Accurate and Computationally Efficient Tensor-based Subspace Approach for Multi-Dimensional Harmonic Retrieval

Weize Sun and H.C. So

Abstract

In this paper, parameter estimation for R-dimensional (R-D) sinusoids with $R > 2$ in additive white Gaussian noise is addressed. With the use of tensor algebra and principal-singular-vector utilization for modal analysis, the sinusoidal parameters at one dimension are first accurately estimated according to an iterative procedure which utilizes the linear prediction property and weighted least squares. The damping factors and frequencies in the remaining dimensions are then solved such that pairing of the R-D parameters is automatically achieved. Algorithm modification for a single R-D tone is made and it is proved that the frequency estimates are asymptotically unbiased while their variances approach Cramér-Rao lower bound at sufficiently high signal-to-noise ratio conditions. Computer simulations are also included to compare the proposed approach with conventional R-D harmonic retrieval schemes in terms of mean square error performance and computational complexity.

Index Terms

multi-dimensional spectral analysis, harmonic retrieval, parameter estimation, tensor algebra, subspace method

I. Introduction

The problem of harmonic retrieval (HR) has been an important topic in science and engineering because many real-world signals can be well described by the sinusoidal model. Although one-dimensional (1-D)

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

spectral analysis is the most common, multi-dimensional HR in fact has numerous applications such as wireless communication channel estimation [1]–[2], nuclear magnetic resonance (NMR) spectroscopy [3] and multiple-input multiple-output (MIMO) radar imaging [4].

As in 1-D spectral estimation, R-dimensional (R-D) HR where $R \geq 2$, can be achieved by means of either nonparametric or parametric approaches [5]. Analogously, the crucial step in the R-D scenarios is to find the damping factor and frequency parameters which are nonlinear functions of the observed data. The nonparametric HR methods, including the periodogram and correlogram, are based directly on the Fourier transform. Although no assumptions are made about the observed sequence, the resolution, or ability to resolve closely spaced frequencies using the nonparametric approach is fundamentally limited by the length of the data available. On the other hand, the parametric approach, which assumes that the signal satisfies a generating model with known functional form, can attain a higher resolution. Conventional R-D parametric HR techniques include maximum likelihood (ML) [6], iterative quadratic ML (IQML) [7] and subspace [1]–[3], [8]–[14] approaches such as estimation of signal parameters via rotational invariance technique (ESPRIT) and multi-dimensional folding (MDF). However, the ML-based methods are only feasible for 2-D HR due to their extremely high computational requirement. On the other hand, the subspace methodology involves a smaller complexity, and its underlying principle is to separate the data into signal and noise subspaces, usually via eigenvalue decomposition (EVD) of the sample covariance matrix or singular value decomposition (SVD) of the raw data matrix, and the parameters of interest are then extracted from the corresponding eigenvectors, singular vectors, eigenvalues or singular values.

For 2-D cases, the data perfectly align with the matrix representation and thus EVD or SVD can be straightforwardly applied. Nevertheless, even signals with $R > 2$ are stored in matrices by means of stacking operations in most of the existing subspace-based approaches. In fact, it is more natural to store and manipulate higher-dimensional signals using tensors [15]–[16]. Although tensor methods such as higher-order SVD (HOSVD) or parallel factor (PARAFAC) have been very popular in some scientific areas, particularly chemometrics and psychometrics [17], they are relatively new in the signal processing discipline. It is worth mentioning that the links of multidimensional HR with tensor algebra and HOSVD were first recognized in [18] and [19], respectively. Other important pioneer works on tensor-based HR include [4], [10]–[11], [20]. To the best of our knowledge, the state-of-the-art subspace-based R-D HR methods cannot provide estimation performance attaining the Cramér-Rao lower bound (CRLB) [21]. It is also desirable that the computational requirement of the HR schemes can be made lower even in processing multi-dimensional data. In this paper, we contribute to the development of an accurate and computationally efficient tensor-based R-D HR approach.

The rest of the paper is organized as follows. In Section II, we first introduce the necessary notation and then formulate the problem of R-D HR in the presence of white Gaussian noise. By exploiting the subspace methodology of principal-singular-vector utilization for modal analysis (PUMA) and tensor algebra, the proposed estimation algorithm is developed in Section III. Note that we have already applied PUMA for a single 2-D tone in [14] where the key ideas are to exploit the rank-one property of the corresponding 2-D noise-free data matrix and to find the damping factor as well as frequency parameters for each dimension from the left and right principal singular vectors in a separable manner. With the use of tensor algebra, 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). By employing the estimated parameters and observed tensor data, we are able to decompose the tensor into damped single-tone sequences which correspond to the damping factors and frequencies in the remaining dimensions. They are then estimated according the iterative procedure and the pairing of the R-D parameters is automatically achieved. As tensor algebra and PUMA are exploited, the devised R-D HR approach is referred to as tensor PUMA (TPUMA) algorithm. We also devise a computationally attractive scheme to determine the appropriate dimension for processing in the first stage. In Section IV, the proposed approach is modified for a single R-D undamped tone and it is proved that the frequency estimates are asymptotically unbiased and their variances attain CRLB when signal-to-noise ratio (SNR) is sufficiently high. In Section V, extensive simulation results are included to evaluate the performance of the PUMA approach by comparing with the ESPRIT [8]–[11] and MDF [12]–[13] algorithms as well as CRLB. Finally, conclusions are drawn in Section VI.

II. Notation and Problem Formulation

The notation used in this paper is first introduced as follows. Scalars, vectors, matrices and tensors are denoted by italic, bold lower-case, bold upper-case and bold calligraphic symbols, respectively. The rth unfolding of $\boldsymbol{\mathcal{A}}$ is written as $[\boldsymbol{\mathcal{A}}]_{(r)} \in \mathbb{C}^{M_r \times (M_1 M_2 \cdots M_{r-1} M_{r+1} \cdots M_R)}$ where the order of the columns is chosen according to [15]. The r-mode product of tensor $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_R}$ and matrix $\mathbf{U} \in \mathbb{C}^{N_r \times M_r}$ along the rth dimension is expressed as $\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{A}} \times_r \mathbf{U} \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_{r-1} \times N_r \times M_{r+1} \times \cdots \times M_R}$ where $[\boldsymbol{\mathcal{B}}]_{(r)} = \mathbf{U}[\boldsymbol{\mathcal{A}}]_{(r)}$. Note that $\boldsymbol{\mathcal{B}}$ can be interpreted as multiplying all r-mode vectors of $\boldsymbol{\mathcal{A}}$ by $\mathbf{U}$. The outer product of two tensors $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_P}$ and $\boldsymbol{\mathcal{B}} \in \mathbb{C}^{N_1 \times N_2 \times \cdots \times N_Q}$ is a $(P+Q)$-D tensor of the form $\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{A}} \circ \boldsymbol{\mathcal{B}} \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_P \times N_1 \times N_2 \times \cdots \times N_Q}$ where $c_{m_1,m_2,\ldots,m_R,n_1,n_2,\ldots,n_R} = a_{m_1,m_2,\ldots,m_R} \cdot b_{n_1,n_2,\ldots,n_R}$ and $\circ$ is the outer product operator. The symbol $\sqcup$ represents the concatenation operator where $\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{A}}_1 \sqcup_r \boldsymbol{\mathcal{A}}_2$ is obtained by stacking $\boldsymbol{\mathcal{A}}_2 \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_{r-1} \times L_2 \times M_{r+1} \times \cdots \times M_R}$ to the end of

$\boldsymbol{\mathcal{A}}_1 \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_{r-1} \times L_1 \times M_{r+1} \times \cdots \times M_R}$ along the rth dimension. The HOSVD of $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_R}$ is $\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{S}} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \cdots \times_R \mathbf{U}_R$ where $\boldsymbol{\mathcal{S}} \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_R}$ is the core ordered tensor satisfying the all-orthogonality and $\mathbf{U}_r \in \mathbb{C}^{M_r \times M_r}, r = 1, 2, \cdots, R$, are the unitary matrices of the r-mode singular vectors [15]. Finally, the remaining mathematical symbols used in the paper are listed in Table 1.

Symbol	**Meaning**
$\angle(a)$	angle of a
$\lvert a \rvert$	magnitude of a
^{T}	transpose
^{H}	conjugate transpose
*	complex conjugate
$^{-1}$	inverse
†	pseudoinverse
vec	vectorization operator
$\otimes$	Kronecker product
$\odot$	Khatri-Rao product
$\mathbb{E}$	expectation operator
$\mathbb{C}^{M_1 \times M_2 \times \cdots \times M_R}$	set of $M_1 \times M_2 \times \cdots \times M_R$ complex tensors
$\mathbf{I}_i$	$i \times i$ identity matrix
$\mathbf{0}_{i \times j}$	$i \times j$ zero matrix
$\widetilde{\mathbf{A}}$	variable of $\mathbf{A}$
$\bar{\mathbf{A}}$	noise-free value of $\mathbf{A}$
$\widehat{\mathbf{A}}$	estimate of $\mathbf{A}$
$[\mathbf{a}]_i$	ith element of $\mathbf{a}$
$[\mathbf{A}]_{i,j}$	(i, j) entry of $\mathbf{A}$
$a_{m_1, m_2, \ldots, m_R}$	$(m_1, m_2, \ldots, m_R)$ entry of a R-D tensor $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_R}$
$\mathrm{diag}(\mathbf{a})$	diagonal matrix with vector $\mathbf{a}$ as main diagonal
$\mathrm{rank}(\mathbf{A})$	rank of $\mathbf{A}$
$\mathrm{span}(\mathbf{A})$	span of $\mathbf{A}$
$\mathrm{Tr}(\mathbf{A})$	trace of the square matrix $\mathbf{A}$
$\mathrm{blkdiag}(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n)$	block diagonal matrix with submatrices $\mathbf{A}_1, \mathbf{A}_2, \cdots \mathbf{A}_n$
$\mathrm{Toeplitz}(\mathbf{a}, \mathbf{b}^T)$	Toeplitz matrix with first column $\mathbf{a}$ and first row $\mathbf{b}^T$

Table 1: Mathematical symbols

The observed R-D sinusoidal signal is modeled as:

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{Q}} \tag{1}$$

where

$$x_{m_1,m_2,\ldots,m_R} = \sum_{f=1}^{F} \gamma_f \prod_{r=1}^{R} \alpha_{r,f}^{m_r} e^{j\omega_{r,f} m_r}, \quad m_r = 1,2,\ldots,M_r, \quad r = 1,2,\ldots,R, \quad f = 1,2,\ldots,F \quad (2)$$

The $\boldsymbol{\mathcal{Y}} \in \mathbb{C}^{M_1 \times M_2 \times \cdots \times M_R}$ is the tensorial structured data set with length M_r along the rth dimension. The tensor $\boldsymbol{\mathcal{X}}$ is the signal component where γ_f, $\omega_{r,f} \in (-\pi, \pi)$ and $\alpha_{r,f} \in (0,1]$, $f = 1,2,\cdots,F$, represent the unknown complex amplitudes, frequencies and damping factors in the rth dimension, and F is the number of R-D frequencies which is assumed known $a\ priori$. It is assumed that at least one of $\{M_r\}$ is larger than F with distinct frequencies. For ease of presentation but without loss of generality, we let $M_R > F$ with all $\{\omega_{R,f}\}$ being distinct. On the other hand, the elements in $\boldsymbol{\mathcal{Q}}$ are zero-mean complex white Gaussian noises with unknown variances σ^2. The task is to find $\{\omega_{r,f}\}$ and $\{\alpha_{r,f}\}$ from the $M = \prod_{r=1}^{R} M_r$ entries of $\boldsymbol{\mathcal{Y}}$. Note that $\{\gamma_f\}$ can be easily estimated by applying a least squares (LS) fit [14] on (1) after the frequencies and damping factors have been determined.

