Subspace Approach for Fast and Accurate Single-Tone Frequency Estimation

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

Abstract

A new signal subspace approach for estimating the frequency of a single complex tone in additive white noise is proposed in this paper. Our main ideas are to use a matrix without repeated elements to represent the observed signal and exploit the principal singular vectors of this matrix for frequency estimation. It is proved that for small error conditions, the frequency estimate is approximately unbiased and its variance is equal to Cramér-Rao lower bound. Computer simulations are included to compare the proposed approach with the generalized weighted linear predictor, periodogram and phase-based maximum likelihood estimators in terms of estimation accuracy, computational complexity and threshold performance.

Index Terms

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

I. INTRODUCTION

Frequency estimation of sinusoids from a finite number of noisy discrete-time measurements has been an important topic in science and engineering because many real-world signals can be well described by the sinusoidal model. For example, it has numerous applications in the instrumentation and measurement community such as testing of sine waves [1] and software radio receivers [2], where a standardized algorithm for the former is provided in [3]. In this work, estimating the frequency of a single complex tone in the presence of additive white noise is tackled. Although it is a fundamental and well-studied problem, efforts have continually been made [4]–[12] to derive estimators which can outperform existing works

The authors are with Department of Electronic Engineering, City University of Hong Kong. The work described in this paper was supported by a grant from CityU (Project no. 7002570).

in terms of accuracy, threshold signal-to-noise ratio (SNR), uniform performance and/or implementation complexity.

In the presence of white Gaussian noise, a maximum likelihood (ML) estimate of frequency is obtained from the periodogram maximum [5], and its variance approaches the performance benchmark of Cramér-Rao lower bound (CRLB). As it corresponds to a multimodal surface search, a practical way is to use the discrete Fourier transform (DFT) peak frequency as the coarse estimate which is then refined by a numerical method, and thus there is no guarantee of attaining the global maximum point. Apart from the periodogram, single-tone frequency estimation can also be achieved by autocorrelation or linear prediction (LP) [4], [7], [9]–[10] and phase-based [6]–[8], [11]–[12] approaches. Although they are similar in the sense that they both extract angle information, their basic distinction is that the former utilizes the phase of the autocorrelation function of the observed signal while the latter exploits the signal phase to estimate the frequency. Moreover, the phased-based methodology is more complex than the autocorrelation approach in the sense that phase unwrapping is required due to the modulo- 2π ambiguity in the phase measurements. In this correspondence, we contribute to devise a fast and accurate LP-based single frequency estimator with the use of the subspace technique.

The rest of the paper is organized as follows. In Section II, the proposed single frequency estimator is derived and analyzed in white Gaussian noise with an observation length of K . By stacking the received samples into a $M \times N$ data matrix such that $K = MN$, it is shown that the corresponding left principal singular vector is a tone sequence of the same frequency while the right principal singular vector has a frequency equals M times the original one. It is worthy to mention that this principal singular value decomposition (SVD) idea, which is referred to as principal-singular-vector utilization for modal analysis (PUMA) [13], has recently been applied to the two-dimensional (2D) sinusoid signals, which results in efficient 2D parameter estimation. We then utilize the generalized weighted linear predictor (GWLP) [10] estimator to find the frequencies of the principal singular vectors and optimally combine the two estimates to give the final frequency estimate. It is also proved that for small error conditions, the frequency estimate is approximately unbiased and its variance is equal to the CRLB. Simulation results are included in Section III to corroborate the analytical development and to evaluate the performance of the proposed approach by comparing with the GWLP [10], periodogram [5] and phase-based ML [11] estimators as well as CRLB. Finally, conclusions are drawn in Section IV.

II. PROPOSED FREQUENCY ESTIMATOR

The observed signal is modeled as

$$r_k = s_k + q_k, \quad k = 1, 2, \dots, K \quad (1)$$

where

$$s_k = \gamma e^{j\omega k} \quad (2)$$

The γ is the unknown complex-valued amplitude and $\omega \in (-\pi, \pi)$ is the unknown frequency of s_k , while q_k is the additive measurement noise which is a zero-mean complex white Gaussian process with unknown variance of σ^2 . Given the K samples of $\{r_k\}$, our objective is to find ω . Here we assume that K can be factorized as $K = MN$ with $M \leq N$, where M and N are integers. Note that even if K is not factorizable, one simple way is to discard a few samples and find M and N such that their product is closest to K , and the performance loss will be negligible for a sufficiently large data length.

