A Generalized Weighted Linear Predictor Frequency Estimation Approach for a Complex Sinusoid

H. C. So, *Member, IEEE,* and Frankie Kit Wing Chan

***Abstract*—Based on linear prediction and weighted least squares, three simple iterative algorithms for frequency estimation of a complex sinusoid in additive white noise are devised. The proposed approach, which utilizes the first-order as well as higher order linear prediction terms simultaneously but does not require phase unwrapping, can be considered as a generalized version of the weighted linear predictor frequency estimator. In particular, convergence as well as mean and variance analysis of the most computationally efficient frequency estimator, namely, GWLP 2, are provided. Computer simulations are included to contrast the performance of the proposed algorithms with several conventional computationally attractive frequency estimators and Cramér–Rao lower bound for different frequencies, observation lengths, and signal-to-noise ratios.**

***Index Terms*—Frequency estimation, iterative algorithm, linear prediction, low complexity.**

I. Introduction

PARAMETER estimation of sinusoids in noise has been a classical problem for more than 200 years [1] and is still an important research topic because of its numerous applications in multiple disciplines such as control theory, signal processing, digital communications, biomedical engineering, instrumentation and measurement. Estimation of the frequencies is often the crucial step in the problem because they are nonlinear functions in the received data sequence. Once the frequencies have been determined, the remaining parameters, namely, amplitudes and phases, can then be computed straightforwardly [2]. It is noteworthy that for exponentially damped sinusoids, we also need to determine the extra parameters of damping factors, which can be estimated jointly with the frequencies [3]. For comprehensive readings on frequency estimation and tracking, the interested reader is referred to [2] and [4]–[6].

In this paper, we consider the most basic form of the frequency estimation problem, namely, finding the frequency of a pure complex tone in additive white noise. Mathematically, the single tone model is

$$x_n = s_n + q_n, n = 1, 2, \ldots, N \tag{1}$$

where

$$s_n = Ae^{j(\omega n + \phi)}. \tag{2}$$

Manuscript received June 13, 2004; revised May 17, 2005. The work described in this paper was supported by a grant from CityU (Project no. 7 001 708). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Tulay Adali.

The authors are with the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong (e-mail: hcso@ee.cityu.edu.hk; k.w.chan@student.cityu.edu.hk).

Digital Object Identifier 10.1109/TSP.2005.863119

The sinusoidal amplitude, frequency, and phase are denoted by $A > 0$, ω, and ϕ, respectively, and they are considered as deterministic but unknown constants. While the noise q_n is assumed to be a zero-mean complex white process of the form $q_n = Au_n + jAv_n$, where Au_n and Av_n are zero-mean real white processes with identical but unknown variances of $\sigma^2/2$ and uncorrelated with each other. Although estimating the single frequency is a fundamental and well-studied problem, efforts have continually been made [7]–[16] to derive estimators that can attain high estimation performance but with low computational cost. Our objective is also to estimate $\omega \in (-\pi, \pi)$ accurately in a computationally simple manner, from the N discrete-time noisy measurements of $\{x_n\}$.

In the presence of white Gaussian noise, the maximum-likelihood (ML) estimate of frequency is obtained from the periodogram maximum [17] but it involves extensive computations. To avoid high computational requirement, autocorrelation or linear prediction [7], and phase-based [8] approaches are widely used choices. 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 $\{x_n\}$, denoted by $R_{xx}(l)$, where l is the lag, while the latter considers the signal phase to achieve frequency estimation. Founded on [7] and [8], many computationally efficient frequency estimators with suboptimal performance have been proposed in the literature, to name but a few [9]–[11], [13]–[16]. Kay [9] has proposed the so-called weighted linear predictor (WLP) frequency estimator [18], which introduces different weights in computing a generalized version of $R_{xx}(1)$. Frequency estimation from a set of $\{R_{xx}(l)\}$ has been investigated in [10], [13], [15]. Recently, Brown and Wang [16] have suggested to use linear prediction together with low-pass filtering, decimation and heterodyning iteratively for single frequency estimation. On the other hand, an alternative to the phase-based approach [8] is devised by using the differenced phase data, which is known as the weighted phase averager (WPA), and this technique has been extended via the use of simple low-pass filtering and a set of filter banks in [11] and [14], respectively. However, as discussed in [16], most of these computationally attractive schemes have the demerits of poor threshold performance, nonuniform estimation performance across the admissible frequency range, limited frequency operation range, and/or requirement of phase unwrapping which becomes prone to errors at low signal-to-noise ratio (SNR).