III. Proposed Estimator

Employing the connection between the PARAFAC and HR in [4], the noise-free tensor $\boldsymbol{\mathcal{X}}$ can be expressed as

$$\boldsymbol{\mathcal{X}} = \sum_{f=1}^{F} \gamma_f \mathbf{g}_{1,f} \circ \mathbf{g}_{2,f} \cdots \circ \mathbf{g}_{R,f} \quad (3)$$

where

$$\mathbf{g}_{r,f} = \begin{bmatrix} a_{r,f} & a_{r,f}^2 & \cdots & a_{r,f}^{M_r} \end{bmatrix}^T, \quad a_{r,f} = \alpha_{r,f} e^{j\omega_{r,f}} \quad (4)$$

According to (3), $\boldsymbol{\mathcal{X}}$ can also be written as

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{C}}_{\gamma_1,\gamma_2,\ldots,\gamma_F} \times_1 \mathbf{G}_1 \times_2 \mathbf{G}_2 \cdots \times_R \mathbf{G}_R \quad (5)$$

where

$$\mathbf{G}_r = \begin{bmatrix} \mathbf{g}_{r,1} & \mathbf{g}_{r,2} & \cdots & \mathbf{g}_{r,F} \end{bmatrix} \quad (6)$$

and $\boldsymbol{\mathcal{C}}_{\gamma_1,\gamma_2,\ldots,\gamma_F} \in \mathbb{C}^{F \times F \times \cdots \times F}$ is a R-D tensor whose $(f,f,\cdots,f)$ entry equals γ_f and zero otherwise. Based on (3) and (5), the rth unfolding of $\boldsymbol{\mathcal{X}}$ is

$$[\boldsymbol{\mathcal{X}}]_{(r)} = \mathbf{G}_r \boldsymbol{\Sigma} (\mathbf{G}_{r+1} \odot \mathbf{G}_{r+2} \odot \cdots \mathbf{G}_R \odot \mathbf{G}_1 \odot \mathbf{G}_2 \odot \cdots \mathbf{G}_{r-1})^T \quad (7)$$

where

$$\boldsymbol{\Sigma} = \mathrm{diag}\left(\begin{bmatrix} \gamma_1, \gamma_2, \ldots, \gamma_F \end{bmatrix}\right) \quad (8)$$

On the other hand, computing the HOSVD of $\boldsymbol{\mathcal{Y}}$ yields

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{S}} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \cdots \times_R \mathbf{U}_R \quad (9)$$

where each $\mathbf{U}_r$, $r = 1, 2, \cdots, R$, is related to the SVD of $[\boldsymbol{\mathcal{Y}}]_{(r)}$ as

$$[\boldsymbol{\mathcal{Y}}]_{(r)} = \mathbf{U}_r \boldsymbol{\Lambda}_r \mathbf{V}_r^H \tag{10}$$

where $\boldsymbol{\Lambda}_r = \mathrm{diag}\left(\left[\lambda_{r,1}, \lambda_{r,2}, \cdots, \lambda_{r,N_r}\right]\right)$, $N_r = \min\{M_r, M/M_r\}$, is the square diagonal matrix containing singular values with $\lambda_{r,1} \geq \lambda_{r,2} \geq \cdots \geq \lambda_{r,N_r} \geq 0$ while $\mathbf{U}_r = \begin{bmatrix} \mathbf{u}_{r,1} & \mathbf{u}_{r,2} & \cdots & \mathbf{u}_{r,N_r} \end{bmatrix} \in \mathbb{C}^{M_r \times N_r}$ and $\mathbf{V}_r \in \mathbb{C}^{(M/M_r) \times N_r}$ are orthonormal matrices whose columns are the corresponding left and right singular vectors of $[\boldsymbol{\mathcal{Y}}]_{(r)}$, respectively.

As the rank of tensor $\boldsymbol{\mathcal{X}}$ is F, we use the observed $\boldsymbol{\mathcal{Y}}$ to obtain a low multi-linear rank approximation or the truncated HOSVD of $\boldsymbol{\mathcal{X}}$, denoted by $\widehat{\boldsymbol{\mathcal{X}}}$ [15]:

$$\widehat{\boldsymbol{\mathcal{X}}} = \boldsymbol{\mathcal{S}}^{[s]} \times_1 \mathbf{U}_1^{[s]} \times_2 \mathbf{U}_2^{[s]} \cdots \times_R \mathbf{U}_R^{[s]} \tag{11}$$

where $\mathbf{U}_r^{[s]} = \begin{bmatrix} \mathbf{u}_{r,1} & \mathbf{u}_{r,2} & \cdots & \mathbf{u}_{r,l_r} \end{bmatrix} \in \mathbb{C}^{M_r \times l_r}$, $r = 1, 2, \cdots, R$, $l_r = \min\{F, M_r, M/M_r\}$, and $\boldsymbol{\mathcal{S}}^{[s]}$ has only the first l_r elements of $\boldsymbol{\mathcal{S}}$ in the rth dimension. Although (11) is not the best rank-F approximation, optimum estimation performance can be achieved, which is demonstrated in Sections IV and V. Noted that the best rank-$(l_1, l_2, \cdots, l_R)$ approximation [22] of $\boldsymbol{\mathcal{X}}$ can be employed but its computational complexity is much higher.

Comparing (5) and (11), we deduce that when $M_r \geq F$, $\mathrm{span}(\bar{\mathbf{U}}_r^{[s]}) \subseteq \mathrm{span}(\mathbf{G}_r)$, and thus we can write

$$\bar{\mathbf{U}}_r^{[s]} = \mathbf{G}_r \boldsymbol{\Omega}_r, \quad r = 1, 2, \cdots, R \tag{12}$$

where each $\boldsymbol{\Omega}_r \in \mathbb{C}^{l_r \times F}$ is an unknown nonsingular matrix. Assuming that all frequencies are distinct in the rth dimension which implies $\mathrm{rank}(\mathbf{G}_r) = F$, each column of $\bar{\mathbf{U}}_r^{[s]} = \begin{bmatrix} \bar{\mathbf{u}}_{r,1} & \bar{\mathbf{u}}_{r,2} & \cdots & \bar{\mathbf{u}}_{r,l_r} \end{bmatrix}$ is a different linear combination of F damped cisoids, namely, $\alpha_{r,1} e^{j\omega_{r,1}}, \alpha_{r,2} e^{j\omega_{r,2}}, \cdots, \alpha_{r,F} e^{j\omega_{r,F}}$. According to the sinusoidal LP property, we have:

$$[\bar{\mathbf{u}}_{r,f}]_m + \sum_{i=1}^{F} c_{r,i} [\bar{\mathbf{u}}_{r,f}]_{m-i} = 0, \quad m = F+1, F+2, \ldots, M_r, \quad f = 1, 2, \ldots, l_r \tag{13}$$

where $\{c_{r,i}\}$, which are characterized by $\{\alpha_{r,f} e^{j\omega_{r,f}}\}$ only, are the LP coefficients for the rth dimension. That is, $\{a_{r,f}\}$ are given by the F roots of $z^F + \sum_{i=1}^{F} c_{i,r} z^{F-i} = 0$. Since $\{c_{r,i}\}$ are common for all columns in $\bar{\mathbf{U}}_r^{[s]}$, we can formulate a multi-channel estimation problem [23] for their determination by considering the l_r column vectors as l_r channel outputs.

Basically, we can exploit (12)–(13) to determine the frequencies and damping factors at each dimension individually from the noisy $\mathbf{U}_r^{[s]}$, $r = 1, 2, \cdots, R$. However, this straightforward idea has three drawbacks. First, we have to compute R SVDs according to (10). Second, parameter matching for the R-D parameters

is needed. Both correspond to extensive computational requirement particularly for large R, F and/or $\{M_r\}$. The last disadvantage is that $M_r > F$ for $r = 1, 2, \cdots, R$, and distinct frequencies at each of the R dimensions are required. In order to achieve computational efficiency and reduce the restriction, we propose an alternative estimation procedure as follows. The main strategy is first to estimate the frequencies and damping factors at only one dimension by exploiting (12). We then make use of these estimates and the inherent structure of $\boldsymbol{\mathcal{X}}$ as well as tensor algebra to determine the parameters of the remaining dimensions such that R-D frequency matching is automatically achieved. According to our assumptions in Section II that $M_R > F$ and $\omega_{R,i} \neq \omega_{R,j}$, $i \neq j$, $i, j = 1, 2, \cdots, F$, we start the estimation at the Rth dimension from $\mathbf{U}_R^{[s]}$ by following [23] which utilizes the LP property and WLS. Nevertheless, we will discuss the selection of the first dimension for processing at the end of this section.

For sufficiently small noise condition, each column of $\mathbf{U}_R^{[s]}$ will approximately satisfy (13). Defining $\mathbf{c} = \begin{bmatrix} c_{R,1} & c_{R,2} & \cdots & c_{R,F} \end{bmatrix}$ and combining all the $l_R = \min\{F, M_R, M/M_R\} = \min\{F, M/M_R\}$ columns in $\mathbf{U}_R^{[s]}$, the LP error vector, denoted by $\mathbf{e}$, is constructed as:

$$\mathbf{e} = \mathbf{D}\tilde{\mathbf{c}} - \mathbf{d} \tag{14}$$

where

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1^T & \mathbf{D}_2^T & \cdots & \mathbf{D}_{l_R}^T \end{bmatrix}^T \tag{15}$$

$$\mathbf{d} = \begin{bmatrix} \mathbf{d}_1^T & \mathbf{d}_2^T & \cdots & \mathbf{d}_{l_R}^T \end{bmatrix}^T \tag{16}$$

$$\mathbf{D}_f = \begin{bmatrix} [\mathbf{u}_{R,f}]_F & [\mathbf{u}_{R,f}]_{F-1} & \cdots & [\mathbf{u}_{R,f}]_1 \\ [\mathbf{u}_{R,f}]_{F+1} & [\mathbf{u}_{R,f}]_F & \cdots & [\mathbf{u}_{R,f}]_2 \\ \vdots & \vdots & \ddots & \cdots \\ [\mathbf{u}_{R,f}]_{M_R-1} & [\mathbf{u}_{R,f}]_{M_R-2} & \cdots & [\mathbf{u}_{R,f}]_{M_R-F} \end{bmatrix} \tag{17}$$

$$\mathbf{d}_f = -\begin{bmatrix} [\mathbf{u}_{R,f}]_{F+1} & [\mathbf{u}_{R,f}]_{F+2} & \cdots & [\mathbf{u}_{R,f}]_{M_R} \end{bmatrix}^T, \quad f = 1, 2, \cdots, l_R \tag{18}$$

The WLS estimate of $\mathbf{c}$ is computed as

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

where $\mathbf{W} \in \mathbb{C}^{l_R(M_R-F) \times l_R(M_R-F)}$ represents the weighting matrix. In this study, we apply the Gauss-Markov theorem [21] to derive $\mathbf{W}$ which is optimum in the sense of producing the minimum variance. Let

$$\mathbf{A}(\mathbf{c}) = \text{Toeplitz}\left(\begin{bmatrix} c_{R,1} & \mathbf{0}_{1\times(M_R-F-1)} \end{bmatrix}^T, \begin{bmatrix} c_{R,F} & c_{R,F-1} & \cdots & c_{R,1} & 1 & \mathbf{0}_{1\times(M_R-F-1)} \end{bmatrix}\right) \tag{20}$$

The optimum weight is computed using the covariance matrix of the residual error vector, which is a function of $\mathbf{c}$, denoted by $\mathbf{W}(\mathbf{c})$:

$$\mathbf{W}(\mathbf{c}) = \sigma^2 \left[\mathbb{E}\left\{(\mathbf{D}\mathbf{c}-\mathbf{d})(\mathbf{D}\mathbf{c}-\mathbf{d})^H\right\}\right]^{-1} = \sigma^2 \left[\mathbb{E}\left\{\text{vec}\left(\mathbf{A}(\mathbf{c})\mathbf{U}_R^{[s]}\right)\text{vec}\left(\mathbf{A}(\mathbf{c})\mathbf{U}_R^{[s]}\right)^H\right\}\right]^{-1} \quad (21)$$

Expressing $\mathbf{U}_R^{[s]} = \bar{\mathbf{U}}_R^{[s]} + \boldsymbol{\Delta}\mathbf{U}_R^{[s]}$ and using (1) and (10), we apply [24] to obtain

$$\boldsymbol{\Delta}\mathbf{U}_R^{[s]} \approx \bar{\mathbf{U}}_R^{[s]}\boldsymbol{\Theta} + [\boldsymbol{\mathcal{Q}}]_{(R)}\bar{\mathbf{V}}_R^{[s]}(\bar{\boldsymbol{\Lambda}}_R^{[s]})^{-1} - \bar{\mathbf{U}}_R^{[s]}\bar{\mathbf{U}}_R^{[s]H}[\boldsymbol{\mathcal{Q}}]_{(R)}\bar{\mathbf{V}}_R^{[s]}(\bar{\boldsymbol{\Lambda}}_R^{[s]})^{-1} \quad (22)$$

where $\bar{\mathbf{V}}_R^{[s]}$ and $\bar{\boldsymbol{\Lambda}}_R^{[s]} = \text{diag}(\left[\bar{\lambda}_{R,1}^2, \bar{\lambda}_{R,2}^2, \cdots, \bar{\lambda}_{R,F}^2\right])$ are the noise-free signal subspace components and $\boldsymbol{\Theta} = \mathbf{Z} \odot \left(\bar{\mathbf{U}}_R^{[s]H}[\boldsymbol{\mathcal{Q}}]_{(R)}\bar{\mathbf{V}}_R^{[s]}\bar{\boldsymbol{\Lambda}}_R^{[s]} + \bar{\boldsymbol{\Lambda}}_R^{[s]}\bar{\mathbf{V}}_R^{[s]H}[\boldsymbol{\mathcal{Q}}]_{(R)}^H\bar{\mathbf{U}}_R^{[s]}\right)$ with $[\mathbf{Z}]_{m,n}$ being equal to $1/(\bar{\lambda}_{R,n} - \bar{\lambda}_{R,m})$ for $m \neq n$ and zero otherwise. Using the fact that $\mathbf{A}(\mathbf{c})\bar{\mathbf{U}}_R^{[s]} = \mathbf{0}_{(M-K)\times M}$ and (22), we have

$$\mathbf{A}(\mathbf{c})\mathbf{U}_R^{[s]} = \mathbf{A}(\mathbf{c})\boldsymbol{\Delta}\mathbf{U}_R^{[s]} \approx \mathbf{A}(\mathbf{c})[\boldsymbol{\mathcal{Q}}]_{(R)}\bar{\mathbf{V}}_R^{[s]}(\bar{\boldsymbol{\Lambda}}_R^{[s]})^{-1} \quad (23)$$