Our first novel idea is to express $\{r_k\}$ as a $M \times N$ data matrix, denoted by $\mathbf{R}$, which has the form of:

$$\mathbf{R} = \begin{bmatrix} r_1 & r_{M+1} & \cdots & r_{M(N-1)+1} \\ r_2 & r_{M+2} & \cdots & r_{M(N-1)+2} \\ \vdots & \vdots & \ddots & \vdots \\ r_M & r_{2M} & \cdots & r_{MN} \end{bmatrix} \quad (3)$$

The (m, n) entry of $\mathbf{R}$ is represented as $[\mathbf{R}]_{m,n} = r_{m+(n-1)M}$. Unlike the conventional Prony and SVD-based methods where a Hankel-style matrix is constructed for parameter estimation, there are no repeated elements in (3). This indicates that the latter has the potential of achieving a more computationally efficient solution than the former which deals with a larger size matrix. In matrix form, (1) becomes

$$\mathbf{R} = \mathbf{S} + \mathbf{Q} \quad (4)$$

where $\mathbf{S}$ and $\mathbf{Q}$ contain $\{s_k\}$ and $\{q_k\}$ accordingly. By expressing $[\mathbf{S}]_{m,n} = \gamma e^{jm\omega} e^{j(n-1)M\omega}$, it can be factorized as:

$$\mathbf{S} = \gamma \mathbf{g} \mathbf{h}^T \quad (5)$$

where

$$\mathbf{g} = [e^{j\omega} \quad e^{j2\omega} \quad \cdots \quad e^{jM\omega}]^T \quad (6)$$

and

$$\mathbf{h} = [1 \quad e^{jM\omega} \quad \cdots \quad e^{jM(N-1)\omega}]^T \quad (7)$$

with T denotes the transpose operator. This implies that frequency estimation can be achieved by making use of $\mathbf{g}$ and $\mathbf{h}$ as they are pure tones with frequencies ω and $M\omega$, respectively. Let $a = e^{j\omega}$ and $b = e^{j\psi}$ with $\psi = M\omega$, we will employ the LP relations:

$$[\mathbf{g}]_m = a[\mathbf{g}]_{m-1}, \quad m = 2, 3, \dots, M \quad (8)$$

and

$$[\mathbf{h}]_n = b[\mathbf{h}]_{n-1}, \quad n = 2, 3, \dots, N \quad (9)$$

to find ω , where $[\]_m$ denotes the m element in a vector.

As it is not trivial to obtain $\mathbf{g}$ and $\mathbf{h}$ from $\{r_k\}$, we follow the PUMA approach [13] to exploit the principal singular vectors of $\mathbf{R}$ for frequency estimation as follows. First, $\mathbf{R}$ is decomposed using SVD as

$$\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H \quad (10)$$

where $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$ denotes the diagonal matrix of singular values with $\lambda_1 \geq \lambda_2 \geq \dots = \lambda_N \geq 0$ while $\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_M]$ and $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_N]$ are orthonormal matrices whose columns are the corresponding left and right singular vectors, respectively, and H represents conjugate transpose. From (5)–(7), it is obvious that the rank of $\mathbf{S}$ is one and thus the best estimate of $\mathbf{S}$ according to (10), denoted by $\hat{\mathbf{S}}$, is

$$\hat{\mathbf{S}} = \lambda_1 \mathbf{u}_1 \mathbf{v}_1^H \quad (11)$$

Comparing (5) and (11) and noting $\mathbf{u}_1$ and $\mathbf{v}_1$ are unit vectors, it is shown that [13] in the absence of noise, $\tilde{\lambda}_1 = \sqrt{MN}|\gamma|$, $\tilde{\mathbf{u}}_1 = \mathbf{g}e^{-j\varphi_g}/\sqrt{M}$ and $\tilde{\mathbf{v}}_1^* = \mathbf{h}e^{-j\varphi_h}/\sqrt{N}$ where φ_g and φ_h are unknown constants, $\sim$ denotes the noise-free component and $*$ is conjugate operator. This means that $\tilde{\mathbf{u}}_1$ and $\tilde{\mathbf{v}}_1$ possess the same LP property as in (8)–(9) and we will utilize $\mathbf{u}_1$ and $\mathbf{v}_1$ to estimate ω . It is worthy to point out that there is a SNR improvement in $\mathbf{u}_1$ and $\mathbf{v}_1$ over r_k , where the latter SNR is $|\gamma|^2/\sigma^2$, which can lead to a better threshold performance as demonstrated in Section III. Taking $\mathbf{u}_1$ as an illustration, the signal power in $[\mathbf{u}_1]_m$, $m = 1, 2, \dots, M$, is $\mathbb{E}\{|\tilde{\mathbf{u}}_1|_m|^2\} = 1/M$, where $\mathbb{E}$ is the expectation operator. Denoting $\Delta\mathbf{u}_1 = \mathbf{u}_1 - \tilde{\mathbf{u}}_1$, the noise power in $[\mathbf{u}_1]_m$ is [13]:

$$\mathbb{E}\{|\Delta\mathbf{u}_1|_m|^2\} = \frac{\sigma^2}{\tilde{\lambda}_1} - \frac{\sigma^2}{\tilde{\lambda}_1} |[\tilde{\mathbf{u}}_1]_m|^2 = \frac{(M-1)\sigma^2}{|\gamma|^2 M^2 N} \quad (12)$$

As a result, the SNR in $\mathbf{u}_1$ is $MN|\gamma|^2/((M-1)\sigma^2)$ which corresponds to a SNR gain of $MN/(M-1)$ over the raw data.

