Subspace Approach for Two-Dimensional Parameter Estimation of Multiple Damped Sinusoids

Frankie K.W. Chan, H.C. So and Weize Sun

ckwf@hkexperts.com, hcso@ee.cityu.edu.hk, weizesun2@student.cityu.edu.hk

Abstract

In this paper, we tackle the two-dimensional (2-D) parameter estimation problem for a sum of $K \ge 2$ real/complex damped sinusoids in additive white Gaussian noise. According to the rank-K property of the 2-D noise-free data matrix, the damping factor and frequency information is contained in the Kdominant left and right singular vectors. Using the sinusoidal linear prediction property of these vectors, the frequencies and damping factors of the first dimension are first estimated. For each frequency of the first dimension, the corresponding parameter in the second dimension is then obtained to achieve autopairing. Computer simulations are included to compare the proposed approach with several conventional 2-D estimators in terms of mean square error performance and computational complexity.

Index Terms

dominant singular vector, linear prediction, weighted least squares, two-dimensional parameter estimation, spectral analysis

I. INTRODUCTION

The problem of parameter estimation for $K \ge 1$ 2-D noisy sinusoids has received a great deal of attention. It is because in many applications such as source localization [1]–[2], radar imaging [3], vibrational analysis of circularly shaped objects [4], nuclear magnetic resonance (NMR) spectroscopy [5], wireless communication channel estimation [6], the corresponding signals can be well described by the 2-D sinusoidal model.

The authors are with Department of Electronic Engineering, City University of Hong Kong.

1

of its computational attractiveness when using fast Fourier transform, it suffers from poor resolution in resolving closely-spaced frequencies and high-sidelobe effects [7]-[8]. In order to achieve higher resolution, the parametric approach, which assumes that the signal satisfies a generating model with known functional form, is a standard choice. Well known 2-D parametric solutions include maximum likelihood (ML) method [9] and subspace-based estimators such as multiple signal classification (MUSIC) [2], [3], [5], matrix enhancement and matrix pencil (MEMP) [10] and estimation of signal parameters via rotational invariance techniques (ESPRIT) [4], [6], [11], [12], [13]. In the presence of additive white Gaussian noise, the ML scheme [9], which corresponds to a multi-dimensional peak search, can produce optimum estimation performance, that is, its mean square error (MSE) attains Cramér-Rao lower bound (CRLB). Comparing with the ML estimator, the subspace methodology whose underlying principle is to separate the received data into signal and noise subspaces via eigenvalue decomposition (EVD) or singular value decomposition (SVD), is more computationally efficient at the expense of suboptimality. The MUSIC algorithm [3] requires to find the K peaks in a 2-D cost function constructed from the noise eigenvectors. By constructing a Hankel-block-Hankel matrix whose size is larger than that of the data matrix, the MEMP method [10] decomposes the 2-D estimation problem into two one-dimensional problems related to each dimension, where generalized EVD and 2-D frequency pairing are needed. The

2-D Fourier transform is the direct nonparametric approach to address 2-D spectral estimation. In spite

ESPRIT algorithm [11] is similar to [10] in the sense that the Hankel-block-Hankel matrix is exploited but it provides auto-pairing of the 2-D frequencies by making use of joint diagonalization and can deal with damped sinusoids. While [4], which addresses X-texture modes, that is, real-valued 2-D sinusoids with damping only in one dimension, can be considered as a modification to [11] by applying partial forward-backward averaging.

In this paper, we contribute to accurate and fast 2-D parameter estimation of multiple complex/real damped sinusoids based on the subspace methodology. We refer our approach to as principal-singular-vector utilization for modal analysis (PUMA), meaning that the principal singular vectors of the data matrix are effectively exploited in the estimation process. This work is a follow-up of [14] where the PUMA algorithm for a single damped/undamped real/complex tone or K = 1 is devised and analyzed. The key ideas in [14] are to make use of the rank-one property of the corresponding 2-D noise-free data matrix and find the damping factor as well as frequency parameters for each dimension from the left and right principal singular vectors in a separable manner. In this paper, we extend [14] to the general case of $K \ge 2$ which is not a straightforward task. In particular, the damping factors and frequencies in each dimension are not directly related to the K principal singular vectors and 2-D parameter pairing needs

to be addressed.

The rest of the paper is organized as follows. The algorithm development for multiple damped cisoids is provided in Section II. According to the rank-*K* property of the 2-D noise-free data matrix, the corresponding left and right dominant singular vectors are characterized by the damping factors and frequencies in the first and second dimension, respectively. Making use of the dominant singular vectors, the parameters of interest at one dimension will first be accurately estimated according to an iterative procedure which utilizes the sinusoidal linear prediction (LP) property and weighted least squares (WLS). The damping factors and frequencies in another dimension are solved via another similar iterative algorithm such that pairing of the 2-D parameters is automatically achieved. Estimation for real tones is addressed in Section III. In Section VI, simulation results are included to evaluate the performance of the PUMA approach by comparing with the ML [9] and ESPRIT algorithms [4], [11] as well as CRLB. Finally, conclusions are drawn in Section V. A list of mathematical symbols that are used in the paper is given in Table 1.

Symbol	Meaning
†	pseudoinverse
vec	vectorization operator
\otimes	Kronecker product
\odot	Hadamard product
0	Khatri-Rao product
\mathbf{I}_i	$i \times i$ identity matrix
$0_{i imes j}$	$i \times j$ zero matrix
ã	noise-free value of a
â	estimate of a
$\left[\mathbf{a} ight]_{i}$	<i>i</i> th element of \mathbf{a}
$\left[\mathbf{A} ight]_{i,j}$	(i,j) entry of ${f A}$
$\operatorname{diag}(\mathbf{a})$	diagonal matrix with vector a as main diagonal
$\mathrm{blkdiag}(\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_n)$	block diagonal matrix with its diagonal elements are square matrices of $\mathbf{A}_1, \mathbf{A}_2, \cdots \mathbf{A}_n$
$\text{Toeplitz}(\mathbf{a}, \mathbf{b}^T)$	Toeplitz matrix with first column \mathbf{a} and first row \mathbf{b}^T
$\operatorname{vec}(\mathbf{A})$	$\begin{bmatrix} \mathbf{a}_1^T & \mathbf{a}_2^T & \cdots & \mathbf{a}_K^T \end{bmatrix}^T \text{ where } \mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_K \end{bmatrix}$

