdots \\ \vdots \\ \cos(2\pi f_1 (N-1)) & -\sin(2\pi f_1 (N-1)) & \cos(2\pi f_3 (N-1)) & -\sin(2\pi f_3 (N-1)) & \cdots \end{bmatrix}$$

(52)

$$\mu = \begin{bmatrix} \text{Re}(\mu_1) \\ \text{Im}(\mu_1) \\ \text{Re}(\mu_3) \\ \text{Im}(\mu_3) \\ \cdots \\ \text{Re}(\mu_{2K-1}) \\ \text{Im}(\mu_{2K-1}) \end{bmatrix}$$
The matrix $H$ is defined as:

$$
H = \begin{bmatrix}
'y_1' & 'y_2' & \cdots & 'y_{N-1}'
\end{bmatrix}^T.
$$

Let $G$ be an $(N - M) \times K$ matrix as shown in (52), which appears at the bottom of the previous page. It can be shown that maximizing the approximate or conditional likelihood function in [2] for real data sequences is equivalent to minimizing

$$
D_1(f_k, \alpha_k, \beta_m) = \|y_{\text{jpe}} + Ha + G\mu\|^2
$$

(53)

where $\mu$ is shown in the equation at the bottom of the previous page, with $\mu_{m-1} = \mu_{2k-1}$, $k = 1, 2, \ldots, K/2$, and

$$
\mu_k = -\alpha_k \left(1 + \sum_{m=1}^{M} a_m e^{-2\pi i f_m/m}\right).
$$

(54)

Then, the frequency estimates can be found by minimizing the following cost function [2]:

$$
D_2(f_k) = y_{\text{jpe}}^T [P_H^{-1} - P_H G (G^T P_H G)^{-1} G^T P_H] y_{\text{jpe}}.
$$

(55)

We also have

$$
\hat{\mu} = -(G^T P_H G)^{-1} G^T P_H y_{\text{jpe}},
$$

(56)

and

$$
\hat{\alpha} = -(H^T H)^{-1} H^T (y_{\text{jpe}} + G\hat{\mu}).
$$

(57)

Hence, $\alpha$ is found from (54) by replacing $\{\mu_k\}$, $\{f_k\}$, and $\{a_m\}$ there with $\{\hat{\mu}_k\}$, $\{\hat{f}_k\}$ and $\{\hat{a}_m\}$, respectively.

REFERENCES