Define $\mathbf{x}_1 = [[\mathbf{u}_1]_1 \quad [\mathbf{u}_1]_2 \quad \cdots \quad [\mathbf{u}_1]_{M-1}]^T$ and $\mathbf{x}_2 = [[\mathbf{u}_1]_2 \quad [\mathbf{u}_1]_3 \quad \cdots \quad [\mathbf{u}_1]_M]^T$. We have:

$$\mathbf{x}_1 a \approx \mathbf{x}_2 \quad (13)$$

As in the GWLP approach, we make use of the LP property in (13) and $a = e^{j\omega}$, and apply weighted least squares to solve for ω . The corresponding frequency estimate, denoted by $\hat{\omega}_L$, is [10]:

$$\hat{\omega}_L = \angle (\mathbf{x}_1^H \mathbf{W}(\omega) \mathbf{x}_2) \quad (14)$$

where $\angle$ is the phase angle operator and the optimum weighting matrix $\mathbf{W}(\omega)$ is characterized by the unknown frequency, which is computed as [13]:

$$\mathbf{W}(\omega) = \frac{\sigma^2}{\bar{\lambda}_1^2} \left[\mathbb{E} \left\{ (\mathbf{x}_1 e^{j\omega} - \mathbf{x}_2) (\mathbf{x}_1 e^{j\omega} - \mathbf{x}_2)^H \right\} \right]^{-1} = (\mathbf{A} \mathbf{A}^H)^{-1} \quad (15)$$

where

$$\mathbf{A} = \text{Toeplitz} \left(\left[-e^{j\omega} \quad \mathbf{0}_{1 \times (M-2)} \right]^T, \left[-e^{j\omega} \quad 1 \quad \mathbf{0}_{1 \times (M-2)} \right] \right) \quad (16)$$

with $^{-1}$ denotes matrix inverse, $\text{Toeplitz}(\mathbf{a}, \mathbf{b}^T)$ is a Toeplitz matrix with first column $\mathbf{a}$ and first row $\mathbf{b}^T$ and $\mathbf{0}_{i \times j}$ represents a $i \times j$ zero matrix. In fact, $\mathbf{W}(\omega)$ has a closed-form expression with entries [10]:

$$[\mathbf{W}(\omega)]_{m,n} = \frac{M \min(m, n) - mn}{M} e^{j(m-n)\omega}, \quad 1 \leq m \leq M-1, \quad 1 \leq n \leq M-1 \quad (17)$$

As the frequency is unknown, $\hat{\omega}_L$ is obtained using an iterative relaxation procedure which is summarized in Table 1. Note that the algorithm convergence has been proved in [10]. At small error conditions and upon the convergence such that $\hat{\omega}_L$ is located at a reasonable proximity of ω , we have shown that (See Appendix A)

$$\mathbb{E}\{\hat{\omega}_L\} \approx \omega \quad (18)$$

which indicates the approximate unbiasedness of the frequency estimate. Furthermore, its variance, denoted by $\text{var}(\hat{\omega}_L)$, is derived as (See Appendix A):

$$\text{var}(\hat{\omega}_L) \approx \frac{6\sigma^2}{NM(M^2 - 1)|\gamma|^2} \quad (19)$$

In a similar manner, the conceptual estimate of ψ according to the GWLP framework, denoted by $\hat{\psi}$, is:

$$\hat{\psi} = \angle (\mathbf{y}_1^H \mathbf{W}(\psi) \mathbf{y}_2) \quad (20)$$

where $\mathbf{y}_1 = [[\mathbf{v}_1]_1 \quad [\mathbf{v}_1]_2 \quad \cdots \quad [\mathbf{v}_1]_{M-1}]^H$, $\mathbf{y}_2 = [[\mathbf{v}_1]_2 \quad [\mathbf{v}_1]_3 \quad \cdots \quad [\mathbf{v}_1]_M]^H$, the weighting matrix $\mathbf{W}(\psi) \in \mathbb{C}^{(N-1) \times (N-1)}$ has the same definition as in (15)–(16) and $\hat{\psi}$ is obtained by iterative updates of (20) and $\mathbf{W}(\psi)$ as in Table 1. In a similar manner, $\hat{\psi}$ is an approximately unbiased estimate of ψ upon convergence and its variance is:

$$\text{var}(\hat{\psi}) \approx \frac{6\sigma^2}{MN(N^2 - 1)|\gamma|^2} \quad (21)$$

