EFFICIENT PARAMETER ESTIMATION OF MULTIPLE DAMPED SINUSOIDS BY COMBINING SUBSPACE AND WEIGHTED LEAST SQUARES TECHNIQUES

Weize Sun and H.C. So

Department of Electronic Engineering, City University of Hong Kong, Hong Kong SAR, China

weizesun2@student.cityu.edu.hk, hcso@ee.cityu.edu.hk

ABSTRACT

A new signal subspace approach for sinusoidal parameter estimation of multiple tones is proposed in this paper. Our main ideas are to arrange the observed data into a matrix without reuse of elements and exploit the principal singular vectors of this matrix for parameter estimation. Comparing with the conventional subspace methods which employ Hankel-style matrices with redundant entries, the proposed approach is more computationally efficient. Computer simulations are also included to compare the proposed methodology with the weighted least squares and ESPRIT approaches in terms of estimation accuracy and computational complexity.

Index Terms— frequency estimation, subspace method, singular value decomposition, linear prediction, weighted least squares

1. INTRODUCTION

Estimating the parameters of sinusoidal signals from noisy observations is an important research topic in science and engineering. The crucial step is to find the damping factors and frequencies which are nonlinear functions in the observed data. Once they have been estimated, computation of the remaining parameters reduces to a least squares (LS) fit.

Generally speaking, parameter estimation can be achieved by either nonparametric or parametric methodologies [1]. In most of the cases, the parametric approach, which assumes that the signal satisfies a generating model with known functional form, will have a higher resolution than the nonparametric ones. In estimating multiple nonlinear parameters, the maximum-likelihood based methodology requires extensive computations for a multi-dimensional search, which may not be suitable in many applications. The subspace methods that separate the data into signal and noise subspaces via eigenvalue decomposition of the sample covariance matrix or the singular value decomposition (SVD) of the raw data matrix, such as MUSIC [2] and ESPRIT [3], can achieve a high resolution with a moderate complexity. On the other hand, linear prediction (LP) approach, which includes the weighted least square (WLS) estimator [4], can attain optimum performance when signal-to-noise ratio (SNR) is sufficiently high. In this work, we propose to combine subspace and WLS techniques to achieve sinusoidal parameter estimation with low computational complexity and high accuracy.

The rest of this paper is organized as follows. The notation and formulation for sinusoidal parameter estimation are given in Section 2, then the proposed estimator is derived in Section 3. In Section 4, simulation results are included to evaluate the performance of the developed approach by comparing with the WLS [4] and ESPRIT [5] algorithms, as well as Cramér-Rao lower bound (CRLB). Finally, conclusions are drawn in Section 5.

The work described in this paper was fully supported by a grant from CityU (Project no. 7002570).

2. NOTATION AND DATA MODEL

Throughout this paper, bold upper/lower case symbols denote matrices/vectors. The $N_1 \times N_2$ zero matrix, $N \times N$ identity matrix and Kronecker product are represented by $\mathbf{0}_{N_1 \times N_2}$, $\mathbf{I}_N$ and $\otimes$, and superscripts ^{T}, ^{H}, *, $^{-1}$ and † denote transpose, Hermitian transpose, complex conjugation, matrix inversion and pseudo-inverse, respectively. Moreover, we use $\widetilde{\mathbf{A}}$ and $\widehat{\mathbf{A}}$ to represent the noise-free counterpart and estimate of $\mathbf{A}$.

The observed damped sinusoidal signal is:

$$x_n = s_n + \xi_n, \quad n = 1, 2, \cdots, N \tag{1}$$

where

$$s_n = \sum_{k=1}^{K} \gamma_k \alpha_k^n e^{j\omega_k n}, \quad k = 1, 2, \cdots, K \tag{2}$$