TABLE I List of Symbols

II. ESTIMATION FOR COMPLEX SINUSOIDS

In this Section, we first devise the PUMA algorithm for estimating the parameters of multiple damped cisoids in additive noise. The signal model is:

$$r_{m,n} = s_{m,n} + q_{m,n}, \quad m = 1, 2, \cdots, M, \quad n = 1, 2, \cdots, N$$
 (1)

$$s_{m,n} = \sum_{k=1}^{K} \gamma_k \alpha_k^m \beta_k^n \exp\left\{j(\mu_k m + \nu_k n)\right\}$$
(2)

is the noise-free signal. The γ_k is the complex amplitude, $\mu_k \in (-\pi, \pi)$ and $\alpha_k > 0$ are the frequency and damping factor in the first dimension while $\nu_k \in (-\pi, \pi)$ and $\beta_k > 0$ are the corresponding parameters in the second dimension, of the *k*th cisoid. The number of damped cisoids, namely, $K \ge 2$, is assumed known. Here we consider that the frequencies are distinct for at least one dimension. Without loss of generality, we assume that $M \ge N > K$ and all frequencies in the first dimension are not identical, that is, $\mu_k \ne \mu_l$, $k \ne l$. The additive noises $\{q_{m,n}\}$ are uncorrelated zero-mean complex white Gaussian processes with unknown variances σ_q^2 . The task is to find $\{\mu_k\}$, $\{\nu_k\}$, $\{\alpha_k\}$, $\{\beta_k\}$ and $\{\gamma_k\}$, from the *MN* samples of $\{r_{m,n}\}$.

We first express (1)–(2) in matrix form as:

$$\mathbf{R} = \mathbf{S} + \mathbf{Q} \tag{3}$$

where $[\mathbf{R}]_{m,n} = r_{m,n}$, $[\mathbf{S}]_{m,n} = s_{m,n}$ and $[\mathbf{Q}]_{m,n} = q_{m,n}$. Considering $s_{m,n}$ as a sum of K components of $(\alpha_k \exp\{j\mu_k\})^m \gamma_k (\beta_k \exp\{j\nu_k\})^n$, $k = 1, 2, \cdots, K$, it is easy to see that S can be factorized as:

$$\mathbf{S} = \mathbf{G} \boldsymbol{\Gamma} \mathbf{H}^T \tag{4}$$

where $\mathbf{G} = \begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \cdots & \mathbf{g}_K \end{bmatrix}$, $\mathbf{\Gamma} = \operatorname{diag}(\gamma_1, \gamma_2, \cdots, \gamma_K)$, $\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \cdots & \mathbf{h}_K \end{bmatrix}$, $\mathbf{g}_k = \begin{bmatrix} a_k & a_k^2 & \cdots & a_k^M \end{bmatrix}^T$, $\mathbf{h}_k = \begin{bmatrix} b_k & b_k^2 & \cdots & b_k^N \end{bmatrix}^T$, $a_k = \alpha_k \exp\{j\mu_k\}$ and $b_k = \beta_k \exp\{j\nu_k\}$. On the other hand, \mathbf{R} can be decomposed using SVD as

$$\mathbf{R} = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^H \tag{5}$$

where $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$ is the square diagonal matrix containing singular values of \mathbf{R} with $\lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_N \ge 0$ while $\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_N] \in \mathbb{C}^{M \times N}$ and $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_N] \in \mathbb{C}^{N \times N}$ are orthonormal matrices whose columns are the corresponding left and right singular vectors, respectively. From the decomposition in (4), it is obvious that $\operatorname{rank}(\mathbf{S}) = K$ and thus the best rank-*K* approximation of \mathbf{R} according to (5), denoted by $\hat{\mathbf{S}}$, is

$$\hat{\mathbf{S}} = \mathbf{U}_s \mathbf{\Lambda}_s \mathbf{V}_s^H \tag{6}$$

where

DRAFT

5

Here, the matrix with subscript "s" contains the dominant components of the original one, corresponds to the signal subspace components. Comparing (4) and (6), it is clear that G, Γ and H correspond to U_s, Λ_s and V^{*}_s, respectively. To perform parameter estimation, we first let matrix with $\tilde{\cdot}$ be the noise-free versions of the matrices in (6). As G and \tilde{U}_s span the same column space, we have

$$\tilde{\mathbf{U}}_s = \mathbf{G} \boldsymbol{\Omega}_G \tag{7}$$

where $\Omega_G \in \mathbb{C}^{K \times K}$ is an unknown matrix. That is, each column of $\tilde{\mathbf{U}}_s$, namely, $\tilde{\mathbf{u}}_k$, $k = 1, 2, \dots, K$, is a sum of K damped cisoids such that the frequencies and damping factors in $\{\tilde{\mathbf{u}}_k\}$ are identical but they have different amplitudes, which corresponds to a multi-channel spectral estimation problem [15]. More precisely, each element of $\tilde{\mathbf{u}}_k$ has the form of $[\tilde{\mathbf{u}}_k]_m = \sum_{k=1}^K \zeta_k \alpha_k^m \exp\{j\mu_k m\}$ where $\{\zeta_k\}$ are parameters not of interest, and thus $[\tilde{\mathbf{u}}_k]_m$ can be expressed as a linear combination of its previous K samples [7]. As a result, we have the following LP property:

$$\sum_{i=0}^{K} c_i [\tilde{\mathbf{u}}_k]_{m-i} = 0, \quad c_0 = 1, \quad k = 1, 2, \cdots, K, \quad m = K+1, \cdots, M$$
(8)

where $\{c_i\}$ are called the LP coefficients. The frequencies ω_k and damping factors α_k are related to the following polynomial:

$$\sum_{i=0}^{K} c_i z^{K-i} = 0 \tag{9}$$

whose roots are $z = a_k$, k = 1, 2, ..., K [16]. When noise is absent, expressing (8), for a specific $\tilde{\mathbf{u}}_k$, as matrix form yields