To estimate ω using (20), we notice that $\hat{\psi}$ corresponds to $2\lfloor M/2 \rfloor + 1$ possible estimates of ω , where $\lfloor \cdot \rfloor$ rounds the value to the nearest integer towards $-\infty$, denoted by $\hat{\omega}_{R,i}$, $i = -\lfloor M/2 \rfloor, -\lfloor M/2 \rfloor + 1, \dots, \lfloor M/2 \rfloor$:

$$\hat{\omega}_{R,i} = \frac{\hat{\psi} + 2\pi i}{M} \quad (22)$$

In this study, we calculate the absolute difference between $\hat{\omega}_L$ and each of $\{\hat{\omega}_{R,i}\}$, and the frequency estimate according to $\hat{\psi}$, denoted by $\hat{\omega}_R$, is given by the latter with the smallest difference. Mathematically, $\hat{\omega}_R = \hat{\omega}_{R,k}$ where k is obtained from

$$k = \arg \min_{i \in \{-\lfloor M/2 \rfloor, -\lfloor M/2 \rfloor + 1, \dots, \lfloor M/2 \rfloor\}} |\hat{\omega}_{R,i} - \hat{\omega}_L| \quad (23)$$

From (21) and $\psi = M\omega$, it follows that the variance of $\hat{\omega}_R$ is:

$$\text{var}(\hat{\omega}_R) \approx \frac{6\sigma^2}{M^3N(N^2 - 1)|\gamma|^2} \quad (24)$$

Let $\hat{\omega}$ be the final estimate based on optimally combining $\hat{\omega}_L$ and $\hat{\omega}_R$. Since $\mathbf{u}_1$ and $\mathbf{v}_1^*$ are independent (See Appendix B), $\hat{\omega}_L$ is uncorrelated with $\hat{\omega}_R$. As a result, from (19) and (24), $\hat{\omega}$ is simply computed as:

$$\hat{\omega} = \frac{\frac{\hat{\omega}_L}{\text{var}(\hat{\omega}_L)} + \frac{\hat{\omega}_R}{\text{var}(\hat{\omega}_R)}}{\frac{1}{\text{var}(\hat{\omega}_L)} + \frac{1}{\text{var}(\hat{\omega}_R)}} \approx \frac{1}{M^2N^2 - 1} [(M^2 - 1)\hat{\omega}_L + M^2(N^2 - 1)\hat{\omega}_R] \quad (25)$$

The complete algorithm for estimating ω is summarized in Table 2. The variance of $\hat{\omega}$ is:

$$\text{var}(\hat{\omega}) = \left[\frac{1}{\text{var}(\hat{\omega}_L)} + \frac{1}{\text{var}(\hat{\omega}_R)} \right]^{-1} \approx \frac{6\sigma^2}{MN(M^2N^2 - 1)|\gamma|^2} \quad (26)$$

which is the CRLB for a complex single-tone [5], indicating the optimality of the proposed frequency estimator. From (25), we also notice that $\hat{\omega}_R \rightarrow \hat{\omega}$ for sufficiently large N .

Finally, the computational complexity for $\hat{\omega}$ is studied as follows. Calculating $\mathbf{u}_1$ and $\mathbf{v}_1$ from $\mathbf{R}$ using Lanczos method corresponds to a complexity of $\mathcal{O}(MN)$ [14] while the GWLP procedures for $\hat{\omega}_L$ and $\hat{\omega}_R$ require $\mathcal{O}(M^2)$ and $\mathcal{O}(N^2)$ computations [10], respectively. In particular, when we assign

$M \approx N \approx \sqrt{K}$, the overall algorithm complexity is approximately $O(K)$, which is much simpler than that of [10], namely, $O(K^2)$, and the proposed estimator can be regarded as a computationally efficient algorithm [7].

-
-
- (i) Obtain an initial frequency estimate using (14) and (17) with $[\mathbf{W}(\omega)]_{m,n} = 0$ for $m \neq n$. That is,

$$\hat{\omega}_L = \angle \left(\sum_{m=1}^{M-1} \left(\frac{Nm - m^2}{N} \right) [\mathbf{u}_1]_m^* [\mathbf{u}_1]_{m+1} \right)$$
 Note that the off-diagonal elements of $\mathbf{W}(\omega)$ are not involved as they are characterized by ω .
 - (ii) Construct $\mathbf{W}(\omega)$ according to (17) with $\omega = \hat{\omega}_L$.
 - (iii) Compute an updated $\hat{\omega}_L$ using (14).
 - (iv) Repeat Steps (ii)-(iii) until a stopping criterion is reached.
-