The primary motivation of this work is to develop computationally simple and accurate frequency estimators which do not have the above drawbacks. Based on the linear prediction property of a complex tone, we have devised three frequency

1053-587X/$20.00 © 2006 IEEE

estimators which can be considered as a generalization of the WLP approach [9], [18]. All the algorithms, namely, GWLP 1, GWLP 2, and GWLP 3, can be easily programmed on computers and their computational requirement is comparable to that of the autocorrelation and phased-based methods. In particular, we have proved the convergence of GWLP 2 and have derived its variance, which can attain Cramér-Rao lower bound (CRLB) for white Gaussian noise.

The rest of the paper is organized as follows. The development of the generalized weighted linear predictor (GWLP) frequency estimation approach is given in Section II. The basic algorithm GWLP 1 is first devised via the use of linear prediction and weighted least squares (WLS). Two alternative realizations, namely, GWLP 2 and GWLP 3, are then proposed. Furthermore, the relationships between the proposed estimators and existing approaches are discussed. In Section III, a detailed study of GWLP 2 is provided, which includes its computational requirement, convergence as well as mean and variance analysis. Numerical examples are presented in Section IV to corroborate the analytical development and to evaluate the performance of the proposed algorithms by comparing with the minimal order linear predictor (LP) [7], WPA and WLP [9], as well as CRLB. Finally, conclusions are drawn in Section V.

II. Algorithm Development

In this section, three GWLP frequency estimators for a complex noisy sinusoid will be developed and their relationships with several well-known frequency estimation methods will also be illustrated.

A. Basic Algorithm

The linear prediction property of s_n can be expressed as

$$s_n = \rho s_{n-1}, \rho = e^{j\omega}. \tag{3}$$

Based on (3), the linear prediction error is

$$\epsilon_n = x_n - \tilde{\rho} x_{n-1}, \quad n = 2, 3, \cdots, N \tag{4}$$

where $\tilde{\rho}$ is a variable corresponds to ρ, which is to be determined. Expressing (4) into vector form yields

$$\boldsymbol{\epsilon} = \mathbf{X_1} - \tilde{\rho}\mathbf{X_2} \tag{5}$$

where

$$\begin{aligned} \boldsymbol{\epsilon} &= [\epsilon_N, \epsilon_{N-1}, \ldots, \epsilon_2]^T \\ \mathbf{X_1} &= [x_N, x_{N-1}, \ldots, x_2]^T \\ \mathbf{X_2} &= [x_{N-1}, x_{N-2}, \ldots, x_1]^T \end{aligned}$$

and T denotes the transpose operation. The WLS cost function constructed from the linear prediction error is then

$$J(\tilde{\rho}) = \boldsymbol{\epsilon}^H \mathbf{W} \boldsymbol{\epsilon} = (\mathbf{X_1} - \tilde{\rho}\mathbf{X_2})^H \mathbf{W} (\mathbf{X_1} - \tilde{\rho}\mathbf{X_2}) \tag{6}$$

where H represents the conjugate transpose and $\mathbf{W} \in \mathcal{C}^{(N-1)\times(N-1)}$ is a weighting matrix which satisfies $\mathbf{W} = \mathbf{W}^H$. An ideal choice of $\mathbf{W}$ is obtained from the Markov estimate [13], [19]

$$\mathbf{W} = \sigma^2 \left[E\left\{\boldsymbol{\epsilon}_q \boldsymbol{\epsilon}_q^H\right\}\right]^{-1} = \mathbf{P}^{-1} \tag{7}$$

where

$$\mathbf{P} = \begin{bmatrix} 1+|\rho|^2 & -\rho & 0 & 0 & \cdots & 0 \\ -\rho^* & 1+|\rho|^2 & -\rho & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & -\rho^* & 1+|\rho|^2 & -\rho \\ 0 & 0 & \cdots & 0 & -\rho^* & 1+|\rho|^2 \end{bmatrix} \tag{8}$$

with $\boldsymbol{\epsilon}_q = [q_N - \rho q_{N-1}, q_{N-1} - \rho q_{N-2}, \ldots, q_2 - \rho q_1]^T$, E is the expectation operator, $*$ represents conjugate, and $^{-1}$ denotes matrix inverse. Differentiating $J(\tilde{\rho})$ with respect to $\tilde{\rho}$ and then setting the resultant expression to zero, we get the estimate of ρ, denoted by $\hat{\rho}$

$$\hat{\rho} = \frac{\mathbf{X_2}^H \mathbf{W} \mathbf{X_1}}{\mathbf{X_2}^H \mathbf{W} \mathbf{X_2}}. \tag{9}$$