Vectorizing both sides of (23) yields

$$\text{vec}\left(\mathbf{A}(\mathbf{c})\mathbf{U}_R^{[s]}\right) \approx \left((\bar{\boldsymbol{\Lambda}}_R^{[s]})^{-1}\bar{\mathbf{V}}_R^{[s]T} \otimes \mathbf{A}(\mathbf{c})\right)\text{vec}([\boldsymbol{\mathcal{Q}}]_{(R)}). \quad (24)$$

As $\boldsymbol{\mathcal{Q}}$ is independent and identical distributed (IID), its Rth unfolding $[\boldsymbol{\mathcal{Q}}]_{(R)}$ is also IID, then we have $\mathbb{E}\{\text{vec}([\boldsymbol{\mathcal{Q}}]_{(R)})\text{vec}([\boldsymbol{\mathcal{Q}}]_{(R)})^H\} = \sigma^2\mathbf{I}_M$. With the use of (24), we obtain

$$\mathbb{E}\left\{\text{vec}\left(\mathbf{A}(\mathbf{c})\mathbf{U}_R^{[s]}\right)\text{vec}\left(\mathbf{A}(\mathbf{c})\mathbf{U}_R^{[s]}\right)^H\right\} \approx \left[\boldsymbol{\Lambda}_R^{-2} \otimes \mathbf{A}(\mathbf{c})\mathbf{A}(\mathbf{c})^H\right]^{-1} \quad (25)$$

where $\{\bar{\lambda}_{R,f}\}$ are replaced by $\{\lambda_{R,f}\}$ in practice. Substituting (25) into (21) yields

$$\mathbf{W}(\mathbf{c}) \approx (\boldsymbol{\Lambda}_R^{[s]})^2 \otimes \left(\mathbf{A}(\mathbf{c})\mathbf{A}(\mathbf{c})^H\right)^{-1} \quad (26)$$

As $\mathbf{W}(\mathbf{c})$ is block diagonal, (19) can be simplified to

$$\hat{\mathbf{c}} = \left(\sum_{f=1}^{l_R}\lambda_{R,f}^2\mathbf{D}_f^H\left(\mathbf{A}(\mathbf{c})\mathbf{A}(\mathbf{c})^H\right)^{-1}\mathbf{D}_f\right)^{-1}\left(\sum_{f=1}^{l_R}\lambda_{R,f}^2\mathbf{D}_f^H\left(\mathbf{A}(\mathbf{c})\mathbf{A}(\mathbf{c})^H\right)^{-1}\mathbf{d}_f\right) \quad (27)$$

which is more computational attractive. As (26) is not known $a\ priori$, we employ the iterative relaxation procedure [30] to solve for $\{\omega_{R,f}\}$ and $\{\alpha_{R,f}\}$, which is summarized in Table 2. Note that this relaxation approach corresponds to the IQML technique [25] or the Steiglitz-McBride algorithm [26], which has local convergence property with linear rate of convergence [27].

When $\{\hat{a}_{R,f}\}$ are available, the parameters in the remaining dimensions are then determined as follows. According to (1), (2) and (5), we express $\boldsymbol{\mathcal{X}}$ as

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{X}}_{\text{sub}} \times_R \mathbf{G}_R \quad (28)$$

(i) Compute the SVD of $[\boldsymbol{\mathcal{Y}}]_{(R)}$ to obtain $\mathbf{U}_R^{[s]}$
(ii) Set $\mathbf{A}(\hat{\mathbf{c}})\mathbf{A}^H(\hat{\mathbf{c}}) = \mathbf{I}_{M_R - F}$
(iii) Calculate $\hat{\mathbf{c}}$ using (27) with $\mathbf{A}(\mathbf{c}) = \mathbf{A}(\hat{\mathbf{c}})$
(iv) Compute an updated version of $\mathbf{A}(\hat{\mathbf{c}})$ using (20)
(v) Repeat Steps (iii)-(iv) until a stopping criterion is reached. In our study, we choose a fixed number of iterations κ as the criterion
(vi) Solve all roots of $z^F + \sum_{i=1}^{F} \hat{c}_{R,i} z^{F-i} = 0$, denoted by $\hat{a}_{R,f}$, $f = 1, 2, \cdots, F$
(vii) Estimate the frequencies and damping factors of the Rth dimension as $\hat{\omega}_{R,f} = \angle(\hat{a}_{R,f})$ and $\hat{\alpha}_{R,f} = \lvert\hat{a}_{R,f}\rvert$, $f = 1, 2, \cdots, F$

Table 2: Estimation algorithm at first stage

where
$$\begin{aligned}\boldsymbol{\mathcal{X}}_{\text{sub}} &= \boldsymbol{\mathcal{C}}_{\lambda_1,\lambda_2,\ldots,\lambda_F} \times_1 \mathbf{G}_1 \times_2 \mathbf{G}_2 \cdots \times_{R-1} \mathbf{G}_{R-1} \\ &= \boldsymbol{\mathcal{I}}^R \times_1 \mathbf{G}_1 \times_2 \mathbf{G}_2 \cdots \times_{R-1} \mathbf{G}_{R-1} \times_R \boldsymbol{\Sigma} \end{aligned} \tag{29}$$

with $\boldsymbol{\mathcal{I}}^R \in \mathbb{C}^{F\times F\times\cdots\times F}$ being a R-D tensor whose diagonal elements are one and zero otherwise. Therefore, the sub-tensor $\boldsymbol{\mathcal{X}}_{\text{sub}} \in \mathbb{C}^{M_1\times M_2\times\cdots\times M_{R-1}\times F}$ can be represented as a concatenation of F tensors in the Rth dimension:

$$\boldsymbol{\mathcal{X}}_{\text{sub}} = \boldsymbol{\mathcal{X}}_{\text{sub}1} \sqcup_R \boldsymbol{\mathcal{X}}_{\text{sub}2} \sqcup_R \cdots \sqcup_R \boldsymbol{\mathcal{X}}_{\text{sub}F} \tag{30}$$

where

$$\boldsymbol{\mathcal{X}}_{\text{sub}f} = \gamma_f \left(\mathbf{g}_{1,f} \circ \mathbf{g}_{2,f} \circ \cdots \circ \mathbf{g}_{R-1,f}\right) \tag{31}$$

Note that the Rth unfolding of the sub-tensor $\boldsymbol{\mathcal{X}}_{\text{sub}}$ is:

$$[\boldsymbol{\mathcal{X}}_{\text{sub}}]_{(R)} = \left[[\boldsymbol{\mathcal{X}}_{\text{sub}1}]_{(R)}^T \quad [\boldsymbol{\mathcal{X}}_{\text{sub}2}]_{(R)}^T \quad \cdots \quad [\boldsymbol{\mathcal{X}}_{\text{sub}F}]_{(R)}^T\right]^T \tag{32}$$

where
$$[\boldsymbol{\mathcal{X}}_{\text{sub}f}]_{(R)}^T = \gamma_f \mathbf{g}_{1,f} \odot \mathbf{g}_{2,f} \odot \cdots \odot \mathbf{g}_{R-1,f} \tag{33}$$

Substituting $\{a_{R,f}\}$ with $\{\hat{a}_{R,f}\}$ in (4) and (6), we obtain an estimate of $\mathbf{G}_R$, namely, $\widehat{\mathbf{G}}_R$. Furthermore, we replace $\boldsymbol{\mathcal{X}}$ and $\mathbf{G}_R$ by the observed $\boldsymbol{\mathcal{Y}}$ and $\widehat{\mathbf{G}}_R$, respectively, in (28), an estimate of $\boldsymbol{\mathcal{X}}_{\text{sub}}$, denoted by $\boldsymbol{\mathcal{Z}}$, is then obtained as

$$\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{Y}} \times_R \widehat{\mathbf{G}}_R^{\dagger} = \boldsymbol{\mathcal{Z}}_1 \sqcup_R \boldsymbol{\mathcal{Z}}_2 \sqcup_R \cdots \sqcup_R \boldsymbol{\mathcal{Z}}_F \tag{34}$$

where $\boldsymbol{\mathcal{Z}}_f$ is a $(R-1)$-D tensor corresponding to $\boldsymbol{\mathcal{X}}_{\text{sub}f}$, $f = 1, 2, \cdots, F$. In practice, we compute (34) with the use of matrix operations as

$$[\boldsymbol{\mathcal{Z}}]_{(R)} = \widehat{\mathbf{G}}_R^{\dagger}[\boldsymbol{\mathcal{Y}}]_{(R)} = \left[[\boldsymbol{\mathcal{Z}}_1]_{(R)}^T \quad [\boldsymbol{\mathcal{Z}}_2]_{(R)}^T \quad \cdots \quad [\boldsymbol{\mathcal{Z}}_F]_{(R)}^T\right]^T \tag{35}$$

where the Rth unfolding components of $\{[\boldsymbol{\mathcal{Z}}_f]_{(R)}\}$ are vectors. Comparing (30)–(31) and (34), we deduce that $\boldsymbol{\mathcal{Z}}_f$ can be approximated as:

$$\boldsymbol{\mathcal{Z}}_f \approx \mathbf{h}_{1,f} \circ \mathbf{h}_{2,f} \cdots \circ \mathbf{h}_{R-1,f} \tag{36}$$

where
$$\mathbf{h}_{r,f} = \beta_{r,f}\mathbf{g}_{r,f}, \quad r = 1, 2, \cdots, R-1, \quad f = 1, 2, \cdots, F \tag{37}$$

with $\{\beta_{r,f}\}$ being unknown constants. According to (36), the rth unfolding of $\boldsymbol{\mathcal{Z}}_f$, $r = 1, 2, \cdots, R-1$, is thus:

$$[\boldsymbol{\mathcal{Z}}_f]_{(r)} \approx \mathbf{h}_{r,f}\left(\mathbf{h}_{r+1,f} \odot \cdots \odot \mathbf{h}_{R-1,f} \odot \mathbf{h}_{1,f} \odot \cdots \odot \mathbf{h}_{r-1,f}\right)^T \tag{38}$$

indicating its rank is one in the absence of noise. Based on (37)–(38), we utilize the SVD of $[\boldsymbol{\mathcal{Z}}_f]_{(r)}$ to estimate each of the remaining $\{a_{r,f}\}$ in a separable manner as follows. According to the rank-1 property, an approximation of $[\boldsymbol{\mathcal{Z}}_f]_{(r)}$ by SVD truncation is:

$$[\boldsymbol{\mathcal{Z}}_f]_{(r)} \approx \lambda_{r,f,1}\mathbf{u}_{r,f,1}\mathbf{v}_{r,f,1}^H \tag{39}$$

where $\lambda_{r,f,1}$, $\mathbf{u}_{r,f,1}$ and $\mathbf{v}_{r,f,1}$ are the largest singular value and the corresponding left and right singular vectors of $[\boldsymbol{\mathcal{Z}}_f]_{(r)}$, respectively. That is to say, $\mathbf{u}_{r,f,1}$ is a noisy version of the 1-D single-tone sequence $\mathbf{h}_{r,f}$ up to an unknown multiplication scalar. As a result, we can apply the same iterative relaxation procedure in Table 2 to find $\omega_{r,f}$ and $\alpha_{r,f}$ from each $\mathbf{u}_{r,f,1}$, and the steps are summarized in Table 3.

(i) Construct $\widehat{\mathbf{G}}_R$ with $a_{R,f} = \hat{a}_{R,f}$, $f = 1, 2, \cdots, F$, according to (4) and (6)

(ii) Compute $[\boldsymbol{\mathcal{Z}}]_{(R)}$ using (35)

(iii) For each f, $f = 1, 2, \cdots, F$, extract $\boldsymbol{\mathcal{Z}}_f$ from $[\boldsymbol{\mathcal{Z}}]_{(R)}$; for each r, $r = 1, 2, \cdots, R-1$, do the rth unfolding of $[\boldsymbol{\mathcal{Z}}_f]$, denoted by $[\boldsymbol{\mathcal{Z}}_f]_{(r)}$, and compute its left principal singular vector, namely, $\mathbf{u}_{r,f,1}$

(iv) Set $\mathbf{W}(\hat{a}_{r,f}) = \mathbf{I}_{M_r-1}$

(v) Compute $\hat{a}_{r,f} = (\mathbf{d}_{r,f,1}^H\mathbf{W}(a_{r,f})\mathbf{d}_{r,f,2})/(\mathbf{d}_{r,f,1}^H\mathbf{W}(a_{r,f})\mathbf{d}_{r,f,1})$ where $\mathbf{d}_{r,f,1}$, $\mathbf{d}_{r,f,2}$ being $\mathbf{u}_{r,f,1}$ without the last and the first element and $\mathbf{W}(a_{r,f}) = \mathbf{W}(\hat{a}_{r,f})$

(vi) Find an updated $\mathbf{W}(\hat{a}_{r,f}) = (\mathbf{A}(\hat{a}_r)\mathbf{A}^H(\hat{a}_r))^{-1}$ where $\mathbf{A}(\hat{a}_r) = \text{Toeplitz}\left(\begin{bmatrix}-\hat{a}_r & \mathbf{0}_{1\times(M_r-2)}\end{bmatrix}^T, \begin{bmatrix}-\hat{a}_r & 1 & \mathbf{0}_{1\times(M_r-2)}\end{bmatrix}\right)$

(vii) Repeat Steps (v)-(vi) for κ iterations

(viii) Estimate the frequencies and damping factors of the rth dimension as $\hat{\omega}_{r,f} = \angle(\hat{a}_{r,f})$ and $\hat{\alpha}_{r,f} = |\hat{a}_{r,f}|$

Table 3: Estimation algorithm at second stage

From Table 3, it is observed that even the frequencies are identical in all remaining dimensions, $r = 1, 2, \cdots, R-1$, the algorithm still works properly because each of the $\{\hat{a}_{r,f}\}$ is separately estimated.