TABLE I
ALGORITHM FOR $\hat{\omega}_L$

-
-
- (i) Use the observed signal to construct $\mathbf{R}$ of (3) and then compute its principal left and right singular vectors, namely, $\mathbf{u}_1$ and $\mathbf{v}_1$.
 - (ii) Use the algorithm in Table 1 to obtain $\hat{\omega}_L$ and $\hat{\psi}$. Note that for the estimation of ψ , $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{W}(\omega)$ are replaced with $\mathbf{y}_1$, $\mathbf{y}_2$ and $\mathbf{W}(\psi)$, respectively.
 - (iii) Determine $\hat{\omega}_R$ from $\hat{\psi}$ according to (22)–(23).
 - (iv) Compute $\hat{\omega}$ using (25).
-

TABLE II
ALGORITHM FOR $\hat{\omega}$
III. NUMERICAL EXAMPLES

Computer simulations have been conducted to evaluate the frequency estimation performance of the proposed methodology by comparing with the GWLP [10], periodogram [5] and phase-based ML [11] schemes. Three iterations are employed as the stopping criterion in the GWLP and proposed methods because no significant improvement is observed for more iterations while the golden section search is used for refining the DFT estimate for the periodogram. The mean square error (MSE) is assigned as the performance measure and CRLB is also included to validate the algorithm optimality. The tone magnitude is chosen as $|\gamma| = 1$ and we properly scale the complex-valued zero-mean white Gaussian noise sequences $\{q_k\}$ to produce different SNR conditions, where $\text{SNR} = |\gamma|^2/\sigma^2 = 1/\sigma^2$. The observed data length is $K = 256$. All results provided are averages of 2000 independent runs using a computer with Intel Core i7 2.67 GHz processors and 3GB RAM.

In the first test, the main focus is to study different combinations of M and N in the proposed algorithm. In each independent trial, $\angle(\gamma)$ and ω are uniformly chosen between $(-\pi, \pi)$. Figure 1 plots

the MSE performance of different frequency estimators versus SNR. Although the proposed estimator with different combinations of M and N attains optimality for sufficiently high SNRs, the choices of $(M, N) = (16, 16)$ and $(M, N) = (8, 32)$ are the best in the sense of achieving the threshold SNR of -4 dB, and the periodogram has the best threshold performance which agrees with [12]. The variance expression of (26) is also verified. Moreover, when the difference between M and N is not very large, the subspace estimator has better threshold performance than that of the GWLP algorithm and this phenomenon also aligns with our SNR improvement analysis in Section II. On the other hand, the phase-based ML method is suboptimal in the mean sense because large errors are introduced in the phase unwrapping step when the frequency values are close to $-\pi$ or π [11]. Regarding computational requirement in a single MATLAB run, the average computation times of the proposed estimator with (M, N) equals $(2, 128)$, $(4, 64)$, $(8, 32)$ and $(16, 16)$ as well as GWLP, periodogram and phase-based ML schemes, are measured as 4.9×10^{-3} s, 2.2×10^{-3} s, 8.91×10^{-4} s, 7.54×10^{-4} s, 9.6×10^{-3} s, 4.2×10^{-3} s and 8.4×10^{-3} s, respectively. To summarize, the proposed approach with $M = N = 16$ is the best estimator in terms of estimation performance and/or computational complexity in this experiment.

Figures 2 to 5 plot the frequency versus SNR contours of MSE for the proposed, GWLP, periodogram and phase-based ML methods, respectively, with $M = N = 16$ in order to investigate the estimation performance in more detail. It is observed that all methods except the last one have the desirable merit of uniform performance. The threshold performance also agrees with the first test, that is, the periodogram is the best, followed by the proposed, GWLP and phase-based ML methods.

IV. CONCLUSION

A fast and accurate single frequency estimation approach based on principal singular value decomposition of the corresponding data matrix has been devised. As the frequency information is embedded in the principal singular vectors, we apply linear prediction and weighted linear squares on them to perform estimation, which results in computational attractiveness and signal-to-noise ratio (SNR) improvement. It is proved that the performance of the estimator attains Cramér-Rao lower bound at sufficiently small error conditions. Furthermore, it is demonstrated that the proposed subspace scheme is superior to the generalized weighted linear predictor, periodogram and phase-based maximum likelihood approaches in terms of accuracy, computational complexity and/or threshold performance.