The γ_k, $\alpha_k \in (0, 1]$, $\omega_k \in [-\pi, \pi)$ are the complex amplitudes, damping factors and frequencies while $\{\xi_n\}$ are zero mean complex white Gaussian noises with unknown variance σ^2. The number of sinusoids, denoted by K, is assumed known $a\ priori$. It is also assumed that N, the length of the data, can be factorized as $N = N_1 N_2$, where $N_1 > K$ and N_2 are integers. Note that even if N is not factorizable, we can simply discard a few samples and find N_1 and N_2 such that their product is closest to N, and the performance loss will be negligible for a sufficiently large data length. Under this definition, stacking x_n into a matrix $\mathbf{X} \in \mathbb{C}^{N_1 \times N_2}$ yields:

$$\mathbf{X} = \mathbf{S} + \mathbf{\Xi} \tag{3}$$

where
$$\mathbf{S} = \begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \cdots & \mathbf{s}_{N_2} \end{bmatrix} \tag{4}$$

and $\mathbf{s}_{n_2} = \begin{bmatrix} \mathbf{s}_{(n_2-1)N_1+1} & \mathbf{s}_{(n_2-1)N_1+2} & \cdots & \mathbf{s}_{n_2 N_1} \end{bmatrix}$ for $n_2 = 1, 2, \cdots, N_2$ while $\mathbf{X}$ and $\mathbf{\Xi}$ contain $\{x_n\}$ and $\{\xi_n\}$ accordingly.

3. PROPOSED ESTIMATOR

Following [7], $\mathbf{S}$ can be factorized as:

$$\mathbf{S} = \mathbf{G}\mathbf{\Gamma}\mathbf{H}^T \tag{5}$$

where
$$\mathbf{\Gamma} = \mathrm{diag}(\begin{bmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_K \end{bmatrix}) \tag{6}$$

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \cdots & \mathbf{g}_K \end{bmatrix} \tag{7}$$

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \cdots & \mathbf{h}_K \end{bmatrix} \tag{8}$$

$$\mathbf{g}_k = \begin{bmatrix} g_k & g_k^2 & \cdots g_k^{N_1} \end{bmatrix}^T \tag{9}$$

$$\mathbf{h}_k = \begin{bmatrix} h_k & h_k^2 & \cdots h_k^{N_2} \end{bmatrix}^T \tag{10}$$

$$g_k = \alpha_{L,k} e^{j\omega_{L,k}} \quad \text{and} \quad h_k = \beta_k e^{j\mu_k} \tag{11}$$

Apparently, we have $\alpha_{L,k} = \alpha_k$, $\omega_{L,k} = \omega_k$, $\beta_k = \alpha_k^{N_1}$ and $\mu_k = (N_1\omega_k) \mod (2\pi)$ for all k. On the other hand, decomposing $\mathbf{X}$ using SVD gives:

$$\mathbf{X} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H = \begin{bmatrix} \mathbf{U}_s & \mathbf{U}_n \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_n \end{bmatrix} \begin{bmatrix} \mathbf{V}_s & \mathbf{V}_n \end{bmatrix}^H \tag{12}$$

where $\mathbf{U}_s \in \mathbb{C}^{N_1 \times l_K}$, $\mathbf{\Lambda}_s \in \mathbb{C}^{l_K \times l_K}$ and $\mathbf{V}_s \in \mathbb{C}^{N_2 \times l_K}$, $l_K = \min\{N_2, K\}$, are the signal subspace components. According to the decomposition in (5)–(11), the best rank-l_K approximation of $\mathbf{S}$ according to (12), denoted by $\widehat{\mathbf{S}}$, is

$$\widehat{\mathbf{S}} = \mathbf{U}_s \mathbf{\Lambda}_s \mathbf{V}_s^H \tag{13}$$