Moreover, since at least two samples are needed in single-tone frequency estimation, the operating requirement is $\min\{M_1, M_2, \ldots, M_{R-1}\} \geq 2$ with $M_R > F$. This identifiability is inferior to the existing results [13]. Although the proposed scheme shows inferiority in identifiability, it is able to outperform the conventional methods in terms of computational complexity and accuracy.

The last issue we have to address for the proposed approach is to determine the appropriate dimension, denoted by r^o, instead of R, in the first estimation stage. Based on the above development, it is necessary that $\mathbf{G}_{r^o}$ spans full column rank, which means that all frequencies in the r^oth dimension are distinct and $M_{r^o} \geq F+1$. The former requirement also suggests that the minimum separation between any two adjacent frequencies at the r^oth dimension should be large among all dimensions. As a result, we suggest to compute r^o as:

$$r^o = \arg \max_{r \in [1,2,\ldots,R]} \left((M_r - 1) \min_{i \neq j} |\hat{\omega}_{r,i} - \hat{\omega}_{r,j}| \right), \quad \text{subject} \quad \text{to} \quad M_{r^o} > F \tag{40}$$

That is, for each r such that $M_r > F$, we compute the weighted minimum frequency separation between $\{\hat{\omega}_{r,i}\}$, $i = 1, 2, \cdots, F$, and r^o corresponds to the dimension with the maximum smallest difference. Analogous to the standard ESPRIT method [9], rough values of all $\{\hat{\omega}_{r,i}\}$ are determined using one SVD in our study in order to achieve computational efficiency in the dimension selection step. As the conventional ESPRIT approach maximizes the data reuse via spatial smoothing to boost estimation performance, an obvious way for complexity reduction is to decrease the data redundancy. Inspiring by [8], [28], we construct $\boldsymbol{\mathcal{Y}}_{ss} \in \mathbb{C}^{L_1 \times L_2 \times \cdots \times L_R \times (K_1 K_2 \ldots K_R)}$ with less redundancy from $\boldsymbol{\mathcal{Y}}$, which has the form of:

$$\begin{aligned} \boldsymbol{\mathcal{Y}}_{ss} = [&\boldsymbol{\mathcal{Y}}_{1,1,\ldots,1} \sqcup_{R+1} \boldsymbol{\mathcal{Y}}_{2,1,\ldots,1} \sqcup_{R+1} \cdots \sqcup_{R+1} \boldsymbol{\mathcal{Y}}_{K_1,1,\ldots,1} \sqcup_{R+1} \boldsymbol{\mathcal{Y}}_{1,2,\ldots,1} \\ &\sqcup_{R+1} \cdots \sqcup_{R+1} \boldsymbol{\mathcal{Y}}_{K_1,2,\ldots,1} \sqcup_{R+1} \boldsymbol{\mathcal{Y}}_{1,K_2,\ldots,K_R} \sqcup_{R+1} \cdots \sqcup_{R+1} \boldsymbol{\mathcal{Y}}_{K_1,K_2,\ldots,K_R}] \end{aligned} \tag{41}$$

where $\boldsymbol{\mathcal{Y}}_{k_1,k_2,\ldots,k_R} = \boldsymbol{\mathcal{Y}} \times_1 \mathbf{J}_{k_1}^{K_1} \times_2 \mathbf{J}_{k_2}^{K_2} \cdots \times_R \mathbf{J}_{k_R}^{K_R}$, $\mathbf{J}_{k_r}^{K_r} = \begin{bmatrix} \mathbf{0}_{L_r \times (k_r-1)p_r} & \mathbf{I}_{L_r} & \mathbf{0}_{L_r \times (K_r-k_r)p_r} \end{bmatrix}$ and $(K_r - 1)p_r + L_r = M_r$, with K_r, L_r and p_r, $r = 1, 2, \ldots, R$, being integers to be determined. The $\{p_r\}$ whose values are between 1 and L_r, can be viewed as the reuse factors. When $p_r = 1$ for all r, $\boldsymbol{\mathcal{Y}}_{ss}$ involves maximum data reuse as in [10]. On the other hand, there is no redundancy in $\boldsymbol{\mathcal{Y}}_{ss}$ if $p_r = L_r$, corresponding to a smoothed tensor with the smallest size for a particular set of $\{L_r\}$. Considering that the $(R+1)$-D tensor $\boldsymbol{\mathcal{Y}}_{ss}$ is composed of R-D sinusoids with $K_1 K_2 \cdots K_R$ snapshots, then we can straightforwardly apply the matrix-based standard ESPRIT algorithm [8] to solve for all $\{\hat{\omega}_{r,i}\}$ with $M_r > F$. Basically, one SVD for $[\boldsymbol{\mathcal{Y}}_{ss}]_{(R+1)}^T$ is required and the interested reader is referred

to [8] for the detailed estimation procedure. To reduce the redundancy, we simply assign $K_r = \lceil F/2^R \rceil$, $r = 1, 2, \ldots, R$ and $p_i \approx L_i$ subject to $\prod_{r=1}^{R} L_r \geq F$ in our study.

Finally, the major computational complexity of the proposed algorithm is studied as follows. In the step of finding r^o, we apply the standard ESPRIT method [8] to $\boldsymbol{\mathcal{Y}}_{ss}$, which has a complexity of $\mathcal{O}(k_t F \prod_{r=1}^{R} L_r K_r)$, where k_t is a constant depends on the design of the algorithm [29]. With appropriate choices of L_r, K_r and p_r, we have $p_r \approx L_r$ which results in $\mathcal{O}(k_t F \prod_{r=1}^{R} L_r K_r) \approx \mathcal{O}(k_t F \prod_{r=1}^{R} M_r) = \mathcal{O}(k_t F M)$. The main calculations in Tables 2 and 3 are computing the SVD and matrix inverse. In the first stage of computing $\hat{\omega}_{r^o,f}$ and $\hat{\alpha}_{r^o,f}$, the SVD operation in (10) has an order of $\mathcal{O}(k_t F \prod_{r=1}^{R} M_r) = \mathcal{O}(k_t F M)$ while that of (27) is less than $\mathcal{O}((F+1)M_{r^o}^3)$. In estimation of the remaining frequencies and damping factors, the complexity of computing the principal left singular vector of $\mathbf{u}_{r,f,1}$ from $[\boldsymbol{\mathcal{Z}}_f]_{(r)}$, $r = 1, 2, \cdots, r^o - 1, r^o + 1, \cdots, R$, $f = 1, 2, \cdots, F$ has a complexity of $O(k_t M F (R-1)/M_{r^o}) < O(k_t M (R-1))$ as $F < M_{r^o}$, while the cost for each matrix inverse in Table 3 is $\mathcal{O}(M_r^3)$. As a result, the computational requirement of the major steps of the TPUMA approach is less than $O(k_t M (R + F - 1)) + \sum_{r=1}^{R} \mathcal{O}(\kappa (F+1) M_r^3)$ where κ denotes the required number of iterations. It is found from extensive computer simulations that the TPUMA algorithm converges only in a few iterations, say, $\kappa = 3$. As a comparison with the ESPRIT methodology in terms of SVD operations, the matrix-based and tensor-based ESPRIT algorithms have complexities of $\mathcal{O}(k_t L^E K^E F)$ and $\mathcal{O}(k_t L^E K^E F (R+1) + L^E K^E F R + L^E F^2 R)$, respectively [10], where $L_r^E + K_r^E - 1 = M_r$, $L_r^E > 1$, $K_r^E > 1$, $L^E = \prod_{r=1}^{R} L_r^E$ and $K^E = \prod_{r=1}^{R} K_r^E$. Furthermore, extra parameter pairing procedure is needed in the ESPRIT approach and thus we expect that the proposed algorithm is more computationally attractive.

IV. Modification for Single-Tone

In this section, the special case when $\boldsymbol{\mathcal{X}}$ is an R-D undamped single-tone is investigated. We will show that the TPUMA algorithm for single-tone estimation can be simplified, which leads to significant reduction in complexity. Although there are no closed-form expressions for the performance of the proposed algorithm in the general case, we are able to provide its closed-form variance expression for the single-tone case, which is shown to be equal to CRLB.

The signal model in (2) is now simplified as:

$$x_{m_1,m_2,\ldots,m_R} = \gamma \prod_{r=1}^{R} e^{j\omega_r m_r} \tag{42}$$

where the parameters of interest are ω_r, $r = 1, 2, \cdots, R$. Without loss of generality and for presentation consistency, we start the estimation at the Rth dimension. According to (3)–(4), $\boldsymbol{\mathcal{X}}$ becomes

$$\boldsymbol{\mathcal{X}} = \gamma \mathbf{g}_1 \circ \mathbf{g}_2 \cdots \circ \mathbf{g}_R \tag{43}$$

where

$$\mathbf{g}_r = \begin{bmatrix} a_r & a_r^2 & \cdots & a_r^{M_r} \end{bmatrix}^T, \quad a_r = e^{j\omega_r} \tag{44}$$

Noting that the rank of $\boldsymbol{\mathcal{X}}$ is one and following (9)–(12), the estimate of $[\boldsymbol{\mathcal{X}}]_{(R)}$ is:

$$[\widehat{\boldsymbol{\mathcal{X}}}]_{(R)} \approx \lambda_{R,1} \mathbf{u}_{R,1} \mathbf{v}_{R,1}^H \tag{45}$$

where

$$\mathbf{u}_{R,1} \approx \beta_u \mathbf{g}_R \tag{46}$$

$$\mathbf{v}_{R,1}^* \approx \beta_v \mathbf{g}_1 \odot \mathbf{g}_2 \odot \cdots \odot \mathbf{g}_{R-1} \tag{47}$$

That is, $\lambda_{R,1}$, $\mathbf{u}_{R,1}$ and $\mathbf{v}_{R,1}$ are the largest singular value and the corresponding left and right singular vectors of $[\widehat{\boldsymbol{\mathcal{X}}}]_{(R)}$, respectively. On the other hand, β_u and β_v are unknown scalars which are analogous to $\boldsymbol{\Omega}_r$ in (12) and $\beta_{r,f}$ in (37), although the relationships in (46)–(47) are simpler. Note that we can follow [14] to obtain $\bar{\lambda}_{R,1} = \sqrt{M}|\gamma|$ and $\bar{\mathbf{u}}_{R,1} = \mathbf{g}_R e^{-j\varphi_g}/\sqrt{M_R}$ where $\varphi_g \in (-\pi, -\pi]$ is unknown, in the noise-free case. As a result, $\mathbf{u}_{R,1}$ corresponds to a noisy tone sequence whose frequency is ω_R. Based on the sinusoidal LP property and WLS, we follow the development in (13)–(21) to obtain the estimate of a_R as:

$$\hat{a}_R = (\mathbf{x}_1^H \mathbf{W}(a_R) \mathbf{x}_2)^{-1} (\mathbf{x}_1^H \mathbf{W}(a_R) \mathbf{x}_1) \tag{48}$$

where

$$\mathbf{x}_1 = \begin{bmatrix} [\mathbf{u}_{R,1}]_1 & [\mathbf{u}_{R,1}]_2 & \cdots & [\mathbf{u}_{R,1}]_{M_R-1} \end{bmatrix}^T \tag{49}$$

$$\mathbf{x}_2 = \begin{bmatrix} [\mathbf{u}_{R,1}]_2 & [\mathbf{u}_{R,1}]_3 & \cdots & [\mathbf{u}_{R,1}]_{M_R} \end{bmatrix}^T \tag{50}$$

$$\mathbf{W}(a_R) = \sigma^2/\bar{\lambda}_1^2 \left[\mathbb{E}\left\{ \mathbf{A}(a_R) \mathbf{u}_{R,1} \mathbf{u}_{R,1}^H \mathbf{A}^H(a_R) \right\} \right]^{-1} \approx (\mathbf{A}(a_R) \mathbf{A}^H(a_R))^{-1} \tag{51}$$

with $\mathbf{A}(a_R) = \text{Toeplitz}\left(\begin{bmatrix} -a_R & \mathbf{0}_{1\times(M_R-2)} \end{bmatrix}^T, \begin{bmatrix} -a_R & 1 & \mathbf{0}_{1\times(M_R-2)} \end{bmatrix} \right)$.

Using the substitution of $a_R = e^{j\omega_R}$ yields a closed-form expression for $\mathbf{W}(a_R) = \mathbf{W}(\omega_R)$ with elements [30]:

$$[\mathbf{W}(\omega_R)]_{m,n} = \frac{M_R \min(m,n) - mn}{M_R} e^{j(m-n)\omega_R}, \quad m = 1, 2, \cdots, M_R - 1, \quad n = 1, 2, \cdots, M_R - 1 \tag{52}$$

As $\mathbf{x}_1^H \mathbf{W}(\omega_R) \mathbf{x}_1$ is real and positive [30], we simplify (48) to obtain the conceptual estimate of ω_R as:

$$\hat{\omega}_R = \angle \left(\mathbf{x}_1^H \mathbf{W}(\omega_R) \mathbf{x}_2 \right) \tag{53}$$

which is practically solved via an iterative and relaxation manner as in Table 3.