APPENDIX A

In this Appendix, the bias and variance of $\hat{\omega}_L$ in (14) are derived. According to (13), the WLS estimate of a , denoted by $\hat{a}$, is:

$$\hat{a} = \arg \min_{\tilde{a}} J(\tilde{a}) \quad (\text{A.1})$$

$$\text{where } J(\tilde{a}) = (\mathbf{x}_1 \tilde{a} - \mathbf{x}_2)^H \mathbf{W}(a) (\mathbf{x}_1 \tilde{a} - \mathbf{x}_2), \quad a = e^{j\omega_L} \quad (\text{A.2})$$

Note that the weighting matrix is now a function of a instead of ω_L . At sufficiently high SNR conditions such that $\hat{a}$ is located at a reasonable proximity of a and assuming that $J''(\tilde{a})$ is smooth enough around a , we expand $J'(\hat{a})$ using Taylor series to yield [15]:

$$0 = J'(\hat{a}) = J'(a) + J''(a)(\hat{a} - a) \approx J'(a) + \mathbb{E}\{J''(a)\}(\hat{a} - a) \quad (\text{A.3})$$

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

$$\begin{aligned} J'(a) &= 2\mathbf{x}_1^H \mathbf{W}(a) (\mathbf{x}_1 a - \mathbf{x}_2) = 2(\tilde{\mathbf{x}}_1 + \Delta \mathbf{x}_1)^H \mathbf{W}(a) [(\tilde{\mathbf{x}}_1 + \Delta \mathbf{x}_1)a - (\tilde{\mathbf{x}}_2 + \Delta \mathbf{x}_2)] \\ &\approx 2\tilde{\mathbf{x}}_1^H \mathbf{W}(a) (\Delta \mathbf{x}_1 a - \Delta \mathbf{x}_2) = 2\tilde{\mathbf{x}}_1^H \mathbf{W}(a) \mathbf{A} \Delta \mathbf{u}_1 \end{aligned} \quad (\text{A.4})$$

From [16], $\Delta \mathbf{u}_1$ is approximated as

$$\Delta \mathbf{u}_1 \approx \tilde{\lambda}_1^{-1} \tilde{\mathbf{U}}_n \tilde{\mathbf{U}}_n^H \mathbf{Q} \tilde{\mathbf{v}}_1 = \tilde{\lambda}_1^{-1} \left(\tilde{\mathbf{v}}_1^T \otimes \tilde{\mathbf{U}}_n \tilde{\mathbf{U}}_n^H \right) \text{vec}(\mathbf{Q}) = \tilde{\lambda}_1^{-1} \left(\tilde{\mathbf{v}}_1^T \otimes (\mathbf{I}_M - \tilde{\mathbf{u}}_1 \tilde{\mathbf{u}}_1^H) \right) \text{vec}(\mathbf{Q}) \quad (\text{A.5})$$

where $\mathbf{I}_M$ is the $M \times M$ identity matrix, $\otimes$ is the Kronecker product and vec denotes vectorization and $\tilde{\mathbf{U}}_n = [\tilde{\mathbf{u}}_2 \ \tilde{\mathbf{u}}_3 \ \cdots \ \tilde{\mathbf{u}}_N]$ is the noise-free version of $\mathbf{U}_n = [\mathbf{u}_2 \ \mathbf{u}_3 \ \cdots \ \mathbf{u}_N]$, which corresponds to the noise subspace. For zero-mean $\mathbf{Q}$, $\mathbb{E}\{\Delta \mathbf{u}_1\}$ approaches a zero vector. Moreover, $\mathbb{E}\{\Delta \mathbf{u}_1 \Delta \mathbf{u}_1^H\}$ is:

$$\begin{aligned} \mathbb{E}\{\Delta \mathbf{u}_1 \Delta \mathbf{u}_1^H\} &= \tilde{\lambda}_1^{-2} \left(\tilde{\mathbf{v}}_1^T \otimes \tilde{\mathbf{U}}_n \tilde{\mathbf{U}}_n^H \right) \mathbb{E}\{\text{vec}(\mathbf{Q}) \text{vec}(\mathbf{Q})^H\} \left(\tilde{\mathbf{v}}_1^* \otimes \tilde{\mathbf{U}}_n \tilde{\mathbf{U}}_n^H \right) \\ &= \tilde{\lambda}_1^{-2} \left(\tilde{\mathbf{v}}_1^T \otimes \tilde{\mathbf{U}}_n \tilde{\mathbf{U}}_n^H \right) (\sigma^2 \mathbf{I}_M) \left(\tilde{\mathbf{v}}_1^* \otimes \tilde{\mathbf{U}}_n \tilde{\mathbf{U}}_n^H \right) \\ &= \tilde{\lambda}_1^{-2} \sigma^2 \left(\tilde{\mathbf{v}}_1^T \tilde{\mathbf{v}}_1^* \right) \otimes \left(\tilde{\mathbf{U}}_n \tilde{\mathbf{U}}_n^H \tilde{\mathbf{U}}_n \tilde{\mathbf{U}}_n^H \right) = \tilde{\lambda}_1^{-2} \sigma^2 \tilde{\mathbf{U}}_n \tilde{\mathbf{U}}_n^H \end{aligned} \quad (\text{A.6})$$