As span$(\widetilde{\mathbf{U}}_s) \subseteq$ span$(\mathbf{G})$, we have

$$\widetilde{\mathbf{U}}_s = \mathbf{G}\mathbf{\Omega}_G \tag{14}$$

where $\mathbf{\Omega}_G$ is an unknown $K \times l_K$ matrix. Equation (14) shows that each column of $\widetilde{\mathbf{U}}_s$, namely, $\tilde{\mathbf{u}}_k$, $k = 1, 2, \cdots, l_K$, is a sum of K damped cisoids such that the frequencies and damping factors in $\{\tilde{\mathbf{u}}_k\}$ are identical but having different amplitudes, which corresponds to a multi-channel spectral estimation problem [6]. For each $\tilde{\mathbf{u}}_k$, we have the following linear prediction (LP) property:

$$\sum_{i=0}^{K} c_i [\tilde{\mathbf{u}}_k]_{n_1 - i} = 0 \tag{15}$$

for $k = 1, 2, \cdots, l_K$, $n_1 = K+1, \cdots, N_1$ with $c_0 = 1$ and $\mathbf{c} = \begin{bmatrix} c_1 & c_2 & \cdots & c_K \end{bmatrix}^T$ being the LP coefficient vector. By finding the roots of

$$\sum_{i=0}^{K} c_i z^{K-i} = 0 \tag{16}$$

says, $\hat{g}_k$, $k = 1, 2, \ldots, K$. The frequency and damping factor estimates $\hat{\omega}_{L,k}$ and $\hat{\alpha}_{L,k}$ are:

$$\hat{\omega}_{L,k} = \angle(\hat{g}_k) \quad \text{and} \quad \hat{\alpha}_{L,k} = |\hat{g}_k| \tag{17}$$

Then the problem is reduced to finding the LP coefficient vector $\mathbf{c}$. By constructing the LP error vector $\mathbf{e} = \mathbf{D}\mathbf{c} - \mathbf{f}$, $\mathbf{c}$ can be solved by the WLS technique [6]:

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \mathbf{e}^H \mathbf{W} \mathbf{e} = (\mathbf{D}^H \mathbf{W} \mathbf{D})^{-1} \mathbf{D}^H \mathbf{W} \mathbf{f} \tag{18}$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1^T & \mathbf{D}_2^T & \cdots & \mathbf{D}_{l_K}^T \end{bmatrix}^T \tag{19}$$

$$\mathbf{f} = \begin{bmatrix} \mathbf{f}_1^T & \mathbf{f}_2^T & \cdots & \mathbf{f}_{l_K}^T \end{bmatrix}^T \tag{20}$$

$$\mathbf{D}_k = \mathrm{Toeplitz}\Big(\begin{bmatrix} [\mathbf{u}_k]_K & [\mathbf{u}_k]_{K+1} & \cdots & [\mathbf{u}_k]_{N_1-1} \end{bmatrix}^T, \begin{bmatrix} [\mathbf{u}_k]_K & [\mathbf{u}_k]_{K-1} & \cdots & [\mathbf{u}_k]_1 \end{bmatrix} \Big) \tag{21}$$

$$\mathbf{f}_k = -\begin{bmatrix} [\mathbf{u}_k]_{K+1} & [\mathbf{u}_k]_{K+2} & \cdots & [\mathbf{u}_k]_{N_1} \end{bmatrix}^T \tag{22}$$

where $\mathbf{W}$ is a symmetric weighting matrix. Define $\mathbf{A} = \mathrm{Toeplitz}\Big(\begin{bmatrix} c_K & \mathbf{0}_{1 \times (N_1-K-1)} \end{bmatrix}^T, \begin{bmatrix} \mathbf{c}_{\prod} & \mathbf{0}_{1 \times (N_1-K-1)} \end{bmatrix} \Big)$ where $\mathbf{c}_{\prod} = \begin{bmatrix} c_K & c_{K-1} & \cdots & c_1 & 1 \end{bmatrix}$ and $\mathbf{U}_s = \widetilde{\mathbf{U}}_s + \mathbf{\Delta U}_s$, we have the fact that $\mathbf{A}\widetilde{\mathbf{U}}_s = \mathbf{0}_{(N_1-K) \times N_1}$. It is shown in [8] that following the Gauss-Markov theorem [9], the optimal $\mathbf{W}$ is:

$$\begin{aligned} \mathbf{W} &= \sigma^2 \left[\mathbb{E}\left\{ \mathbf{e}\mathbf{e}^H \right\} \right]^{-1} \\ &= \sigma^2 \left[\mathbb{E}\left\{ \mathrm{vec}(\mathbf{A}\mathbf{U}_s)\mathrm{vec}(\mathbf{A}\mathbf{U}_s)^H \right\} \right]^{-1} \\ &= \sigma^2 \left[\mathbb{E}\left\{ \mathrm{vec}(\mathbf{A}\mathbf{\Delta U}_s)\mathrm{vec}(\mathbf{A}\mathbf{\Delta U}_s)^H \right\} \right]^{-1} \\ &= \mathrm{diag}(\tilde{\lambda}_1^2, \tilde{\lambda}_2^2, \cdots, \tilde{\lambda}_K^2) \otimes (\mathbf{A}\mathbf{A}^H)^{-1} \\ &\approx \mathrm{diag}(\lambda_1^2, \lambda_2^2, \cdots, \lambda_K^2) \otimes (\mathbf{A}\mathbf{A}^H)^{-1} \end{aligned} \tag{23}$$

by approximating $\tilde{\lambda}_k$ using λ_k. The estimation of $\hat{\mathbf{c}}$ is done by an iterative procedure of (18) and (23) with an initial choice of $\mathbf{W} = \mathrm{diag}(\lambda_1^2, \lambda_2^2, \cdots, \lambda_K^2) \otimes \mathbf{I}_{N_1}$. According to the definition, we can simply set $\hat{\omega}_k = \hat{\omega}_{L,k}$ and $\hat{\alpha}_k = \hat{\alpha}_{L,k}$ and claim that the work is done. As it is shown in Section 4, setting (17) as the final estimates will not be accurate enough, while a better way is to extract the estimates from $\hat{\mu}_k$ and $\hat{\beta}_k$ [7], which can be obtained by applying (14) to (23) to $\mathbf{V}_s^*$. However, this will lead to two problems. First, although the ω_k's are identical naturally, μ_k's might not be identical as $\mu_k = (N_1\omega_k) \mod (2\pi)$. Second, an extra pairing step between ω_k's and μ_k's is needed, which is costly. In order to overcome these problems, we employ another procedure for estimating μ_k's and β_k's as follows.

From (5), we have

$$\mathbf{X} \approx \widehat{\mathbf{G}}\boldsymbol{\mathcal{H}}^T \tag{24}$$

where $\widehat{\mathbf{G}}$ is the estimate of $\mathbf{G}$ which is constructed by assigning $g_k = \hat{g}_k$ and

$$\boldsymbol{\mathcal{H}}^T = \mathbf{\Gamma}\mathbf{H}^T = \begin{bmatrix} \boldsymbol{h}_1 & \boldsymbol{h}_2 & \cdots & \boldsymbol{h}_K \end{bmatrix}^T \tag{25}$$

$$\boldsymbol{h}_k = \gamma_k \mathbf{h}_k. \tag{26}$$

From (24), the LS estimate of $\boldsymbol{\mathcal{H}}$ is

$$\widehat{\boldsymbol{\mathcal{H}}} = \mathbf{X}^T (\widehat{\mathbf{G}}^\dagger)^T. \tag{27}$$

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

$$\mathbf{T}_1\hat{\boldsymbol{h}}_k h_k \approx \mathbf{T}_2\hat{\boldsymbol{h}}_k \tag{28}$$

where $\mathbf{T}_1 = \begin{bmatrix}\mathbf{I}_{N_2-1} & \mathbf{0}_{(N_2-1)\times 1}\end{bmatrix}^T$ and $\mathbf{T}_2 = \begin{bmatrix}\mathbf{0}_{(N_2-1)\times 1} & \mathbf{I}_{N_2-1}\end{bmatrix}^T$ are selection matrices. Considering sufficiently small error conditions such that $\hat{g}_k \to g_k$, we have $\widehat{\mathbf{G}} \to \mathbf{G}$, and then $\mathbf{X}$ will be independent of $\widehat{\mathbf{G}}$, therefore the disturbances among each vector of $\hat{\boldsymbol{h}}_k$ can be assumed independent and identically distributed. Following [10], the WLS estimate of h_k, $k = 1, 2, \cdots, K$ is computed as:

$$\hat{h}_k = ((\mathbf{T}_1\hat{\boldsymbol{h}}_k)^H \boldsymbol{\Psi}_k \mathbf{T}_2\hat{\boldsymbol{h}}_k)^{-1}(\mathbf{T}_1\hat{\boldsymbol{h}}_k)^H \boldsymbol{\Psi}_k \mathbf{T}_2\hat{\boldsymbol{h}}_k. \tag{29}$$