$$\begin{bmatrix} \mathbf{D}_k & -\mathbf{f}_k \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{c} \end{bmatrix} = \mathbf{0}_{(M-K)\times 1}, \quad k = 1, 2, \cdots, K$$
(10)

where

$$\mathbf{D}_{k} = \begin{vmatrix} [\tilde{\mathbf{u}}_{k}]_{K} & [\tilde{\mathbf{u}}_{k}]_{K-1} & \cdots & [\tilde{\mathbf{u}}_{k}]_{1} \\ [\tilde{\mathbf{u}}_{k}]_{K+1} & [\tilde{\mathbf{u}}_{k}]_{K} & \cdots & [\tilde{\mathbf{u}}_{k}]_{2} \\ \vdots & \vdots & \ddots & \cdots \end{vmatrix}$$
(11)

$$\tilde{\mathbf{u}}_{k}]_{M-1} \quad [\tilde{\mathbf{u}}_{k}]_{M-2} \quad \cdots \quad [\tilde{\mathbf{u}}_{k}]_{M-K} \end{bmatrix}$$
$$\mathbf{c} = \begin{bmatrix} c_{1} \quad c_{2} \quad \cdots \quad c_{K} \end{bmatrix}^{T} \tag{12}$$

$$\mathbf{f}_{k} = -\left[[\tilde{\mathbf{u}}_{k}]_{K+1} \quad [\tilde{\mathbf{u}}_{k}]_{K+2} \quad \cdots \quad [\tilde{\mathbf{u}}_{k}]_{M} \right]^{T}, \quad k = 1, 2, \cdots, K.$$
(13)

Grouping the left hand side of (10) for $k = 1, 2, \dots, K$, yields

$$\mathbf{e} = \begin{bmatrix} \mathbf{e}_1^T & \mathbf{e}_2^T \cdots & \mathbf{e}_K^T \end{bmatrix}^T = \mathbf{D}\mathbf{c} - \mathbf{f} = \mathbf{0}_{(M-K)K \times 1}$$
(14)

where

$$\mathbf{e}_k = \mathbf{D}_k \mathbf{c} - \mathbf{f}_k = \mathbf{A} \tilde{\mathbf{u}}_k = \mathbf{0}_{(M-K)\times 1},\tag{15}$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1^T & \mathbf{D}_2^T & \cdots & \mathbf{D}_K^T \end{bmatrix}^T$$
(16)

$$\mathbf{f} = \begin{bmatrix} \mathbf{f}_1^T & \mathbf{f}_2^T & \cdots & \mathbf{f}_K^T \end{bmatrix}^T, \quad k = 1, 2, \cdots, K$$
(17)

$$\mathbf{A} = \text{Toeplitz} \left(\begin{bmatrix} c_K & \mathbf{0}_{1 \times (M-K-1)} \end{bmatrix}^T, \begin{bmatrix} c_K & c_{K-1} & \cdots & c_1 & 1 & \mathbf{0}_{1 \times (M-K-1)} \end{bmatrix} \right).$$
(18)

In the presence of noise, $\tilde{\mathbf{u}}_k$ is replaced by \mathbf{u}_k and (14) is invalid. To estimate c, we formulate a WLS cost function

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \mathbf{e}^{H} \mathbf{W} \mathbf{e} = \left(\mathbf{D}^{H} \mathbf{W} \mathbf{D} \right)^{-1} \mathbf{D}^{H} \mathbf{W} \mathbf{f}$$
(19)

where \hat{c} denotes the estimate of c. Furthermore, W is a symmetric weighting matrix and the optimal choice according to Gauss-Markov theorem is [17]:

$$\mathbf{W} = \sigma_q^2 \left[\mathbb{E} \left\{ \mathbf{e} \mathbf{e}^H \right\} \right]^{-1} \tag{20}$$

To calculate \mathbf{W} , we first note that \mathbf{e}_k in (15) implies

$$\mathbf{A}\begin{bmatrix} \tilde{\mathbf{u}}_1 & \tilde{\mathbf{u}}_2 & \cdots & \tilde{\mathbf{u}}_K \end{bmatrix} = \mathbf{A}\tilde{\mathbf{U}}_s = \mathbf{0}_{(M-K)\times K}.$$
(21)

Moreover, when noise is present, we can express (14) as

$$\mathbf{e} = \operatorname{vec}\left(\mathbf{A}\mathbf{U}_{s}\right) = \operatorname{vec}\left(\mathbf{A}\tilde{\mathbf{U}}_{s} + \mathbf{A}\Delta\mathbf{U}_{s}\right)$$
(22)

where $\Delta U_s = U_s - \tilde{U}_s$ is the perturbation caused by noise. Using (22), we can express (20) as

$$\mathbf{W} = \sigma_q^2 \left[\mathbb{E} \left\{ \operatorname{vec}(\mathbf{A}\mathbf{U}_s) \operatorname{vec}(\mathbf{A}\mathbf{U}_s)^H \right\} \right]^{-1}$$
(23)

To compute (23), we first apply [18] to obtain

$$\Delta \mathbf{U}_s = \tilde{\mathbf{U}}_s \boldsymbol{\Theta} + \mathbf{Q} \tilde{\mathbf{V}} \tilde{\boldsymbol{\Lambda}}^{-1} - \tilde{\mathbf{U}}_s \tilde{\mathbf{U}}_s^H \mathbf{Q} \tilde{\mathbf{V}} \tilde{\boldsymbol{\Lambda}}^{-1}$$
(24)

where $\Theta = \mathbf{Z} \odot \left(\tilde{\mathbf{U}}_s^H \mathbf{Q} \tilde{\mathbf{V}} \tilde{\mathbf{\Lambda}} + \tilde{\mathbf{\Lambda}} \tilde{\mathbf{V}}^H \mathbf{Q}^H \tilde{\mathbf{U}}_s \right)$ with $[\mathbf{Z}]_{m,n}$ being equal to $1/(\tilde{\lambda}_n - \tilde{\lambda}_m)$ and zero otherwise. Using (21) and (24), we have

$$\mathbf{A}\mathbf{U}_s = \mathbf{A}\mathbf{\Delta}\mathbf{U}_s = \mathbf{A}\mathbf{Q}\tilde{\mathbf{V}}\tilde{\mathbf{\Lambda}}^{-1}$$
(25)