As the ideal $\mathbf{W}$ is a function of the unknown parameter ρ, we propose to use a relaxation algorithm [20] for iterative frequency estimation, which is denoted as GWLP 1, and the procedure is summarized as follows.

i) Find a coarse estimate of ρ from the WLP frequency estimate, denoted by $\hat{\omega}_{\mathrm{WLP}}$. The initial $\hat{\rho}$ is given by $\hat{\rho} = e^{j\hat{\omega}_{\mathrm{WLP}}}$. The WLP is used because it belongs to the GWLP approach and their relationship will be shown clearly in Section II-C. It is noteworthy to mention that other simple frequency estimators such as the minimal order LP and discrete Fourier transform (DFT) [21] can be used for initialization as well.
ii) Use $\hat{\rho}$ to construct $\mathbf{W}$ from (7) and (8).
iii) Compute an updated version of $\hat{\rho}$ using (9).
iv) Repeat Steps ii) and iii) until parameter convergence.
v) The frequency estimate of GWLP 1, denoted by $\hat{\omega}_1$, is calculated as $\hat{\omega}_1 = \angle(\hat{\rho})$, where $\angle(.)$ represents the phase angle in $[-\pi, \pi)$.

It is seen that a major computational requirement of the GWLP 1 is to perform the matrix inverse in (7) and attempts [22], [23] have been made to determine a closed-form expression for $\mathbf{P}^{-1}$ though unsuccessful. Note that it is possible that a closed-form result exists, which is not known by the authors. Nevertheless, efficient computation of the inverse can be achieved via the use of Cholesky decomposition [24] or exploiting the tri-diagonal property of $\mathbf{P}$ [25]. In particular, the former technique only requires the order N of flops, where each flop means roughly a complex multiplication and a complex addition [26].

B. Alternative Realizations

Since $\rho = e^{j\omega}$, an alternative form for the ideal weighting matrix is then

$$\mathbf{W} = \boldsymbol{\Omega}^{-1} \tag{10}$$

where

$$\mathbf{\Omega} = \begin{bmatrix} 2 & -e^{j\omega} & 0 & 0 & \cdots & 0 \\ -e^{-j\omega} & 2 & -e^{j\omega} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & -e^{-j\omega} & 2 & -e^{j\omega} \\ 0 & 0 & \cdots & 0 & -e^{-j\omega} & 2 \end{bmatrix}. \tag{11}$$

It is worthy to note that in practical implementation, using (7) is different from (10) because the magnitude of $\hat{\rho}$ in (9) is generally not equal to unity due to noise, which implies that the latter is an approximation. Nevertheless, it is easy to show that (10) has a closed form and its (m,n) entry is expressed as

$$[\mathbf{W}]_{m,n} = \frac{N\min(m,n) - mn}{N} e^{j(n-m)\omega}, \quad 1 \le m \le N-1, \quad 1 \le n \le N-1 \tag{12}$$

where $\min(m,n) = m$ if $m < n$ and it is equal to n otherwise. From (12), we also notice that $\mathbf{\Omega}^{-1}$ is a positive-definite Hermitian matrix, which implies that $\mathbf{X_2}^H \mathbf{\Omega}^{-1} \mathbf{X_2}$ is real and always has zero phase. As a result, the denominator of (9) can be removed if the weighting matrix of (10) is employed. Based on these findings, we suggest our second algorithm for single frequency estimation, which is referred to as GWLP 2, as follows.

i) Obtain an initial frequency estimate of GWLP 2 denoted by $\hat{\omega}_2$ using WLP, that is, $\hat{\omega}_2 = \hat{\omega}_{\mathrm{WLP}}$.
ii) Use $\hat{\omega}_2$ to construct $\mathbf{W}$ from (12).
iii) Compute an updated $\hat{\omega}_2$ using

$$\hat{\omega}_2 = \angle\left(\mathbf{X_2}^H \mathbf{W} \mathbf{X_1}\right). \tag{13}$$

iv) Repeat Steps ii) and iii) until parameter convergence.

From GWLP 1 and GWLP 2, we propose the third frequency estimation algorithm, namely, GWLP 3, which replaces $\hat{\rho}$ in (9) by $\hat{\zeta}$, which is defined as

$$\hat{\zeta} = \mathbf{X_2}^H \mathbf{W} \mathbf{X_1} \tag{14}$$

where we ignore the denominator of (9) in the GWLP 1. The steps in GWLP 3 are as follows.

i) Use $\hat{\zeta} = e^{j\hat{\omega}_{\mathrm{WLP}}}$ to obtain an initial estimate.
ii) Substitute $\hat{\zeta}$ for ρ to construct $\mathbf{W}$ from (7) and (8).
iii) Compute an updated version of $\hat{\zeta}$ using (14).
iv) Repeat Steps ii) and iii) until parameter convergence.
v) The frequency estimate of GWLP 3, denoted by $\hat{\omega}_3$, is calculated as $\hat{\omega}_3 = \angle(\hat{\zeta})$.