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

$$\mathbb{E}\{J''(a)\} = \mathbb{E}\{2\mathbf{x}_1^H \mathbf{W}(a) \mathbf{x}_1\} \approx 2\tilde{\mathbf{x}}_1^H \mathbf{W}(a) \tilde{\mathbf{x}}_1 \quad (\text{A.7})$$

Let $\mathbf{g}_1$ be $\mathbf{g}$ without the last element. Using $\tilde{\mathbf{u}}_1 = \mathbf{g} e^{-j\varphi_g} / \sqrt{M}$ and interchanging $\mathbf{W}(\omega)$ and $\mathbf{W}(a)$, $\tilde{\mathbf{x}}_1^H \mathbf{W}(a) \tilde{\mathbf{x}}_1$ is calculated as:

$$\begin{aligned} \tilde{\mathbf{x}}_1^H \mathbf{W}(a) \tilde{\mathbf{x}}_1 &= \frac{1}{M^2} \mathbf{g}_1^H \mathbf{W}(\omega) \mathbf{g}_1 = \frac{1}{M^2} \sum_{m=1}^{M-1} \sum_{n=1}^{M-1} e^{-j\omega m} (M \min(m, n) - mn) e^{j(m-n)\omega} e^{j\omega n} \\ &= \frac{1}{M^2} \left(\sum_{m=1}^{M-1} (Mm - m^2) + \sum_{m=1}^{M-1} \sum_{n=1}^{m-1} (Mn - mn) + \sum_{n=1}^{M-1} \sum_{m=1}^{n-1} (Mm - mn) \right) = \frac{M^2 - 1}{12} \end{aligned} \quad (\text{A.8})$$

Combining (A.7)–(A.8), we get:

$$\mathbb{E}\{J''(a)\} \approx \frac{M^2 - 1}{6} \quad (\text{A.9})$$

With the use of (A.4) and (A.9), taking the expectation on (A.3) yields $\mathbb{E}\{\hat{a}\} \approx a$ because $\mathbb{E}\{J'(a)\} \approx 0$. As $\hat{a} = e^{j\hat{\omega}_L}$, $\hat{\omega}_L$ is an approximately unbiased estimate of ω . Employing (A.3) again, the variance of $\hat{a}$, denoted by $\text{var}(\hat{a})$, is:

$$\text{var}(\hat{a}) = \mathbb{E}\{(\hat{a} - a)(\hat{a} - a)^*\} = \frac{\mathbb{E}\{J'(a)(J'(a))^*\}}{[\mathbb{E}\{J''(a)\}]^2} \quad (\text{A.10})$$

With the use of (A.6), $\tilde{\mathbf{U}}_n \tilde{\mathbf{U}}_n^H = \mathbf{I}_M - \tilde{\mathbf{u}}_1 \tilde{\mathbf{u}}_1^H$, $\mathbf{A} \tilde{\mathbf{u}}_1 = \mathbf{0}_{(M-1) \times 1}$ and $\mathbf{W}(a) = \mathbf{W}^H(a) = (\mathbf{A} \mathbf{A}^H)^{-1}$, the numerator of (A.10) is:

$$\begin{aligned} \mathbb{E}\{J'(a)(J'(a))^*\} &= 4\tilde{\mathbf{x}}_1 \mathbf{W}(a) \mathbf{A} \mathbb{E}\{\{\Delta \mathbf{u}_1 \Delta \mathbf{u}_1^H\} \mathbf{A}^H \mathbf{W}^H(a) \tilde{\mathbf{x}}_1\} \\ &= 4\tilde{\lambda}_1^{-2} \sigma^2 \tilde{\mathbf{x}}_1 \mathbf{W}(a) \mathbf{A} \tilde{\mathbf{U}}_n \tilde{\mathbf{U}}_n^H \mathbf{A}^H \mathbf{W}^H(a) \tilde{\mathbf{x}}_1 = 4\tilde{\lambda}_1^{-2} \sigma^2 \tilde{\mathbf{x}}_1 \mathbf{W}(a) \mathbf{A} (\mathbf{I}_M - \tilde{\mathbf{u}}_1 \tilde{\mathbf{u}}_1^H) \mathbf{A}^H \mathbf{W}(a) \tilde{\mathbf{x}}_1 \\ &= 4\tilde{\lambda}_1^{-2} \sigma^2 \tilde{\mathbf{x}}_1 \mathbf{W}(a) \mathbf{W}^{-1}(a) \mathbf{W}(a) \tilde{\mathbf{x}}_1 = \frac{4\tilde{\lambda}_1^{-2} \sigma^2 (M^2 - 1)}{12} \end{aligned} \quad (\text{A.11})$$

Substituting (A.9) and (A.11) with $\tilde{\lambda}_1 = \sqrt{MN}|\gamma|$ into (A.10) yields:

$$\text{var}(\hat{a}) \approx \frac{12\sigma^2}{NM(M^2 - 1)|\gamma|^2} \quad (\text{A.12})$$

Employing the transformation formula of $\text{var}(\hat{\omega}_L) \approx \text{var}(\hat{a})/(2|a|^2)$ [17], we eventually get (19).