Vectorizing both sides of (25) yields

$$\operatorname{vec}\left(\mathbf{A}\mathbf{U}_{s}\right) = \left(\tilde{\mathbf{\Lambda}}^{-1}\tilde{\mathbf{V}}^{T}\otimes\mathbf{A}\right)\operatorname{vec}(\mathbf{Q}).$$
(26)

As $\mathbb{E}\{\mathbf{Q}\mathbf{Q}^H\} = \sigma_q^2 \mathbf{I}_{MN}$, (26) implies that ¹

$$\mathbb{E}\left\{\operatorname{vec}\left(\mathbf{A}\mathbf{U}_{s}\right)\operatorname{vec}\left(\mathbf{A}\mathbf{U}_{s}\right)^{H}\right\} = \sigma_{q}^{2}\tilde{\boldsymbol{\Lambda}}^{-2}\otimes\mathbf{A}\mathbf{A}^{H}.$$
(27)

Substituting (27) into (23) and approximating $\tilde{\lambda}_k$ by λ_k , we have ²

$$\mathbf{W} \approx \operatorname{diag}(\lambda_1^2, \lambda_2^2, \cdots, \lambda_K^2) \otimes (\mathbf{A}\mathbf{A}^H)^{-1}.$$
(28)

Since W is block diagonal, (19) can be written as

$$\hat{\mathbf{c}} = \left(\sum_{k=1}^{K} \lambda_k^2 \mathbf{D}_k^H (\mathbf{A}\mathbf{A}^H)^{-1} \mathbf{D}_k\right)^{-1} \left(\sum_{k=1}^{K} \lambda_k^2 \mathbf{D}_k^H (\mathbf{A}\mathbf{A}^H)^{-1} \mathbf{f}_k\right).$$
(29)

As (28) depends on the unknown c, we follow [19] to estimate c in an iterative manner and the estimation procedure is:

- (i) Set $\mathbf{A}\mathbf{A}^{H} = \mathbf{I}_{M-K}$ which corresponds to the LS estimator for (26)
- (ii) Calculate \hat{c} using (29)
- (iii) Compute an updated version of W with $c = \hat{c}$ using (28)
- (iv) Repeat Steps (ii)-(iii) until a stopping criterion is reached

Substituting $\hat{\mathbf{c}} = \mathbf{c}$ in (9) and solving for the roots, denoted by \hat{a}_k , $k = 1, 2, \dots, K$, the frequency and damping factor estimates in the first dimension are:

$$\hat{\mu}_k = \angle(\hat{a}_k) \tag{30}$$

$$\hat{\alpha}_k = |\hat{a}_k|. \tag{31}$$

Basically, we can follow (28)–(31) to solve for $\{\hat{\nu}_k\}$ and $\{\hat{\beta}_k\}$ but a matching step is required to determine the correct pairs of (a_k, b_k) . In order to achieve 2-D parameter pairing in an automatic manner, we employ another estimation procedure for the parameters in the second dimension as follows.

From (3)–(4), we have

$$\mathbf{R} \approx \hat{\mathbf{G}} \boldsymbol{\mathcal{H}}^T \tag{32}$$

 ${}^{1}(\mathbf{A}\otimes\mathbf{B}) (\mathbf{C}\otimes\mathbf{D}) = \mathbf{A}\mathbf{C}\otimes\mathbf{B}\mathbf{D}$ ${}^{2}(\mathbf{A}\otimes\mathbf{B})^{-1} = \mathbf{A}^{-1}\otimes\mathbf{B}^{-1}$

where $\hat{\mathbf{G}}$ is the estimate of \mathbf{G} which is constructed according to $a_k = \hat{a}_k$ and

$$\boldsymbol{\mathcal{H}}^{T} = \boldsymbol{\Gamma} \mathbf{H}^{T} = \begin{bmatrix} \boldsymbol{h}_{1} & \boldsymbol{h}_{2} & \cdots & \boldsymbol{h}_{K} \end{bmatrix}^{T}$$
(33)

$$\boldsymbol{h}_k = \gamma_k \mathbf{h}_k. \tag{34}$$

From (32), the least squares (LS) estimate of \mathcal{H} is

$$\hat{\boldsymbol{\mathcal{H}}} = \mathbf{R}^T (\hat{\mathbf{G}}^\dagger)^T.$$
(35)

Noting that the elements of h_k possess the same LP property as in h_k , we extract \hat{h}_k from $\hat{\mathcal{H}}$ to construct the equations

$$\hat{\boldsymbol{h}}_{k,u} b_k \approx \hat{\boldsymbol{h}}_{k,l} \tag{36}$$

where $h_{k,u}$ and $h_{k,l}$ are h_k but without the last and first element, respectively. Considering sufficiently small error conditions such that $\hat{a}_k \rightarrow \hat{a}$, we have $\hat{\mathbf{G}} \rightarrow \mathbf{G}$. In such conditions, \mathbf{R} will be independent of $\hat{\mathbf{G}}$ and thus the disturbances among each vector of \hat{h}_k can be assumed independent and identically distributed (IID). Following [19], the WLS estimate of b_k is computed as:

$$\hat{b}_{k} = \arg\min_{b_{k}} (\hat{\boldsymbol{h}}_{k,u} b_{k} - \hat{\boldsymbol{h}}_{k,l})^{H} \boldsymbol{\Psi}_{k} (\hat{\boldsymbol{h}}_{k,u} b_{k} - \hat{\boldsymbol{h}}_{k,l}) = \frac{\hat{\boldsymbol{h}}_{k,u}^{H} \boldsymbol{\Psi}_{k} \hat{\boldsymbol{h}}_{k,l}}{\hat{\boldsymbol{h}}_{k,u}^{H} \boldsymbol{\Psi}_{k} \hat{\boldsymbol{h}}_{k,u}}, \quad k = 1, 2, \cdots, K$$
(37)