The optimum weighting matrix $\boldsymbol{\Psi}_k$ has the form:

$$\begin{aligned}\boldsymbol{\Psi}_k &= \left[\mathbb{E}\left\{(\mathbf{T}_1\hat{\boldsymbol{h}}_k h_k - \mathbf{T}_2\hat{\boldsymbol{h}}_k)(\mathbf{T}_1\hat{\boldsymbol{h}}_k h_k - \mathbf{T}_2\hat{\boldsymbol{h}}_k)^H\right\}\right]^{-1} \\ &= \left(\mathbf{B}_k\mathbf{B}_k^H\right)^{-1} \end{aligned} \tag{30}$$

where
$$\begin{aligned}\mathbf{B}_k = &\mathrm{Toeplitz}\left(\begin{bmatrix}-h_k & \mathbf{0}_{1\times(N_2-2)}\end{bmatrix}^T,\right. \\ &\left.\begin{bmatrix}-h_k & 1 & \mathbf{0}_{1\times(N_2-2)}\end{bmatrix}\right).\end{aligned} \tag{31}$$

We start with $\boldsymbol{\Psi}_k = \mathbf{I}_{N-1}$ in the iterations between (29) and (30) to obtain $\hat{h}_k$, $k = 1, 2, \cdots, K$. Finally, the frequencies μ_k and damping factors β_k for $k = 1, 2, \cdots, K$ are estimated as

$$\hat{\mu}_k = \angle(\hat{h}_k) \quad \text{and} \quad \hat{\beta}_k = |\hat{h}_k| \tag{32}$$

where $\hat{h}_k$ and $\hat{g}_k$ are automatically paired up.

As it is shown in [7], $\hat{\mu}_k$ corresponds to $2\lfloor N_1/2\rfloor + 1$ possible estimates of ω_k, where $\lfloor\ \rfloor$ rounds the value to the nearest integer towards $-\infty$, denoted by $\hat{\omega}_{R,k,i}$, $i = -\lfloor N_1/2\rfloor, -\lfloor N_1/2\rfloor + 1, \cdots, \lfloor N_1/2\rfloor$:

$$\hat{\omega}_{R,k,i} = \frac{\hat{\mu}_k + 2\pi i}{N_1} \tag{33}$$

Defining $\hat{\omega}_{R,k} = \hat{\omega}_{R,k,f}$ where f is computed from

$$f = \arg \min_{i\in\{-\lfloor N_1/2\rfloor, -\lfloor N_1/2\rfloor+1, \cdots, \lfloor N_1/2\rfloor\}} |\hat{\omega}_{R,k,i} - \hat{\omega}_{L,k}| \tag{34}$$

and

$$\hat{\alpha}_{R,k} = \hat{\beta}_k^{1/N_1} \tag{35}$$

It will be shown in Section 4 that using $\hat{\omega}_{R,k}$ and $\hat{\alpha}_{R,k}$ has a much higher accuracy than that of $\hat{\omega}_{L,k}$ and $\hat{\alpha}_{L,k}$. Therefore, we assign

$$\hat{\omega}_k \approx \hat{\omega}_{R,k} \quad \text{and} \quad \hat{\alpha}_k \approx \hat{\alpha}_{R,k} \tag{36}$$

as the final estimates.

4. SIMULATION RESULTS

Computer simulations have been carried out to evaluate the performance of the proposed algorithm for multiple damped sinusoids in the presence of white Gaussian noise. The average mean square error (MSE) is assigned to evaluate the algorithm performance and the SNR in dB is defined as SNR = $10\log_{10}(\sum_{n=1}^{N}|s_n|^2/\sigma^2)$. All results provided are averages of 2000 independent runs.