Following (28)–(39), we substitute ω_R with $\hat{\omega}_R$ in (44) to construct $\hat{\mathbf{g}}_R$ and then compute the Rth unfolding of the sub-tensor $\boldsymbol{\mathcal{Z}}$ as:

$$[\boldsymbol{\mathcal{Z}}]_{(R)} = \hat{\mathbf{g}}_R^{\dagger}[\boldsymbol{\mathcal{Y}}]_{(R)} \tag{54}$$

Finally, we utilize (53) to solve for ω_r from $\mathbf{u}_{r,1}$, which is the principal left singular vector of $[\boldsymbol{\mathcal{Z}}]_{(r)}$, $r = 1, 2, \cdots, R-1$. The complete estimation procedure for R-D single-tone frequency estimation is summarized in Table 4.

(i) Compute the principal left singular vector of $[\boldsymbol{\mathcal{Y}}]_{(R)}$, namely, $\mathbf{u}_{R,1}$

(ii) Obtain an initial frequency estimate $\hat{\omega}_R$ using (53) with $[\mathbf{W}(\omega_R)]_{m,n} = 0$ for $m \neq n$ in (52)

(iii) Construct $\mathbf{W}(\omega_R)$ according to (52) with $\hat{\omega}_R = \omega_R$

(iv) Compute an updated $\hat{\omega}_R$ using (53)

(v) Repeat Steps (iii)–(iv) for κ iterations

(vi) Use $\hat{\omega}_R$ to construct $[\boldsymbol{\mathcal{Z}}]_{(R)}$ according to (54) and compute the principal left singular vector of each $[\boldsymbol{\mathcal{Z}}]_{(r)}$, namely, $\mathbf{u}_{r,1}$, $r = 1, 2, \cdots, R-1$

(vii) Repeat the iterative procedure in Steps (ii)–(v) to obtain $\hat{\omega}_r$ from $\mathbf{u}_{r,1}$, $r = 1, 2, \cdots, R-1$

Table 4: Frequency estimation algorithm for R-D single-tone

The means and variances of $\{\omega_r\}$ are now analyzed. Upon the global convergence such that $\hat{\omega}_r$ is located at a reasonable proximity of ω_r, $r = 1, 2, \cdots, R$, we have proved that (See Appendix A):

$$\lim_{\text{SNR}\to\infty} \mathbb{E}\{\hat{\omega}_r\} = \omega_r \tag{55}$$

$$\text{var}(\hat{\omega}_r) \approx \frac{6\sigma^2}{M(M_r^2 - 1)|\gamma|^2} \tag{56}$$

which is the CRLB for R-D frequency estimation (see Appendix B). Expressions (55) and (56) indicate that the proposed scheme provides asymptotically unbiased frequency estimates and is efficient, respectively.

V. Numerical Examples

Computer simulations have been carried out to evaluate the performance of the proposed approach by comparing with the standard ESPRIT (SE) [8], unitary ESPRIT (UE) [9], standard tensor ESPRIT (STE), unitary tensor ESPRIT (UTE) [10], single-snapshot unitary tensor ESPRIT (SSUTE) [11], MDF [12] and improved MDF (IMDF) [13] schemes. The stopping criterion of the PUMA algorithm is a fixed

number of iterations, and $\kappa = 3$ iterations are used because no significant improvement is observed for more iterations. All elements in the noise tensor $\boldsymbol{\mathcal{Q}}$ are zero-mean white complex Gaussian processes with identical variances of σ^2 which is adjusted for producing different SNR conditions. The average mean square error (MSE) is employed as the performance measure and CRLB is included as the optimality benchmark. All results provided are averages of 1000 independent runs.

In the first test, estimation of two damped 3-D sinusoids from $8 \times 8 \times 8$ data sets is investigated and the appropriate methods for comparison are SE, STE and MDF algorithms. The complex amplitudes and damping factors are assigned as $\begin{bmatrix}\gamma_1 & \gamma_2\end{bmatrix} = \begin{bmatrix}1e^{j1} & 2e^{j2}\end{bmatrix}$ and $\begin{bmatrix}\alpha_{1,1} & \alpha_{2,1} & \alpha_{3,1}\end{bmatrix} = \begin{bmatrix}0.99 & 0.99 & 0.99\end{bmatrix}$, $\begin{bmatrix}\alpha_{1,2} & \alpha_{2,2} & \alpha_{3,2}\end{bmatrix} = \begin{bmatrix}0.99 & 0.98 & 0.97\end{bmatrix}$. Three different frequency separation cases are investigated and the parameter settings are provided in Table 5. The frequencies of the first tone are fixed. Cases 1, 2 and 3 correspond to the two 3-D frequencies are well separated in one dimension and closely spaced in the remaining dimensions, closely spaced in all dimensions and closely spaced in one dimension and same in the remaining dimensions, respectively. The average MSEs of frequencies and damping factors versus SNR for the three cases are plotted in Figures 1 to 6. Figures 1 and 2 show the results for the first case and we see that the proposed method is superior to the SE, STE and MDF schemes at $\mathrm{SNR} \geq 2$ dB and its performance attains CRLB for $\mathrm{SNR} \geq 6$ dB. The MSEs of frequencies and damping factors for the second case are shown in Figures 3 and 4. It is observed that none of the examined algorithms is optimal, although the STE method performs the best when $\mathrm{SNR} \geq 20$ dB while the TPUMA and MDF methods have the smallest threshold SNR of 12 dB. Figures 5 and 6 show the results for the third case where only the proposed estimator gives the best accuracy for $\mathrm{SNR} \geq 38$ dB while the MDF scheme provides the best threshold performance. The average computation times of the SE, STE, MDF and TPUMA algorithms for a single trial are measured as 0.0119s, 0.1644s, 0.0221s and 0.0038s, respectively, which agree with the complexity analysis in Section III. Combining the findings in Figures 1 to 6, the MDF algorithm performs the best at lower SNR, while the proposed method is the most computationally attractive and is able to attain the highest accuracy at sufficiently high SNR conditions.

In the second test, estimation of two undamped 3-D sinusoids from $8 \times 8 \times 8$ data sets is investigated and the appropriate methods for comparison are UE, UTE, SSUTE and IMDF algorithms. We follow the parameter settings in the first case of the previous experiment except that the damping factors are now equal to unity. The average MSEs for frequency versus SNR are plotted in Figure 7 and it is seen that when $\mathrm{SNR} \geq 6$ dB, the proposed scheme outperforms other methods and its performance is close to CRLB. The average computation times of the UE, UTE, SSUTE, IMDF and TPUMA algorithms are measured as 0.0318s, 0.0509s, 0.5444s, 0.0609s and 0.0040s, indicating the significant complexity

First tone: $\begin{bmatrix}\omega_{1,1} & \omega_{2,1} & \omega_{3,1}\end{bmatrix} = \begin{bmatrix}0.2\pi & 0.1\pi & 0.8\pi\end{bmatrix}$	
Case 1: well separated in one dimension and closely spaced in remaining dimensions Second tone: $\begin{bmatrix}\omega_{1,2} & \omega_{2,2} & \omega_{3,2}\end{bmatrix} = \begin{bmatrix}0.5\pi & 0.04\pi & 0.86\pi\end{bmatrix}$	
Case 2: closely spaced in all dimensions Second tone: $\begin{bmatrix}\omega_{1,2} & \omega_{2,2} & \omega_{3,2}\end{bmatrix} \begin{bmatrix}0.26\pi & 0.04\pi & 0.86\pi\end{bmatrix}$	
Case 3: closely spaced in one dimension and same in remaining dimensions Second tone: $\begin{bmatrix}\omega_{1,2} & \omega_{2,2} & \omega_{3,2}\end{bmatrix} = \begin{bmatrix}0.26\pi & 0.1\pi & 0.8\pi\end{bmatrix}$	

Table 5: Three different frequency separation cases in first test

reduction in the TPUMA method.

The third test examines the single-tone case and the noise-free signal is the first 3-D sinusoid in the above experiment. The average MSEs are shown in Figure 8 and we observe that the proposed method is superior to the remaining schemes at $\mathrm{SNR} \geq -8$ dB. Moreover, its performance attains CRLB for $\mathrm{SNR} \geq -2$ dB, which collaborates the analysis of (56). On the other hand, the average computation times of the UE, UTE, SSUTE, IMDF and TPUMA algorithms are measured as 0.0295s, 0.0477s, 0.5201s, 0.0587s and 0.0009s, which again demonstrate the latter computational advantage.

In the fourth test, we examine the scenario when only one of the $\{M_r\}$ is larger than F. The noise-free signal consists of three undamped 3-D cisoids with $M_1 = M_2 = 2$ and $M_3 = 10$. The values of the complex amplitudes and frequencies are $\begin{bmatrix}\gamma_1 & \gamma_2 & \gamma_3\end{bmatrix} = \begin{bmatrix}1e^{j1} & 2e^{j2} & 1e^{j3}\end{bmatrix}$ and $\begin{bmatrix}\omega_{1,1} & \omega_{2,1} & \omega_{3,1}\end{bmatrix} = \begin{bmatrix}0.2\pi & 0.6\pi & 0.8\pi\end{bmatrix}$, $\begin{bmatrix}\omega_{1,2} & \omega_{2,2} & \omega_{3,2}\end{bmatrix} = \begin{bmatrix}0.4\pi & 0.1\pi & 0.5\pi\end{bmatrix}$, $\begin{bmatrix}\omega_{1,2} & \omega_{2,2} & \omega_{3,2}\end{bmatrix} = \begin{bmatrix}0.6\pi & 0.4\pi & 0.3\pi\end{bmatrix}$. The average MSE performance versus SNR is plotted in Figure 9. Note that in this case, the UE, UTE and SSUTE methods are equivalent. It is seen that when $\mathrm{SNR} \geq 12$ dB, the proposed estimator outperforms the UE and IMDF algorithms and its performance achieves the CRLB. Regarding the complexity, the measured computation times of the UE, IMDF and TPUMA schemes are 0.0047s, 0.0014s and 0.0027s.

In the final test, the computational time of the proposed algorithm versus M_r is examined because its complexity contains the term of $\sum_{r=1}^{R} \mathcal{O}(\kappa(F+1)M_r^3)$. Note that the TPUMA algorithm is still computationally attractive for large F because its complexity order is linearly increasing with F. Figure 10 shows the results of the proposed, UE, UTE, IMDF methods versus M_3 for an undamped 3-D model with fixing $F = 3$ and $M_1 = M_2 = 6$. Note that the SSUTE algorithm is not included here because its complexity is much higher. We again see the computational advantage of the TPUMA scheme for increasing M_3, and this also implies that the term of $O(k_t M(R+F-1))$ dominates $\sum_{r=1}^{R} \mathcal{O}(\kappa(F+1)M_r^3)$.

VI. Conclusion

A subspace-based R-dimensional (R-D) harmonic retrieval (HR) approach with $R > 2$ in additive white Gaussian noise, which is referred to as tensor principal-singular-vector utilization for modal analysis (TPUMA) algorithm, is devised. The sinusoidal parameters at one dimension are first estimated according to an iterative procedure which utilizes the linear prediction property and weighted least squares. The damping factors and frequencies in the remaining dimensions are then solved such that pairing of the R-D parameters is automatically achieved. We also modify the TPUMA method for a single R-D tone and prove that the frequency estimates are asymptotically unbiased and their variances attain Cramér-Rao lower bound (CRLB) at sufficiently high signal-to-noise ratio conditions. Computer simulations show that the TPUMA algorithm is computationally simpler than conventional HR estimators and its variance can attain CRLB even when there are identical frequencies at $(R-1)$ dimensions or the lengths in $(R-1)$ dimensions are less than the number of sinusoids. Nevertheless, the proposed scheme cannot work well when there are identical frequencies in all dimensions. As a future work, we will extend the TPUMA methodology for this challenging scenario.