where the optimum weighting matrix Ψ_k has the form:

$$\Psi_{k} = \left[\mathbb{E}\left\{ (\hat{\boldsymbol{h}}_{k,u}b_{k} - \hat{\boldsymbol{h}}_{k,l})(\hat{\boldsymbol{h}}_{k,u}b_{k} - \hat{\boldsymbol{h}}_{k,l})^{H} \right\} \right]^{-1} = \left(\mathbf{B}_{k}\mathbf{B}_{k}^{H} \right)^{-1}$$
(38)

$$\mathbf{B}_{k} = \text{Toeplitz} \left(\begin{bmatrix} -b_{k} & \mathbf{0}_{1 \times (N-2)} \end{bmatrix}^{T}, \begin{bmatrix} -b_{k} & 1 & \mathbf{0}_{1 \times (N-2)} \end{bmatrix} \right).$$
(39)

Similar to the iterative estimation of c in the first dimension, we start with $\Psi_k = \mathbf{I}_{N-1}$ in the iterations between (37) and (38) to obtain \hat{b}_k , $k = 1, 2, \dots, K$. Finally, the frequencies ν_k and damping factors β_k are estimated as

$$\hat{\nu}_k = \angle(\hat{b}_k) \tag{40}$$

$$\hat{\beta}_k = |\hat{b}_k|, \qquad k = 1, 2, \cdots, K.$$
 (41)

In doing so, a_k and b_k are automatically paired up. It is worthy to point out that we can first estimate $\{b_k\}$ according to (7)–(31) and then base on (37) to find $\{a_k\}$.

The breakdown of the required number of multiplications of the PUMA algorithm is provided in Table II where ℓ_1 and ℓ_2 represent the iteration numbers in (26) and (35), respectively. Considering $M \ge N$ which is much larger than K, ℓ_1 and ℓ_2 , the complexity of the proposed estimator is $\mathcal{O}(M^3)$. On the other hand, the main computational requirement of the ESPRIT method [11] is to perform SVD of a matrix with dimension $PQ \times (M - P + 1)(N - Q + 1)$, where P and Q are typically chosen as M/2 and N/2, respectively, corresponding to $\mathcal{O}(M^3N^3)$ multiplications. While the ML algorithm [9] needs to compute a LS solution involving a matrix with dimension (M - K)N + (M - K - 1)(N - 1) by K(K - 1), which results in $\mathcal{O}(M^3N^3)$ multiplications. As a result, the PUMA estimator is superior to the ESPRIT and ML methods in terms of computationally attractiveness.

Operation	Multiplication	Step
SVD of R	$4M^2N + 8MN^2 + 9N^3$	1
$\mathbf{A}\mathbf{A}^{H}$	$(M-K)^2M$	2
$\mathbf{Z} = (\mathbf{A}\mathbf{A}^H)^{-1} \begin{bmatrix} \mathbf{D}_k & \mathbf{f}_k \end{bmatrix}$	$\frac{(M-K)(2M^2+6M-2-K^2-3K-KM)}{6}$	3
$\mathbf{Z}_k = \lambda_k^2 \mathbf{Z}$	(M-K)(M+1)	4
$[\mathbf{V}_k \;\; \mathbf{v}_k] = \mathbf{D}_k^H \mathbf{Z}_k$	K(M+1)(M-K)	5
$(2+3+4) \times K$	$\frac{(M-K)(8KM^2 - K^2M + 12KM - K^3 + 3K^2 + 4K)}{6}$	6
$\left(\sum_{k=1}^{K} \mathbf{V}_k ight)^{-1} \left(\sum_{k=1}^{K} \mathbf{v}_k ight)$	$\frac{K(K^2+3K-1)}{3}$	Ø
$(6+2) \times \ell$ iterations	$\ell_1 \times \frac{8KM^3 - 9K^2M^2 + 12KM^2 - 9K^2M + 4KM + K^4 + K^3 + 8K^2 - 10K}{6}$	8
$\mathbf{T}_1 = \mathbf{G}^H \mathbf{G}$	K^2M	9
$\mathbf{T}_2 = \mathbf{G}^H \mathbf{R}$	KMN	10
$\mathbf{T}_1^{-1}\mathbf{T}_2$	$\frac{K+9K^2+2K^3+9KN+3K^2N}{6}$	(11)
9+ <u>10</u> + <u>11</u>	$\frac{6KMN + 3K^2N + 9KN + 6K^2M + 2K^3 + 9K^2 + K}{6}$	12
$\boldsymbol{\Psi}_k^{-1} = \mathbf{B}_k \mathbf{B}_k^H$	$N(N-1)^2$	(13)
$oldsymbol{\Psi}_k \left[\hat{oldsymbol{h}}_{k,u} \;\; \hat{oldsymbol{h}}_{k,l} ight]$	$\frac{-3-N+3N^2+N^3}{3}$	(14)
\hat{b}_k	2N - 1	(15)
$(13 + 14 + 15) \times K \times \ell_2$ iterations	$\ell_2 \times K \times \frac{4N^3 - 3N^2 + 8N - 6}{3}$	16

TABLE II

COMPLEXITY BREAKDOWN OF PROPOSED METHOD

With the use of the frequency and damping factor estimates, the estimate of $\gamma = \begin{bmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_K \end{bmatrix}^T$, is computed using a LS fit:

$$\hat{\boldsymbol{\gamma}} = \left(\hat{\mathbf{H}} \circ \hat{\mathbf{G}}\right)^{\dagger} \operatorname{vec}(\mathbf{R})$$
(42)

where $\hat{\mathbf{H}}$ is the estimate of \mathbf{H} which is constructed using $b_k = \hat{b}_k$ and \circ denotes the Khatri-Rao product.