In the first test, we study the performance of the proposed algorithm comparing with the WLS [4] and standard ESPRIT (SE) [5] approaches. The signal parameters are $\begin{bmatrix}\gamma_1 & \gamma_2\end{bmatrix} = \begin{bmatrix}1 & 2e^j\end{bmatrix}$, $\begin{bmatrix}\alpha_1 & \alpha_2\end{bmatrix} = \begin{bmatrix}0.99 & 0.98\end{bmatrix}$, $\begin{bmatrix}\omega_1 & \omega_2\end{bmatrix} = \begin{bmatrix}0.05\pi & 0.36\pi\end{bmatrix}$, $N = 256$ and $N_1 = N_2 = \sqrt{N} = 16$ in (4). Both using (17) and (36) as the final estimates of $\hat{\omega}_k$ and $\hat{\alpha}_k$ are tested, and the average MSEs for frequencies and damping factors versus SNR are plotted in Figures 1 and 2, respectively. It is seen that using (36) as the final estimates outperforms that of (17), therefore we use (36) in the following. It is also shown that the proposed scheme performs almost the same as the SE algorithm, and has a better threshold performance than the GWLP method.

Figure 3 plots the MSEs for frequencies under different choices of $N_1 \times N_2$ for the proposed method with the same data set as in the first test. The plot for damping factor is almost the same therefore is not shown here. It shows that the settings 8×32 and 16×16 have the best performance among all the choices, indicating that the proposed scheme performs the best when the data matrix is approximately square. In the situations when $N_1 >> N_2$ or $N_1 << N_2$, the estimator will suffer from an inaccurate estimation of $\hat{g}_f$ or $\hat{h}_f$ therefore the performance degrades. The average computation times of the WLS, SE and proposed schemes under 4×64, 8×32, 16×16, 32×8 and 64×4 for a single trial are measured as 3.27×10^{-2}s, 3.55×10^{-2}s, 4.39×10^{-3}s, 2.32×10^{-3}s, 2.01×10^{-3}s, 2.31×10^{-3}s and 4.21×10^{-3}s, respectively, showing that the proposed method with the choice of $N_1 = N_2$ is the best estimator in terms of estimation performance and/or computational complexity in this experiment.

Figure 4 investigates the MSEs for frequencies when ω_2 varis from -0.95π to 0.97π at $\omega_1 = -0.99\pi$. The SNR is set to 40dB while other parameters remain unchanged. We see that for a larger N_1, the proposed method can correctly separate more closely spaced sources. It also indicates that with a properly chosen N_1 and N_2, the proposed scheme has almost the same frequency resolution performance as the SE and WLS methods.

5. CONCLUSION

A fast and accurate sinusoidal parameter estimation approach based on principal singular value decomposition of the data matrix has been derived. The key point is to employ the left singular vectors to get a rough estimation of the parameters at first, and then use the whole data matrix together with these values to get a better result. Furthermore, it is demonstrated that the proposed subspace scheme has an outstanding performance in terms of computational complexity and/or estimation accuracy.

6. REFERENCES

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

[2] R. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Transactions on Antennas and Propagation*, vol.34, no.3, pp.276-280, Mar. 1986

[3] R. Roy, A. Paulraj and T. Kailath, “ESPRIT-A subspace rotation approach to estimation of parameters of cisoids in noise,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol.34, no.5, pp.1340-1342, Oct. 1986

[4] H.C.So, Y.T.Chan, K.C.Ho and K.W.Chan, “Unbiased equation-error based algorithms for efficient system identification using noisy measurements,” *Signal Processing*, vol.87, pp.1014-1030, May. 2007

[5] 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

[6] 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

[7] H.C.So, F.K.W.Chan and W.Sun, “Subspace approach for fast and accurate single-tone frequency estimation,” *IEEE Transactions on Signal Processing*, vol.59, no.2, pp.827-831, Feb. 2011