Appendix A

We first produce the bias and variance for $\hat{\omega}_R$ of (53) and then utilize the derived results to the frequency estimates in the remaining dimensions. According to (48), the estimate of $a_R = e^{j\omega_R}$ is given by the $\hat{a}_R = \arg\min_{\tilde{a}_R} J(\tilde{a}_R)$ where

$$J(\tilde{a}_R) = (\mathbf{x}_1\tilde{a}_R - \mathbf{x}_2)^H \mathbf{W}(a_R)(\mathbf{x}_1\tilde{a}_R - \mathbf{x}_2) \tag{A.1}$$

At sufficiently high SNR conditions such that $\hat{a}_R$ is located at a reasonable proximity of a_R and assuming that $J''(\tilde{a}_R)$ is smooth enough around a_R, we expand $J'(\hat{a}_R)$ using Taylor series to yield [14]:

$$0 = J'(\hat{a}_R) = J'(a_R) + J''(a_R)(\hat{a}_R - a_R) \approx J'(a_R) + \mathbb{E}\{J''(a_R)\}(\hat{a}_R - a_R) \tag{A.2}$$

Expressing $\mathbf{x}_1 = \bar{\mathbf{x}}_1 + \mathbf{\Delta}\mathbf{x}_1$ and $\mathbf{x}_2 = \bar{\mathbf{x}}_2 + \mathbf{\Delta}\mathbf{x}_2$ as well as using $\bar{\mathbf{x}}_1 a_R = \bar{\mathbf{x}}_2$, $J'(a_R)$ is approximated as:

$$\begin{aligned} J'(a_R) &= 2\mathbf{x}_1^H \mathbf{W}(a_R)(\mathbf{x}_1 a_R - \mathbf{x}_2) \\ &= 2(\bar{\mathbf{x}}_1 + \mathbf{\Delta}\mathbf{x}_1)^H \mathbf{W}(a_R)\left[(\bar{\mathbf{x}}_1 + \mathbf{\Delta}\mathbf{x}_1)a_R - (\bar{\mathbf{x}}_2 + \mathbf{\Delta}\mathbf{x}_2)\right] \\ &\approx 2\bar{\mathbf{x}}_1^H \mathbf{W}(a_R)\left(\mathbf{\Delta}\mathbf{x}_1 a_R - \mathbf{\Delta}\mathbf{x}_2\right) \\ &= 2\bar{\mathbf{x}}_1^H \mathbf{W}(a_R)\mathbf{A}\mathbf{\Delta}\mathbf{u}_{R,1} \end{aligned} \tag{A.3}$$

where $\boldsymbol{\Delta}\mathbf{u}_{R,1} = \mathbf{u}_{R,1} - \bar{\mathbf{u}}_{R,1}$. According to (45), the noise matrix under this case is $[\boldsymbol{\mathcal{Q}}]_{(R)}$. From [24], $\boldsymbol{\Delta}\mathbf{u}_{R,1}$ can be approximated as

$$\begin{aligned}\boldsymbol{\Delta}\mathbf{u}_{R,1} &\approx \bar{\lambda}_{R,1}^{-1}\bar{\mathbf{U}}_n\bar{\mathbf{U}}_n^H[\boldsymbol{\mathcal{Q}}]_{(R)}\bar{\mathbf{v}}_{R,1} \\ &= \bar{\lambda}_{R,1}^{-1}\left(\bar{\mathbf{v}}_{R,1}^T \otimes \bar{\mathbf{U}}_n\bar{\mathbf{U}}_n^H\right)\mathrm{vec}([\boldsymbol{\mathcal{Q}}]_{(R)}) \\ &= \bar{\lambda}_{R,1}^{-1}\left(\bar{\mathbf{v}}_{R,1}^T \otimes (\mathbf{I}_M - \bar{\mathbf{u}}_{R,1}\bar{\mathbf{u}}_{R,1}^H)\right)\mathrm{vec}([\boldsymbol{\mathcal{Q}}]_{(R)})\end{aligned} \tag{A.4}$$

where $\bar{\mathbf{U}}_n = \left[\bar{\mathbf{u}}_{R,2}\ \bar{\mathbf{u}}_{R,3}\ \cdots\ \bar{\mathbf{u}}_{R,N_R}\right]$, $N_R = \min\{M_R, M/M_R\}$. For zero-mean $\boldsymbol{\mathcal{Q}}$, $\mathbb{E}\{\boldsymbol{\Delta}\mathbf{u}_{R,1}\}$ approaches a zero vector. Moreover, $\mathbb{E}\left\{\boldsymbol{\Delta}\mathbf{u}_{R,1}\boldsymbol{\Delta}\mathbf{u}_{R,1}^H\right\}$ is:

$$\begin{aligned}\mathbb{E}\left\{\boldsymbol{\Delta}\mathbf{u}_{R,1}\boldsymbol{\Delta}\mathbf{u}_{R,1}^H\right\} &= \bar{\lambda}_{R,1}^{-2}\left(\bar{\mathbf{v}}_{R,1}^T \otimes \bar{\mathbf{U}}_n\bar{\mathbf{U}}_n^H\right)\mathbb{E}\left\{\mathrm{vec}([\boldsymbol{\mathcal{Q}}]_{(R)})\mathrm{vec}([\boldsymbol{\mathcal{Q}}]_{(R)})^H\right\}\left(\bar{\mathbf{v}}_{R,1}^* \otimes \bar{\mathbf{U}}_n\bar{\mathbf{U}}_n^H\right) \\ &= \bar{\lambda}_{R,1}^{-2}\left(\bar{\mathbf{v}}_{R,1}^T \otimes \bar{\mathbf{U}}_n\bar{\mathbf{U}}_n^H\right)\left(\sigma^2\mathbf{I}_M\right)\left(\bar{\mathbf{v}}_{R,1}^* \otimes \bar{\mathbf{U}}_n\bar{\mathbf{U}}_n^H\right) \\ &= \bar{\lambda}_{R,1}^{-2}\sigma^2\left(\bar{\mathbf{v}}_{R,1}^T\bar{\mathbf{v}}_{R,1}^*\right) \otimes \left(\bar{\mathbf{U}}_n\bar{\mathbf{U}}_n^H\bar{\mathbf{U}}_n\bar{\mathbf{U}}_n^H\right) \\ &= \bar{\lambda}_{R,1}^{-2}\sigma^2\bar{\mathbf{U}}_n\bar{\mathbf{U}}_n^H\end{aligned} \tag{A.5}$$

On the other hand, when SNR is sufficiently high, $\mathbb{E}\{J''(a_R)\}$ is

$$\mathbb{E}\{J''(a_R)\} = \mathbb{E}\{2\mathbf{x}_1^H\mathbf{W}(a_R)\mathbf{x}_1\} \approx 2\bar{\mathbf{x}}_1^H\mathbf{W}(a_R)\bar{\mathbf{x}}_1 \tag{A.6}$$

Let $\mathbf{g}_{R1}$ be $\mathbf{g}_R$ in (44) without the last element. Using $\bar{\mathbf{u}}_{R,1} = \mathbf{g}_R e^{-j\varphi_g}/\sqrt{M_R}$ and interchanging $\mathbf{W}(\omega_R)$ and $\mathbf{W}(a_R)$, $\bar{\mathbf{x}}_1^H\mathbf{W}(a_R)\bar{\mathbf{x}}_1$ is calculated as:

$$\begin{aligned}\bar{\mathbf{x}}_1^H\mathbf{W}(a_R)\bar{\mathbf{x}}_1 &= \frac{1}{M_R^2}\mathbf{g}_{R1}^H\mathbf{W}(\omega_R)\mathbf{g}_{R1} \\ &= \frac{1}{M_R^2}\sum_{m=1}^{M_R-1}\sum_{n=1}^{M_R-1} e^{-j\omega_R m}\left(M_R\min(m,n) - mn\right)e^{j(m-n)\omega_R}e^{j\omega_R n} \\ &= \frac{1}{M_R^2}\left(\sum_{m=1}^{M_R-1}(M_R m - m^2) + \sum_{m=1}^{M-1}\sum_{n=1}^{m-1}(M_R n - mn) + \sum_{n=1}^{M_R-1}\sum_{m=1}^{n-1}(M_R m - mn)\right) \\ &= \frac{M_R^2 - 1}{12}\end{aligned} \tag{A.7}$$

Combining (A.6)–(A.7), we get:

$$\mathbb{E}\{J''(a_R)\} \approx \frac{M_R^2 - 1}{6} \tag{A.8}$$

From (A.3) and (A.6)–(A.8), it is clear that $\mathbb{E}\{J'(a_R)\} = 0$ and $\mathbb{E}\{J''(a_R)\} = (M_R^2 - 1)/6$ when SNR tends to infinity. Together with (A.2), we obtain $\lim_{\mathrm{SNR}\to\infty}\mathbb{E}\{\hat{a}_R\} = a$. As $\hat{a}_R = e^{j\hat{\omega}_R}$, $\hat{\omega}_R$ is an

asymptotically unbiased estimate of ω_R. Employing (A.2) again, the variance of $\hat{a}_R$ is:

$$\text{var}(\hat{a}_R) = \mathbb{E}\left\{(\hat{a}_R - a_R)(\hat{a}_R - a_R)^*\right\} = \frac{\mathbb{E}\{J'(a_R)(J'(a_R))^*\}}{[\mathbb{E}\{J''(a_R)\}]^2} \tag{A.9}$$

Using (A.5), $\bar{\mathbf{U}}_n\bar{\mathbf{U}}_n^H = \mathbf{I}_{M_R} - \bar{\mathbf{u}}_{R,1}\bar{\mathbf{u}}_{R,1}^H$, $\mathbf{A}(a_R)\bar{\mathbf{u}}_{R,1} = \mathbf{0}_{(M_R-1)\times 1}$ and $\mathbf{W}(a_R) = \mathbf{W}^H(a_R) = (\mathbf{A}(a_R)\mathbf{A}^H(a_R))^{-1}$, the numerator of (A.9) is:

$$\begin{aligned}
\mathbb{E}\{J'(a_R)(J'(a_R))^*\} &= 4\bar{\mathbf{x}}_1\mathbf{W}(a_R)\mathbf{A}(a_R)\mathbb{E}\{\left\{\mathbf{\Delta}\mathbf{u}_{R,1}\mathbf{\Delta}\mathbf{u}_{R,1}^H\right\}\mathbf{A}^H(a_R)\mathbf{W}^H(a_R)\bar{\mathbf{x}}_1 \\
&= 4\bar{\lambda}_{R,1}^{-2}\sigma^2\bar{\mathbf{x}}_1\mathbf{W}(a_R)\mathbf{A}(a_R)\bar{\mathbf{U}}_n\bar{\mathbf{U}}_n^H\mathbf{A}^H(a_R)\mathbf{W}^H(a_R)\bar{\mathbf{x}}_1 \\
&= 4\bar{\lambda}_{R,1}^{-2}\sigma^2\bar{\mathbf{x}}_1\mathbf{W}(a_R)\mathbf{A}(a_R)(\mathbf{I}_{M_R} - \bar{\mathbf{u}}_{R,1}\bar{\mathbf{u}}_{R,1}^H)\mathbf{A}^H(a_R)\mathbf{W}(a_R)\bar{\mathbf{x}}_1 \\
&= 4\bar{\lambda}_{R,1}^{-2}\sigma^2\bar{\mathbf{x}}_1\mathbf{W}(a_R)\mathbf{W}^{-1}(a_R)\mathbf{W}(a_R)\bar{\mathbf{x}}_1 \\
&= \frac{4\bar{\lambda}_{R,1}^{-2}\sigma^2(M_R^2-1)}{12}
\end{aligned} \tag{A.10}$$

Substituting (A.8) and (A.10) with $\bar{\lambda}_{R,1} = \sqrt{M}|\lambda|$ into (A.9) and $\text{SNR} = |\gamma|^2/\sigma^2$ yields:

$$\text{var}(\hat{a}_R) \approx \frac{12\sigma^2}{M(M_R^2-1)|\gamma|^2} = \frac{12}{M(M_R^2-1)\text{SNR}} \tag{A.11}$$

Employing the transformation formula of $\text{var}(\hat{w}_R) \approx \text{var}\left(\hat{a}_R\right)/(2|a_R|^2)$ [32], we eventually get (56) for $r = R$.

According to (54) and the definition of Rth unfolding, the elements of the $(R-1)$ dimensional tensor $\boldsymbol{\mathcal{Z}}$ can be written as:

$$z_{m_1,m_2,\cdots,m_{R-1}} = \hat{\mathbf{g}}_R^\dagger\left(\gamma\prod_{r=1}^{R-1}e^{j\omega_r m_r}\mathbf{g}_R + \mathbf{q}^{(m_1,m_2,\cdots,m_{R-1})}\right) \tag{A.12}$$

where $\gamma\prod_{r=1}^{R-1}e^{j\omega_r m_r}\hat{\mathbf{g}}_R^\dagger\mathbf{g}_R$ and $\hat{\mathbf{g}}_R^\dagger\mathbf{q}^{(m_1,m_2,\cdots,m_{R-1})}$ are the signal and noise components with

$$\mathbf{q}^{(m_1,m_2,\cdots,m_{R-1})} = \begin{bmatrix} q_{m_1,m_2,\cdots,m_{R-1},1} & q_{m_1,m_2,\cdots,m_{R-1},2} & \cdots & q_{m_1,m_2,\cdots,m_{R-1},M_R}\end{bmatrix}^T \tag{A.13}$$