III. ESTIMATION FOR REAL SINUSOIDS

In this Section, we extend our study to real sinusoids and the signal model is now given as:

$$r_{m,n} = s_{m,n} + q_{m,n} (43)$$

$$s_{m,n} = \sum_{k=1}^{K} \gamma_k \alpha_k^m \beta_k^n \cos(\mu_k m + \phi_k) \cos(\nu_k n + \theta_k).$$
(44)

Now $\{\gamma_k\}$ are real and positive, the admissible ranges for $\{\mu_k\}$ and $\{\nu_k\}$ are reduced to $(0, \pi)$, and $\{\phi_k\}$ and $\{\theta_k\}$ are the additional initial phase parameters. While $q_{m,n}$ is a real zero-mean white Gaussian process with unknown variance σ_q^2 . The parameters of interest are μ_k , ν_k , α_k , β_k , ϕ_k , θ_k and γ_k , k = $1, 2, \dots, K$. Expressing (44) in matrix form, it is shown that **S** can be factorized as in (4) but \mathbf{g}_k and \mathbf{h}_k are modified to

$$\mathbf{g}_{k} = \begin{bmatrix} \alpha_{k} \cos\left(\mu_{k} + \phi_{k}\right) & \alpha_{k}^{2} \cos\left(2\mu_{k} + \phi_{k}\right) & \cdots & \alpha_{k}^{M} \cos\left(M\mu_{k} + \phi_{k}\right) \end{bmatrix}^{T}$$
(45)

$$\mathbf{h}_{k} = \begin{bmatrix} \beta_{k} \cos\left(\nu_{k} + \theta_{k}\right) & \beta_{k}^{2} \cos\left(2\nu_{k} + \theta_{k}\right) & \cdots & \beta_{k}^{N} \cos\left(N\nu_{k} + \theta_{k}\right) \end{bmatrix}^{T}$$
(46)

which are characterized by the frequencies, damping factors and phases. Following the development in Section II, the LP property in $\tilde{\mathbf{u}}_k$, $k = 1, 2, \dots, K$, is now:

$$\sum_{i=0}^{2K} c_i [\tilde{\mathbf{u}}_k]_{m-i} = 0, \quad c_0 = 1, \quad m = 2K + 1, \cdots, M.$$
(47)

The roots of the following polynomial:

where

and

$$\sum_{i=0}^{2K} c_i z^{2K-i} = 0 \tag{48}$$

are $z = \alpha_k \exp\{\pm j\mu_k\}$, $k = 1, 2, \dots, K$, from which the frequencies and damping factors in the first dimension can be straightforwardly obtained. Similar to (19)–(28), the conceptual WLS estimate of c is obtained as:

$$\hat{\mathbf{c}} = \left(\sum_{k=1}^{K} \lambda_k^2 \mathbf{D}_k^H (\mathbf{A}\mathbf{A}^H)^{-1} \mathbf{D}_k\right)^{-1} \left(\sum_{k=1}^{K} \lambda_k^2 \mathbf{D}_k^H (\mathbf{A}\mathbf{A}^H)^{-1} \mathbf{f}_k\right)$$
(49)

February 1, 2012

where **D**, **f**, **c** and **W** are defined in (16), (17), (12) and (28), respectively while \mathbf{D}_k , \mathbf{f}_k and **A** are

 $\mathbf{D}_{k} = \begin{bmatrix} [\mathbf{u}_{k}]_{2K} & [\mathbf{u}_{k}]_{2K-1} & \cdots & [\mathbf{u}_{k}]_{1} \\ [\mathbf{u}_{k}]_{2K+1} & [\mathbf{u}_{k}]_{2K} & \cdots & [\mathbf{u}_{k}]_{2} \\ \vdots & \vdots & \ddots & \cdots \\ [\mathbf{u}_{k}]_{M-1} & [\mathbf{u}_{k}]_{M-2} & \cdots & [\mathbf{u}_{k}]_{M-2K} \end{bmatrix}$ (50)

$$\mathbf{f}_{k} = -\left[[\mathbf{u}_{k}]_{2K+1} \quad [\mathbf{u}_{k}]_{2K+2} \quad \cdots \quad [\mathbf{u}_{k}]_{M} \right]^{T}, \quad k = 1, 2, \cdots, K$$

$$(51)$$

$$\mathbf{A} = \text{Toeplitz} \left(\begin{bmatrix} c_{2K} & \mathbf{0}_{1 \times (M-2K-1)} \end{bmatrix}^T, \begin{bmatrix} c_{2K} & c_{2K-1} & \cdots & c_1 & 1 & \mathbf{0}_{1 \times (M-2K-1)} \end{bmatrix} \right).$$
(52)

Initializing W with the identity matrix, we iterate between (49) and (52) to get \hat{c} , from which 2K roots with conjugate relationship are determined using (48). We then employ those phases are within $(0, \pi)$ to find μ_k and α_k according to (30) and (31). The next step is to find $\{\phi_k\}$. With the use of $\cos(u+v) = \cos(u)\cos(v) - \sin(u)\sin(v)$, we write

$$\mathbf{G} = \mathbf{G}_{\mu\alpha}\mathbf{G}_{\phi} \tag{53}$$

where

modified as

$$\mathbf{G}_{\mu\alpha} = \begin{bmatrix} \boldsymbol{c}_1 & -\boldsymbol{s}_1 & \boldsymbol{c}_2 & -\boldsymbol{s}_2 & \cdots & \boldsymbol{c}_K & -\boldsymbol{s}_K \end{bmatrix}$$
(54)

$$\boldsymbol{c}_{k} = \begin{bmatrix} \alpha_{k} \cos\left(\mu_{k}\right) & \alpha_{k}^{2} \cos\left(2\mu_{k}\right) & \cdots & \alpha_{k}^{M} \cos\left(M\mu_{k}\right) \end{bmatrix}^{T}$$
(55)

$$\boldsymbol{s}_{k} = \begin{bmatrix} \alpha_{k} \sin\left(\mu_{k}\right) & \alpha_{k}^{2} \sin\left(2\mu_{k}\right) & \cdots & \alpha_{k}^{M} \sin\left(M\mu_{k}\right) \end{bmatrix}^{T}$$
(56)

$$\mathbf{G}_{\phi} = \text{blkdiag}\left(\boldsymbol{\xi}_{1}, \boldsymbol{\xi}_{2}, \cdots, \boldsymbol{\xi}_{K}\right)$$
(57)

and

$$\boldsymbol{\xi}_{k} = \begin{bmatrix} \cos(\phi_{k}) & \sin(\phi_{k}) \end{bmatrix}^{T}.$$
(58)