APPENDIX B

In this Appendix, we will prove that $\mathbf{u}_1$ and $\mathbf{v}_1^*$ are independent. Expressing $\mathbf{u}_1 = \tilde{\mathbf{u}}_1 + \Delta \mathbf{u}_1$ and $\mathbf{v}_1 = \tilde{\mathbf{v}}_1 + \Delta \mathbf{v}_1$ and using (A.5), we have $\mathbb{E}\{\Delta \mathbf{u}_1\} = \mathbf{0}_{M \times 1}$ and $\mathbb{E}\{\Delta \mathbf{v}_1\} = \mathbf{0}_{N \times 1}$, and it suffices to show that $\mathbb{E}\{\Delta \mathbf{u}_1 \Delta \mathbf{v}_1^H\} = \mathbf{0}_{M \times N}$. The first-order perturbation of $\Delta \mathbf{u}_1$ has been given in (A.5) while that of $\Delta \mathbf{v}_1$ is obtained similarly by considering the SVD of $\mathbf{R}^H$:

$$\Delta \mathbf{v}_1 \approx \tilde{\lambda}_1^{-1} \left(\tilde{\mathbf{u}}_1^T \otimes \tilde{\mathbf{V}}_n \tilde{\mathbf{V}}_n^H \right) \text{vec}(\mathbf{Q}^H) \quad (\text{B.1})$$

where $\tilde{\mathbf{V}}_n = [\tilde{\mathbf{v}}_2 \ \tilde{\mathbf{v}}_3 \ \cdots \ \tilde{\mathbf{v}}_N]$ is the noise-free version of $\mathbf{V}_n$. As $\mathbb{E}\{\text{vec}(\mathbf{Q}) (\text{vec}(\mathbf{Q}^H))^H\} = \mathbf{0}_{M \times N}$, the proof is completed.

REFERENCES

- [1] IEEE Standard for Digitizing Waveform Recorders, IEEE Std. 1057, 2007
- [2] P. Händel and P. Zetterberg, "Receiver I/Q imbalance: Tone test, sensitivity analysis, and the universal software radio peripheral," *IEEE Transactions on Instrumentation and Measurement*, vol.59, no.3, pp.704-714, Mar. 2010
- [3] P. Händel, "Properties of the IEEE-STD-1057 four parameter sine wave fit algorithm," *IEEE Transactions on Instrumentation and Measurement*, vol.49, no.6, pp.1189-1193, Dec. 2000

- [4] G. W. Lank, I. S. Reed and G. E. Pollon, "A semicoherent detection and Doppler estimation statistic," *IEEE Trans. Aerosp. Electron. Syst.*, vol.9, pp.151-165, Mar. 1973
- [5] D. C. Rife and R. R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," *IEEE Transactions on Information Theory*, vol.20, no.5, pp.591-598, Sep. 1974
- [6] S. A. Tretter, "Estimating the frequency of a noisy sinusoid by linear regression," *IEEE Trans. Inform. Theory*, vol.31, pp.832-835, Nov. 1985
- [7] S. Kay, "A fast and accurate single frequency estimator," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.37, pp.1987-1990, Dec. 1989
- [8] M. L. Fowler and J. A. Johnson, "Extending the threshold and frequency range for phase-based frequency estimation," *IEEE Trans. Signal Processing*, vol.47, pp.2857-2863, Oct. 1999
- [9] B. Völcker and P. Händel, "Frequency estimation from proper sets of correlations," *IEEE Trans. Signal Processing*, vol.50, pp.791-802, Apr. 2002
- [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
- [11] H. Fu and P. Y. Kam, "MAP/ML estimation of the frequency and phase of a single sinusoid in noise," *IEEE Transactions on Signal Processing*, vol.55, no.3, pp.834-845, Mar. 2007
- [12] R. G. McKilliam, B. G. Quinn, I. V. L. Clarkson and B. Moran, "Frequency estimation by phase unwrapping," *IEEE Transactions on Signal Processing*, vol.58, no.6, pp.2953-2963, Jun. 2007
- [13] 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
- [14] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Baltimore and London, John Hopkins University Press, 1996
- [15] K. W. Cheung and H. C. So, "A multidimensional scaling framework for mobile location using time-of-arrival measurements," *IEEE Transactions on Signal Processing*, vol.53, no.2, pp.460-470, Feb. 2005
- [16] 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
- [17] 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

Fig. 1. Mean square error of $\hat{\omega}$ versus SNR

Fig. 2. Contour plot of proposed method

Fig. 3. Contour plot of GWLP method

Fig. 4. Contour plot of Periodogram

Fig. 5. Contour plot of Phase-based ML method