For zero-mean white Gaussian disturbance, we have

$$\mathbb{E}\{q_{m_1,m_2,\cdots,m_{R-1},m_R}\cdot q_{m_1,m_2,\cdots,m_{R-1},m_R}^H\} = \sigma^2 \tag{A.14}$$

and

$$\mathbb{E}\{q_{m_1,m_2,\cdots,m_{R-1},m_R}\cdot q_{n_1,n_2,\cdots,n_{R-1},n_R}^H\} = 0, \quad n_r = 1,2,\cdots,M_r \tag{A.15}$$

if $m_r \neq n_r$ for at least one of $r = 1,2,\cdots,R$. Hence $\mathbf{q}^{(m_1,m_2,\cdots,m_{R-1})}$ has the following properties:

$$\mathbb{E}\{\left(\mathbf{q}^{(m_1,m_2,\cdots,m_{R-1})}\right)\left(\mathbf{q}^{(m_1,m_2,\cdots,m_{R-1})}\right)^H\} = \mathbf{I}_{M_R}\sigma^2 \tag{A.16}$$

and

$$\mathbb{E}\{\left(\mathbf{q}^{(m_1,m_2,\cdots,m_{R-1})}\right)\left(\mathbf{q}^{(n_1,n_2,\cdots,n_{R-1})}\right)^H\} = \mathbf{0}_{M_R} \tag{A.17}$$

if $m_r \neq n_r$ for at least one of $r = 1, 2, \cdots, R-1$. Assuming that $\hat{\mathbf{g}}_R^\dagger$ and $\mathbf{q}^{(m_1, m_2, \cdots, m_{R-1})}$ are uncorrelated, we then have:

$$\mathbb{E}\{\hat{\mathbf{g}}_R^\dagger \mathbf{q}^{(m_1, m_2, \cdots, m_{R-1})}\} = \mathbb{E}\{\hat{\mathbf{g}}_R^\dagger\}\mathbb{E}\{\mathbf{q}^{(m_1, m_2, \cdots, m_{R-1})}\} = 0 \tag{A.18}$$

$$\begin{aligned}\mathbb{E}\{\left(\hat{\mathbf{g}}_R^\dagger \mathbf{q}^{(m_1, m_2, \cdots, m_{R-1})}\right)^2\} &= \mathbb{E}\{\left(\hat{\mathbf{g}}_R^\dagger \mathbf{q}^{(m_1, m_2, \cdots, m_{R-1})}\right)\left(\hat{\mathbf{g}}_R^\dagger \mathbf{q}^{(m_1, m_2, \cdots, m_{R-1})}\right)^H\} \\ &= \mathbb{E}\{\hat{\mathbf{g}}_R^\dagger \mathbb{E}\{\left(\mathbf{q}^{(m_1, m_2, \cdots, m_{R-1})}\right)\left(\mathbf{q}^{(m_1, m_2, \cdots, m_{R-1})}\right)^H\}\hat{\mathbf{g}}_R^{\dagger H}\} \\ &= M_R\sigma^2 \end{aligned} \tag{A.19}$$

and

$$\mathbb{E}\{\left(\hat{\mathbf{g}}_R^\dagger \mathbf{q}^{(n_1, n_2, \cdots, n_{R-1})}\right)\left(\hat{\mathbf{g}}_R^\dagger \mathbf{q}^{(m_1, m_2, \cdots, m_{R-1})}\right)^H\} = 0 \tag{A.20}$$

where we see that the noise component in $\boldsymbol{\mathcal{Z}}$ is also an uncorrelated zero mean process with power of $M_R\sigma^2$. On the other hand, by substituting $\hat{\mathbf{g}}_R = \mathbf{g}_R$ which is valid for sufficiently high SNR conditions, the signal component in $\boldsymbol{\mathcal{Z}}$ can be approximated as:

$$\gamma \prod_{r=1}^{R-1} e^{j\omega_r m_r} \mathbf{g}_R^\dagger \mathbf{g}_R \approx M_R \gamma \prod_{r=1}^{R-1} e^{j\omega_r m_r} \tag{A.21}$$

which has a power of $M_R^2|\gamma|^2$. That is, the SNR in $\boldsymbol{\mathcal{Z}}$ is $\text{SNR}_{\boldsymbol{\mathcal{Z}}} = (M_R^2|\gamma|^2)/(M_R\sigma^2) = M_R\text{SNR}$. As (A.12) is analogous to (42), we can apply the development in (A.1)–(A.11) to analyze $\hat{\omega}_r$, $r = 1, 2, \cdots, R-1$. As a result, $\{\hat{\omega}_r\}$ are asymptotically unbiased estimates of $\{\omega_r\}$. Moreover, from (A.11), we have:

$$\text{var}(\hat{a}_r) \approx \frac{12}{\text{SNR}_{\boldsymbol{\mathcal{Z}}}(M/M_R)(M_r^2 - 1)} = \frac{12\sigma^2}{M(M_r^2 - 1)|\gamma|^2}, \quad r = 1, 2, \cdots, R-1 \tag{A.22}$$

Applying $\text{var}(\hat{w}_r) \approx \text{var}\,(\hat{a}_r)/(2|a_r|^2)$ again, we obtain (56) for the remaining frequency estimates.

APPENDIX B

The CRLB for a single undamped R-D cisoid in the presence of white Gaussian noise is derived as follows. First we define $\gamma = be^{j\theta}$ where $b = |\gamma|$ is the magnitude and θ is the phase. Decomposing $\boldsymbol{\mathcal{Y}}$ as real and imaginary components as $\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{Y}}^{\text{real}} + j\boldsymbol{\mathcal{Y}}^{\text{imag}}$, the joint probability density function for $\boldsymbol{\mathcal{Y}}$ with the unknown parameter vector $\boldsymbol{\Phi} = \begin{bmatrix}\omega_1 & \omega_2 & \cdots & \omega_R & \theta & b\end{bmatrix}^T$ is [31]:

$$\begin{aligned} f\left(\boldsymbol{\mathcal{Y}}; \boldsymbol{\Phi}\right) = \left(\frac{1}{\sigma^2\pi}\right)^M \exp\Bigg\{ -\frac{1}{\sigma^2} \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} \cdots \sum_{m_R=1}^{M_R} \\ \left[\left(y_{m_1, m_2, \cdots, m_R}^{\text{real}} - \mu_{m_1, m_2, \cdots, m_R}\right)^2 + \left(y_{m_1, m_2, \cdots, m_R}^{\text{imag}} - \nu_{m_1, m_2, \cdots, m_R}\right)^2\right]\Bigg\} \end{aligned} \tag{B.1}$$

where $\mu_{m_1,m_2,\cdots,m_R} = b\cos\left(\sum_{r=1}^{R} m_r\omega_r + \theta\right)$ and $\nu_{m_1,m_2,\cdots,m_R} = b\sin\left(\sum_{r=1}^{R} m_r\omega_r + \theta\right)$. Using (B.1), the (i,j) entry, $i,j = 1,2,\cdots,R+2$, of the corresponding Fisher information matrix, denoted by $\mathbf{J}$, are:

$$\begin{aligned}[\mathbf{J}]_{i,j} &= \mathbb{E}\{\frac{\partial}{\partial\left[\mathbf{\Phi}\right]_i}\log f\left(\boldsymbol{\mathcal{Y}};\mathbf{\Phi}\right)\frac{\partial}{\partial\left[\mathbf{\Phi}\right]_j}\log f\left(\boldsymbol{\mathcal{Y}};\mathbf{\Phi}\right)\} \\ &= \frac{2}{\sigma^2}\sum_{m_1=1}^{M_1}\sum_{m_2=1}^{M_2}\cdots\sum_{m_R=1}^{M_R}\left[\frac{\partial\mu_{m_1,m_2,\cdots,m_R}}{\partial\left[\Phi\right]_i}\frac{\partial\mu_{m_1,m_2,\cdots,m_R}}{\partial\left[\Phi\right]_j}+\frac{\partial\nu_{m_1,m_2,\cdots,m_R}}{\partial\left[\Phi\right]_i}\frac{\partial\nu_{m_1,m_2,\cdots,m_R}}{\partial\left[\Phi\right]_j}\right] \end{aligned} \quad \text{(B.2)}$$

After some manipulations, it can be shown that

$$\mathbf{J} = \frac{2M}{\sigma^2}\begin{bmatrix}\begin{bmatrix}\mathbf{A}+\mathbf{B}\mathbf{B}^T/|\gamma|^2 & \mathbf{B}\\ \mathbf{B}^T & b^2\end{bmatrix} & \mathbf{0}_{(R+1)\times 1}\\ \mathbf{0}_{1\times(R+1)} & 1\end{bmatrix} \quad \text{(B.3)}$$

where

$$\mathbf{A} = \frac{|\gamma|^2}{12}\mathrm{diag}\left(\begin{bmatrix}M_1^2-1 & M_2^2-1 & \cdots & M_R^2-1\end{bmatrix}\right) \quad \text{(B.4)}$$

$$\mathbf{B} = \frac{|\gamma|^2}{2}\begin{bmatrix}M_1+1 & M_2+1 & \cdots & M_R+1\end{bmatrix}^T \quad \text{(B.5)}$$

Noting that the CRLB for frequencies, denoted by CRLB_{ω_r}, is given as $\left[\mathbf{J}^{-1}\right]_{r,r}$, $r = 1,2,\cdots,R$, and applying the matrix inversion lemma, we have:

$$\mathrm{CRLB}_{\omega_r} = \frac{\sigma^2}{2M}\left[(\mathbf{A}+\frac{1}{|\gamma|^2}\mathbf{B}\mathbf{B}^T-\frac{1}{|\gamma|^2}\mathbf{B}\mathbf{B}^T)^{-1}\right]_{r,r} = \frac{\sigma^2}{2M}\left[\mathbf{A}^{-1}\right]_{r,r} = \frac{6\sigma^2}{M(M_r^2-1)|\gamma|^2} \quad \text{(B.6)}$$

ACKNOWLEDGEMENTS

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China [Project No. CityU 120911].

REFERENCES

[1] 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, Feb. 1998

[2] X. Liu, N.D. Sidiropoulos and T. Jiang, "Multidimensional harmonic retrieval with applications in MIMO wireless channel sounding," in *Space-Time Processing for MIMO Communications*, A.B. Gershman and N.D. Sidiropoulos, Eds., Wiley, 2005

[3] 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, Jan. 1998

[4] D. Nion and N.D. Sidiropoulos, "Tensor algebra and multidimensional harmonic retrieval in signal processing for MIMO radar," *IEEE Transactions on Signal Processing*, vol.58, no.11, pp. 5693-5705, Nov. 2010

[5] P. Stoica and R. Moses, *Spectral Analysis of Signals*, Upper Saddle River, NJ: Prentice-Hall, 2005

[6] C.R. Rao, L.C. Zhao and B. Zhou, "Maximum likelihood of 2-D superimposed exponential signals," *IEEE Transactions on Signal Processing*, vol.42, no.7, pp.1795-1802, Jul. 1994

[7] 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, Jun. 1994

[8] M. Haardt and J.A. Nossek, ”Unitary ESPRIT: How to obtain increased estimation accuracy with a reduced computational burden,” *IEEE Transactions on Signal Processing*, vol.43, no.5, pp.1232-1242, May 1995

[9] M. Haardt and J.A. Nossek, ”Simultaneous Schur decomposition of several nonsymmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems,” *IEEE Transactions on Signal Processing*, vol.46, no.1, pp.161-169, Jan. 1998

[10] M. Haardt, F. Roemer and G. Del Galdo, “Higher-order SVD-based subspace estimation to improve the parameter estimation accuracy in multidimensional harmonic retrieval problems,” *IEEE Transactions on Signal Processing*, vol.56, no.7, pp. 3198-3213, Jul. 2008

[11] A. Thakre, M. Haardt and K. Giridhar, “Single snapshot R-D unitary tensor-ESPRIT using an augmentation of the tensor order,” *Proc. IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, pp.81–84, Dec. 2009, Aruba, Dutch Antilles

[12] X. Liu and N.D. Sidiropoulos, “Almost sure identifiability of multidimensional constant modulus harmonic retrieval,” *IEEE Transactions on Signal Processing*, vol.50, no.9, pp.2366-2368, Sep. 2002

[13] J. Liu and X. Liu, “An eigenvector-based approach for multidimensional frequency estimation with improved identifiability,” *IEEE Transactions on Signal Processing*, vol.54, no.12, pp.4543-4556, Dec. 2006

[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, Apr. 2010

[15] L. de Lathauwer, B. de Moor, and J. Vanderwalle, “A multilinear singular value decomposition,” *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253-1278, 2000

[16] T.G. Kolda and B.W. Bader, “Tensor decompositions and applications,” *SIAM Rev.*, vol.51, no.3, pp.455-500, Sep. 2009

[17] A. Smilde, R. Bro and P. Geladi, *Multi-Way Analysis: Applications in the Chemical Sciences*, Wiley, West Sussex, England, 2004

[18] N.D. Sidiropoulos, ”Generalizing Caratheodory’s uniqueness of harmonic parameterization to N dimensions,” *IEEE Trans. Inf. Theory*, vol.47, no.4, pp.1687-1690, May 2001

[19] J.-M. Papy, L. de Lathauwer, S. van Huffel, “Exponential data fitting using multilinear algebra: The single-channel and multi-channel case,” *Num. Lin. Alg. Appl.*, vol.12, no.8, pp.809–826, Oct. 2005

[20] T. Jiang, N. D. Sidiropoulos and J. M. F. Ten Berge, “Almost-sure identifiability of multidimensional harmonic retrieval,” *IEEE Transactions on Signal Processing*, vol.49, pp.1849-1859, Sep. 2001

[21] S.M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993

[22] L. de Lathauwer, B. de Moor and J. Vanderwalle, “On the best rank-1 and rank- $(r_1, r_2, \cdots, r_n)$ approximation of higher-order tensors,” *SIAM J. Matrix Anal. Appl.*, vol.21, no.4, pp.1324–1342, 2000

[23] 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, Jun. 2010

[24] 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, Jul. 2008

[25] J. Li, P. Stoica and Z.S. Liu, “Comparative study of IQML and MODE direction-of-arrival estimators,” *IEEE Trans. on Signal Processing*, vol.46, no.1, pp.149-160, Jan. 1998

[26] J. H. McClellan and D. Lee, "Exact equivalence of the Steiglitz-McBride iteration and IQML," *IEEE Trans. on Signal Processing*, vol.39, no.2, pp.509-512, Feb. 1991

[27] P. Stoica and T. Söderström, "The Steiglitz-McBride identification algorithm revisited – Convergence analysis and accuracy aspects," *IEEE Trans. on Signal Processing*, vol.26, no.3, pp.712-717, Jun. 1981

[28] W. Sun and H.C. So, "Improvement to ESPRIT-type frequency estimators via reducing data redundancy," *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, pp.4192-4195, May 2011, Prague, Czech Republic

[29] G. H. Golub and C. F. van Loan, *Matrix Computations*, Baltimore, MD: John Hopkins University Press, 1996

[30] 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, Apr. 2006

[31] D. Rife and R. Boorstyn, "Single tone parameter estimation from discrete-time observations," *IEEE Transactions on Information Theory*, vol.20, no.5, pp.591-598, Sep. 1974

[32] Y.-X. Yao and S. M. Pandit, "Variance of least squares estimators for a damped sinusoidal process," *IEEE Transactions on Signal Processing*, vol.42, no.11, pp.3016-3025, Nov. 1994

Weize Sun received the B.S. degree in Electronic Information Science and Technology from SUN YAT-SEN University, China, in 2005. He is currently a Research Student and working towards his PhD degree in City University of Hong Kong. His research interests include statistical signal processing, parameter estimation, tensor algebra, with particular attention to frequency estimation.