After substituting $\hat{\alpha}_k$ and $\hat{\mu}_k$ into $\mathbf{G}_{\mu\alpha}$, we obtin

$$\hat{\boldsymbol{\mathscr{H}}}^{T} \triangleq \hat{\mathbf{G}}_{\mu\alpha}^{\dagger} \mathbf{R} \approx \hat{\mathbf{G}}_{\phi} \hat{\boldsymbol{\Gamma}} \hat{\mathbf{H}}^{T} = \begin{bmatrix} \hat{\mathfrak{h}}_{1} & \hat{\mathfrak{h}}_{2} & \cdots & \hat{\mathfrak{h}}_{2K} \end{bmatrix}^{T}.$$
(59)

Relating $\hat{\mathscr{H}}$ and $\hat{\mathbf{H}}$ in (59) yields

$$\hat{\mathfrak{h}}_{2k-1} \approx \hat{\mathbf{h}}\hat{\gamma}_k \cos(\hat{\phi}_k), \quad \hat{\mathfrak{h}}_{2k} \approx \hat{\mathbf{h}}\hat{\gamma}_k \sin(\hat{\phi}_k), \quad k = 1, 2, \cdots, K$$
 (60)

which implies

$$\hat{\phi}_k = \tan^{-1} \left(\hat{\mathfrak{h}}_{2k-1}^{\dagger} \hat{\mathfrak{h}}_{2k} \right), \quad k = 1, 2, \cdots, K.$$
 (61)

Constructing $\hat{\mathbf{G}}_{\phi}$ from (61) and using the equation $h_{k,i+2} = 2\beta_k \cos(\nu_k)h_{k,i+1} - \beta_k^2 h_{k,i}$, we can follow (35)–(41) to estimate $\hat{\beta}_k$ and $\hat{\nu}_k$ from $\hat{\mathcal{H}}$. After that, the estimation of $\{\theta_k\}$ is achieved by considering **H** which is similar to (53)–(61). Finally, $\{\gamma_k\}$ are estimated using LS which is similar to (42).

DRAFT

IV. SIMULATION RESULTS

Computer simulations have been conducted to evaluate the parameter estimation performance of the PUMA algorithm for multiple damped 2-D sinusoids in the presence of white Gaussian noise. The stopping criterion of the PUMA algorithm is a fixed number of iterations. We use 3 iterations as no significant improvement is observed for more iterations. The average MSE is employed for the performance measure. Apart from CRLB, comparison is made with ESPRIT [4], [11] and ML [9] algorithms. In particular, [11] is employed for cisoids while [4] which applies partial forward-backward averaging is used for the real-valued signals. On the other hand, the ML estimator is realized by the Newton's method where parameter initialization is obtained using [11] and [4] in the complex and real cases, respectively. In the study, synthetic signals are employed although real data can help understanding the actual performance in real-life applications. The signal power is defined as $\sigma_s^2 = (\sum_{m=1}^M \sum_{n=1}^N |s_{m,n}|^2)/(MN)$ and $q_{m,n}$ is scaled to produce different signal-to-noise ratio (SNR) conditions with SNR = σ_s^2/σ_q^2 . The number of 2-D sinusoids is assigned as K = 2 while M = N = 20 correspond to the data size. All results provided are averages of 1000 independent runs based on a computer with Intel Dual Core i7 2.66 GHz processors and 3GB RAM.

In the first test, we study the complex-valued case and the signal parameters are $\gamma_1=2, \ \gamma_2=$ $3\exp(0.15j), \alpha_1 = 0.95, \alpha_2 = 0.96, \beta_1 = 0.99, \beta_2 = 0.95, \mu_1 = 0.1\pi, \mu_2 = 0.36\pi, \nu_1 = 0.24\pi$ and $\nu_2 = 0.36\pi$. The average MSEs for frequencies, damping factors and amplitudes, versus SNR are plotted in Figures 1 to 3, respectively. It is seen that the MSEs of the proposed and ML methods attain the corresponding CRLBs at SNR ≥ -2 dB and at SNR ≥ -6 dB in all three figures, respectively. On the other hand, the ESPRIT method is not optimum in the whole SNR range. Furthermore, we observe that starting with estimation of $\{a_k\}$ and $\{b_k\}$ in the PUMA scheme give similar final results although the former has relatively smaller threshold SNRs. The average CPU computation times of the ESPRIT, ML and PUMA algorithms for a single trial are measured as 2.98×10^{-2} s, 3.148×10^{-1} s and 2.00×10^{-3} s, respectively, which agree with the complexity analysis in Section 2. The more challenging scenarios of closely-spaced and identical frequencies are now examined. Figure 4 shows the average MSEs for frequencies with $\mu_1 = 0.3\pi$, $\mu_2 = 0.36\pi$, $\nu_1 = 0.32\pi$ and $\nu_2 = 0.36\pi$ while the other parameter settings remain unchanged. The results are similar to those of Figure 1 except that estimating $\{a_k\}$ in the first stage is better than starting with $\{b_k\}$ because the frequency separation in the latter is closer. Figure 5 plots the frequency estimation performance when there are identical frequencies in the second dimension. The parameter settings are equal to those of Figure 4 except $\nu_1 = 0.36\pi$. The observations are similar

to those of Figure 1. Note that in this case, we are not able to start with estimating $\{b_k\}$ due to the restriction in utilizing (17). Moreover, as the results for damping factors and amplitudes are similar, they are not included here. From the simulations, it can be observed that though the proposed algorithm is computationally more attractive, it fails to start with the dimension having identical frequency component. To avoid this situation, estimation of the frequency separations of both dimensions has to be performed first. We then begin the algorithm at the dimension with maximum frequency separation.