[8] F.K.W. Chan, H.C. So and Weize Sun, “Utilizing principal singular vectors for two-dimensional parameter estimation of multiple damped sinusoids,” *submitted for publication*

[9] S.M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Englewood Cliffs, NJ: Prentice-Hall, 1993

[10] 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

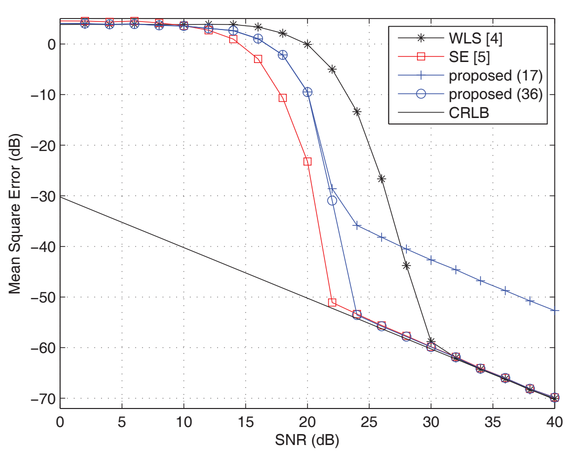

Fig. 1. Average mean square frequency error versus SNR.

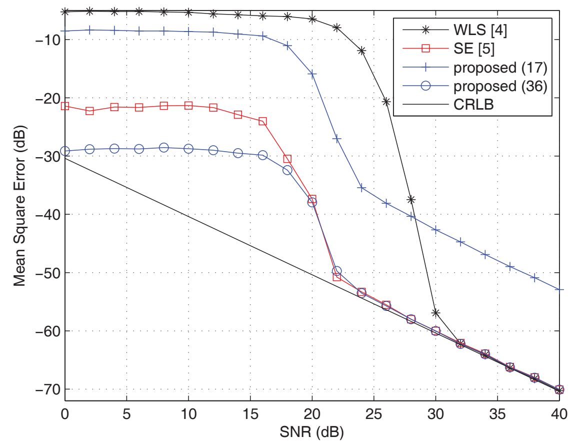

Fig. 2. Average mean square damping factor error versus SNR.

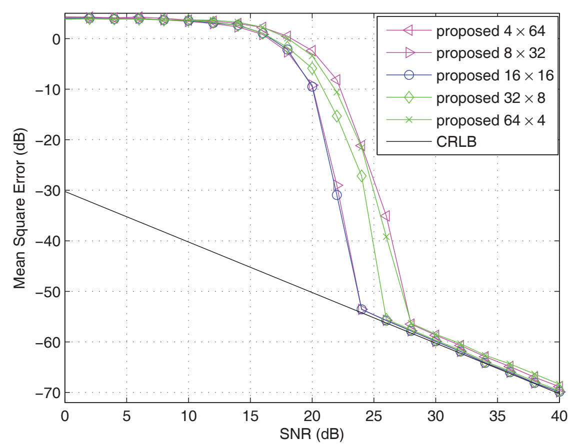

Fig. 3. Average mean square frequency error versus SNR under different $N_1 \times N_2$.

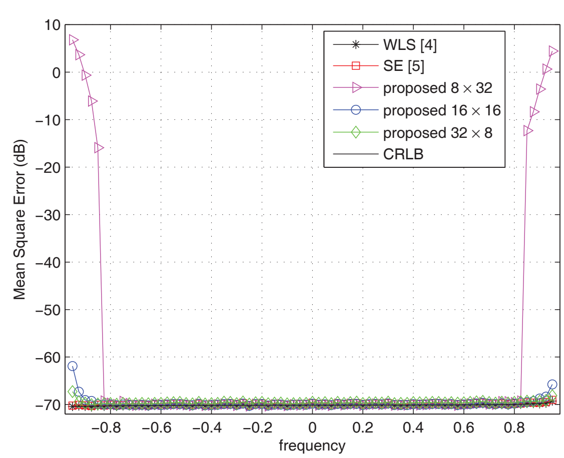

Fig. 4. Average mean square frequency error under $\omega_2 \in (-\pi, \pi)$.