H. C. So was born in Hong Kong. He obtained the B.Eng. degree from City University of Hong Kong and the Ph.D. degree from The Chinese University of Hong Kong, both in electronic engineering, in 1990 and 1995, respectively. From 1990 to 1991, he was an Electronic Engineer at the Research and Development Division of Everex Systems Engineering Ltd., Hong Kong. During 1995-1996, he worked as a Post-Doctoral Fellow at The Chinese University of Hong Kong. From 1996 to 1999, he was a Research Assistant Professor at the Department of Electronic Engineering, City University of Hong Kong, where he is currently an Associate Professor. His research interests include statistical signal processing, fast and adaptive algorithms, signal detection, parameter estimation, and source localization. He has been on the editorial boards of IEEE Transactions on Signal Processing, Signal Processing, Digital Signal Processing and ISRN Applied Mathematics as well as a member in Signal Processing Theory and Methods Technical Committee of the IEEE Signal Processing Society.

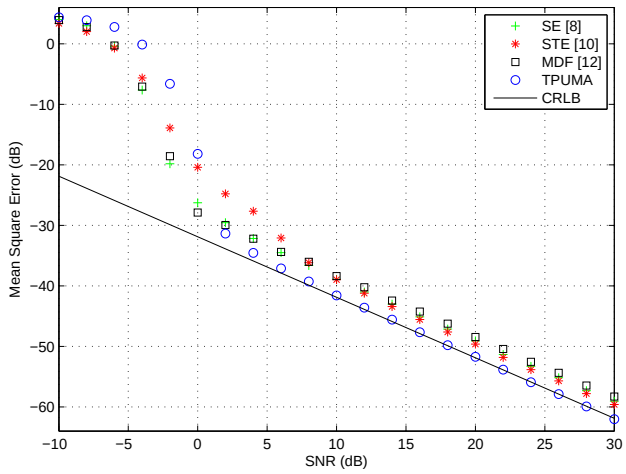

Fig. 1. Average mean square frequency error versus SNR for 3-D damped tones in first case

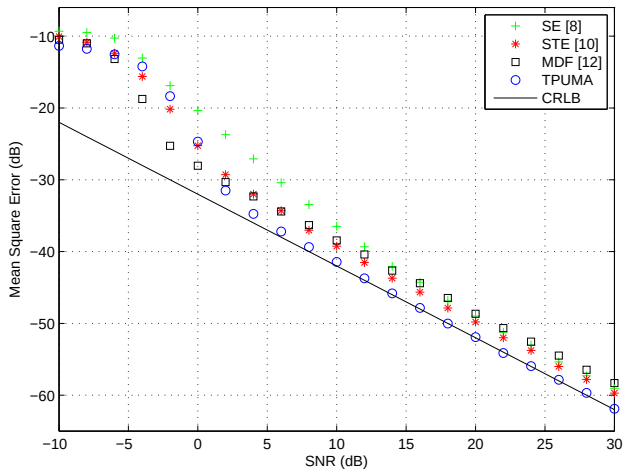

Fig. 2. Average mean square damping factor error versus SNR for 3-D damped tones in first case

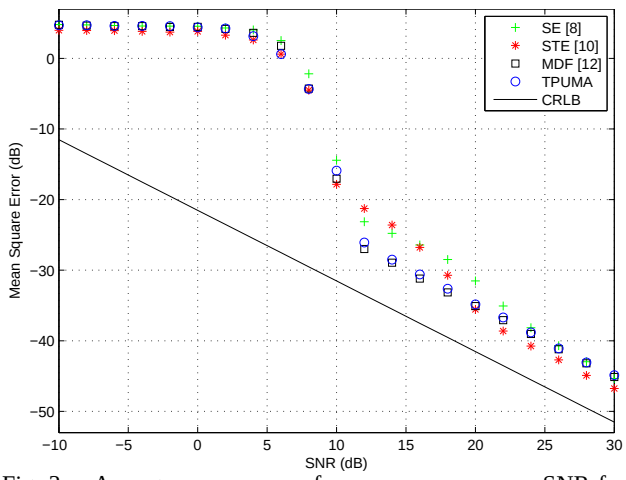

Fig. 3. Average mean square frequency error versus SNR for 3-D damped tones in second case

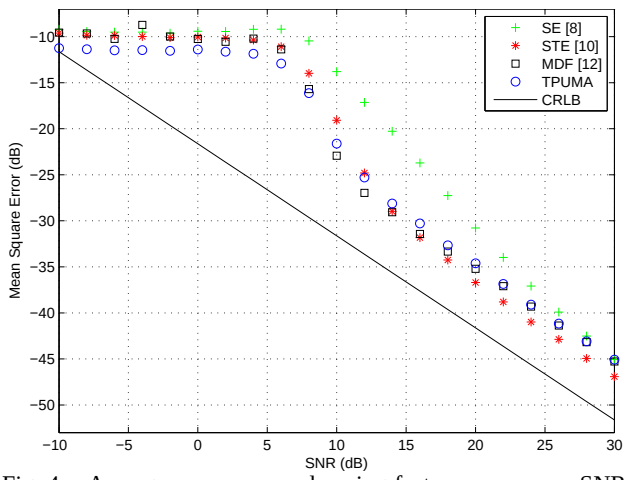

Fig. 4. Average mean square damping factor error versus SNR for 3-D damped tones in second case

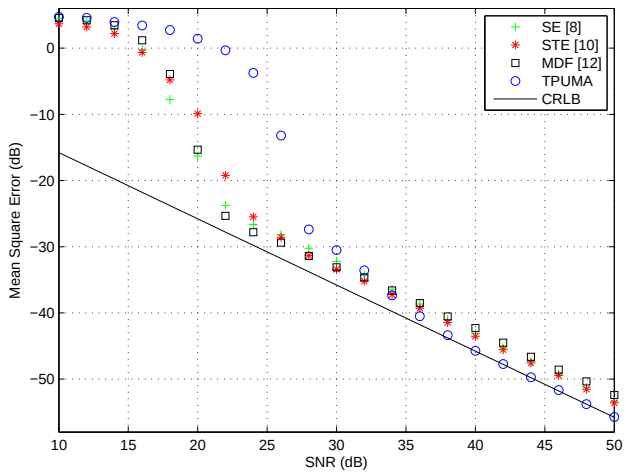

Fig. 5. Average mean square frequency error versus SNR for 3-D damped tones in third case

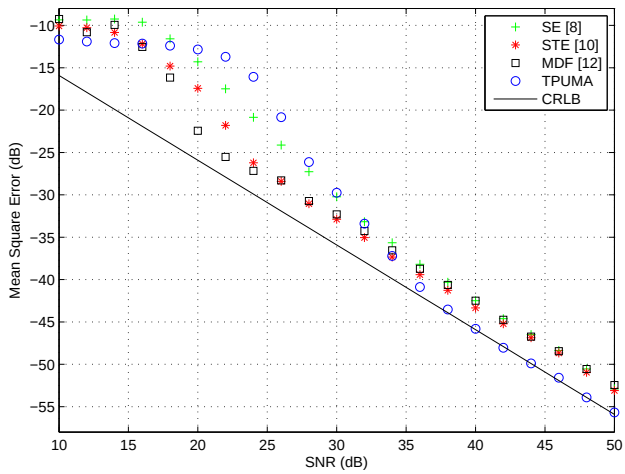

Fig. 6. Average mean square damping factor error versus SNR for 3-D damped tones in third case

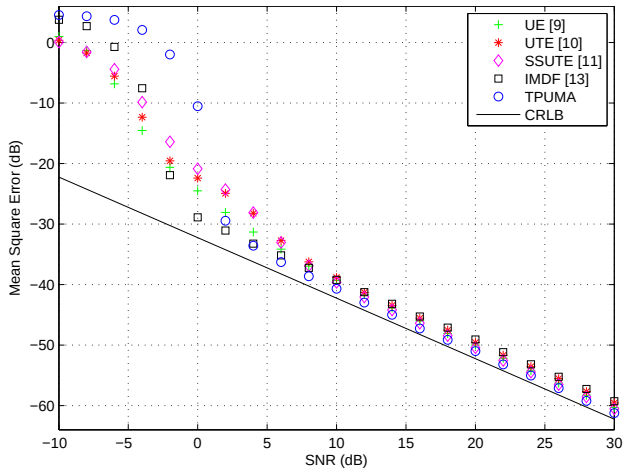

Fig. 7. Average mean square frequency error versus SNR for 3-D undamped tones

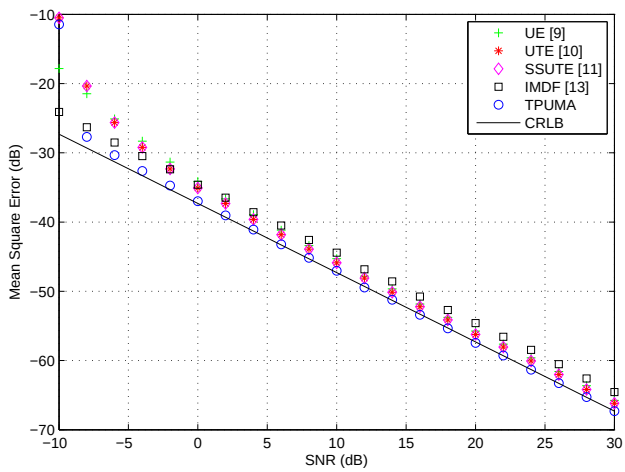

Fig. 8. Average mean square frequency error versus SNR for single 3-D undamped tone

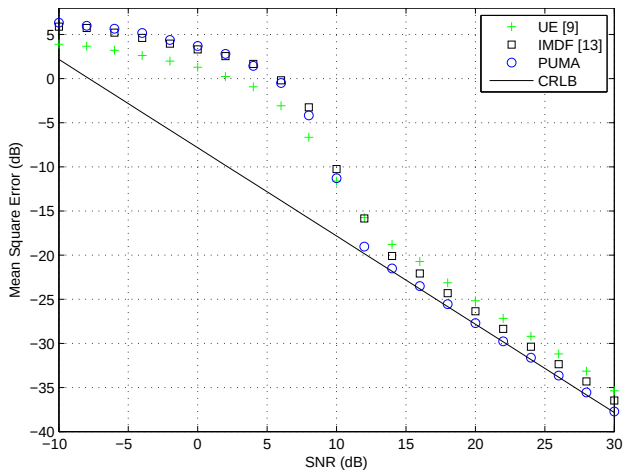

Fig. 9. Average mean square frequency error versus SNR for $2 \times 2 \times 10$ data set

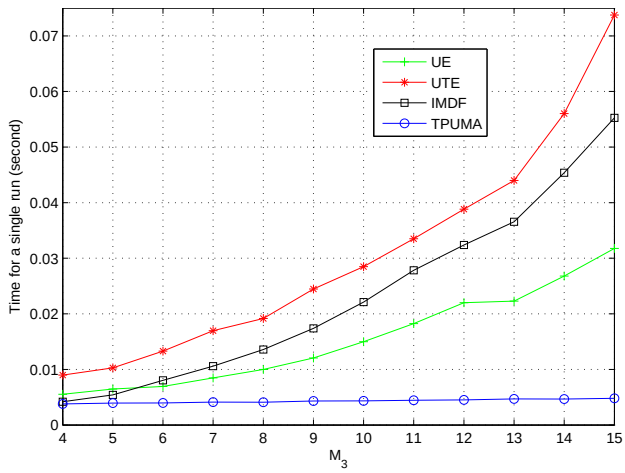

Fig. 10. Average computation time for a single run under $M_1 = M_2 = 6$ and $F = 3$