The real-valued scenario is investigated in the second test and we consider that the 2-D tones are undamped in one dimension corresponding to the X-texture modes [4], which arise naturally in vibrational analysis of circularly shaped objects. The PUMA algorithm in Section 3 is modified accordingly [14] to achieve better estimation performance, and without loss of generality, we first estimate $\{\alpha_k\}$, $\{\mu_k\}$ and $\{\phi_k\}$. The parameter settings are identical to those in the first test except that $\beta_1 = \beta_2 = 1$ and the additional phase values are assigned as $\phi_1 = 0$, $\phi_2 = 0.15$, $\theta_1 = 0.5$ and $\theta_2 = 1$. The average MSEs for frequencies, damping factors, amplitudes and phases, versus SNR are plotted in Figures 6 to 7, respectively. Again, we see that the PUMA algorithm is superior to the ESPRIT method in terms of estimation accuracy. While the MSEs of the proposed and ML methods attain the corresponding CRLBs at SNR ≥ 6 dB and at SNR ≥ 8 dB in all four figures, indicating that the former has a slightly better threshold performance. In addition, the average computation times of the ESPRIT, ML and proposed estimators for a single trial are measured as 1.89×10^{-2} s, 9.257×10^{-1} s and 2.2×10^{-3} s, respectively. These further demonstrate the superiority of the proposed algorithm over the two conventional algorithms in terms of computational complexity and/or accuracy.

V. CONCLUSION

A novel subspace-based parameter estimation approach for $K \ge 2$ 2-D multiple damped real/complex sinusoids in additive white Gaussian noise is devised. With the use of singular value decomposition and weighted least squares techniques, the damping factors and frequencies are obtained by iterative estimation between the two dimensions. The parameter estimates in both dimensions are automatically paired. Computer simulations show that the PUMA approach is computationally attractive and able to attain optimal mean square error performance when the signal-to-noise ratio is sufficiently high.

REFERENCES

S. O. Al-Jazzar, D. C. McLernon, and M. A. Smadi, "SVD-based joint azimuth/elevation estimation with automatic pairing," *Signal Processing*, vol. 90, no. 5, pp. 1669–1675, May 2010.

- [2] M. L. Bencheikh, Y. Wang, and H. He, "Polynomial root finding technique for joint DOA DOD estimation in bistatic mimo radar," *Signal Processing*, vol. 90, no. 9, pp. 2723–2730, Sep. 2010.
- [3] J. W. Odendaal, E. Barnard, and W. I. Pistorius, "Two-dimensional superresolution radar imaging using the MUSIC algorithm," *IEEE Transactions on Antennas and Propagation*, vol. 42, no. 10, pp. 1386–1391, October 1994.
- [4] J. Axmon, M. Hansson, and L. Sornmo, "Partial forward-backward averaging for enhanced frequency estimation of real X-texture modes," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2550–2562, July 2005.
- [5] Y. Li, J. Razavilar, and K. J. R. Liu, "A high-resolution technique for multidimensional NMR spectroscopy," *IEEE Transactions on Biomedical Engineering*, vol. 45, pp. 78–86, January 1998.
- [6] A.-J. van der Veen, M. C. Vanderveen, and A. Paulraj, "Joint angle and delay estimation using shift-invariance techniques," *IEEE Transactions on Signal Processing*, vol. 46, no. 2, pp. 405–418, February 1998.
- [7] P. Stoica and R. Moses, Spectral Analysis of Signals. Upper Saddle River, NJ: Prentice Hall, 2005.
- [8] N. Sandgren, P. Stoica, and F. J. Frigo, "Area-selective signal parameter estimation for two-dimensional MR spectroscopy data," *Journal of Magnetic Resonance*, vol. 183, pp. 50–59, August 2006.
- M. P. Clark and L. L. Scharf, "Two-dimensional modal analysis based on maximum likelihood," *IEEE Transactions on Signal Processing*, vol. 42, no. 6, pp. 1443–1452, June 1994.
- [10] Y. Hua, "Estimating two-dimensional frequencies by matrix enhancement and matrix pencil," *IEEE Transactions on Signal Processing*, vol. 40, no. 9, pp. 2267–2280, September 1992.
- [11] S. Rouquette and M. Najim, "Estimation of frequencies and damping factors by two-dimensional ESPRIT type methods," *IEEE Transactions on Signal Processing*, vol. 49, no. 1, pp. 237–245, January 2001.
- [12] Y. Berthoumieu, M. E. Ansari, B. Aksasse, M. Donias, and M. Najim, "A 2-D robust high-resolution frequency estimation approach," *Signal Processing*, vol. 85, no. 6, pp. 1165–1188, Jun. 2005.
- [13] Y. Zhou, D.-Z. Feng, and J.-Q. Liu, "A novel algorithm for two-dimensional frequency estimation," *Signal Processing*, vol. 87, no. 1, pp. 1165–1188, Jan. 2007.
- [14] H. C. So, F. K. W. Chan, W. H. Lau, and C. F. Chan, "An efficient approach for two-dimensional parameter estimation of a single-tone," *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 1999–2009, April 2010.
- [15] F. K. W. Chan, H. C. So, W. H. Lau, and C. F. Chan, "Efficient approach for sinusoidal frequency estimation of gapped data," *IEEE Signal Processing Letters*, vol. 17, pp. 611–614, June 2010.
- [16] Y. T. Chan, J. M. M. Lavoie, and J. B. Plant, "A parameter estimation approach to estimation of frequencies of sinusoids," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. 29, no. 2, pp. 214–219, April 1981.
- [17] S. M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory. Prentice Hall, 1993.
- [18] J. Liu, X. Liu, and X. Ma, "First-order perturbation analysis of singular vectors in singular value decomposition," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3044–3049, July 2008.
- [19] H. C. So and K. W. Chan, "A generalized weighted linear predictor frequency estimation approach for a complex sinusoid," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1304–1315, April 2006.



Fig. 1. Average mean square frequency error versus SNR for complex tones



Fig. 2. Average mean square damping factor error versus SNR for complex tones



Fig. 3. Average mean amplitude error versus SNR for complex tones



Fig. 4. Average mean frequency error versus SNR for closely-spaced frequencies



Fig. 5. Average mean frequency error versus SNR for identical frequencies



Fig. 6. Average mean square frequency error versus SNR for real tones



Fig. 7. Average mean square damping factor error versus SNR for real tones