It is obvious that among the three proposed estimators, GWLP 1 involves the highest computational requirement whereas GWLP 2 is the most computationally simple, although all algorithms can be easily programmed on computers. Admitting that GWLP 2 and GWLP 3 are approximate forms of GWLP 1, the approximation in GWLP 2 can also be interpreted as application of the unity-magnitude constraint in $\hat{\rho}$ which is justifiable while $|\hat{\zeta}|$ will be increasing or decreasing during the iterative procedure because there is a discrepancy in the magnitudes of both sides of (14) as the denominator term is removed. As a result, it is expected that both GWLP 1 and GWLP 2 perform comparably and are superior to GWLP 3, which will be demonstrated via computer simulations in Section IV.

C. Relations With Existing Approaches

The relationships between the proposed estimators and some well-known estimation methods in the literature are now discussed. Following the development in [27], it can be shown that the minimizer of (6) with $\mathbf{W}$ parameterized by $\tilde{\rho}$ is in fact the ML estimator in the presence of white Gaussian noise. The GWLP 1 can also be viewed as the so-called iterative quadratic maximum-likelihood (IQML) method [28], [29] which relaxes the ML cost function to a quadratic form, for the special case when the source signal is a pure complex sinusoid. It should be noted that although the two methods are equal in the implementation, their derivations are different: Our approach utilizes the WLS technique while the IQML estimate is derived from the ML criterion and assumes white Gaussian noise. Due to their equivalence, we may utilize the IQML properties in [30] to prove the local convergence and derive the theoretical performance of GWLP 1. On the other hand, since it has been shown [31] that the IQML method is identical to Steiglitz–McBride algorithm [32], we can also apply the convergence results of the iterative filtering algorithm for frequency estimation [33], which bears a strong resemblance to [32], to GWLP 1.

The proposed algorithms are also related to the linear prediction approach as follows. When we use the identity matrix as the weighing matrix in GWLP 1, GWLP 2 and GWLP 3, all of them will reduce to the minimal order LP [7]. For the GWLP 2 and GWLP 3, we can express the frequency estimate in a scalar form

$$\hat{\omega}_i = \angle\left(\sum_{m=1}^{N-1}\sum_{n=1}^{N-1} x_{N-m}^* [\mathbf{W}]_{m,n} x_{N+1-n}\right), \quad i = 2,3 \tag{15}$$

where $\mathbf{W}$ is a function of $\hat{\omega}_2$ for $\hat{\omega}_2$ and it is characterized by $\hat{\zeta}$ for $\hat{\omega}_3$. In particular, the estimate of GWLP 2 can be further expanded via summing up the diagonals of $\mathbf{W}$ one by one, as follows:

$$\begin{aligned}\hat{\omega}_2 = \angle\Bigg(&\sum_{m=1}^{N-1} x_m^* [\mathbf{W}]_{m,m} x_{m+1} \\ &+ \sum_{l=1}^{N-2}\sum_{k=l+1}^{N-1} x_{N-k}^* [\mathbf{W}]_{k,k-l} x_{N-k+l+1} \\ &+ \sum_{l=2-N}^{-1}\sum_{k=1}^{N+l-1} x_{N-k}^* [\mathbf{W}]_{k,k-l} x_{N-k+l+1}\Bigg)\end{aligned} \tag{16}$$

with

$$\begin{aligned}[\mathbf{W}]_{m,m} &= \frac{Nm - m^2}{N} \\ &= \frac{N}{4}\left(1 - \left[\frac{m - 1 - \left(\frac{N}{2} - 1\right)}{\frac{N}{2}}\right]^2\right)\end{aligned} \tag{17}$$

TABLE I
COMPUTATIONAL COMPLEXITY OF DIFFERENT LOW-COMPLEXITY FREQUENCY ESTIMATORS

Method	Real-valued Multiplication	Real-valued Addition	Phase Calculation
GWLP 2 (up to 1st iteration)	$6N+12NK-6K^2-14K-14$	$4N+10NK-5K^2-13K-12$	2
GWLP 2 (up to 2nd iteration)	$6N+20NK-10K^2-26K-18$	$4N+18NK-9K^2-25K-16$	3
minimal order LP	$4(N-1)$	$2(N-1)$	1
WLP	$6(N-1)$	$4N-6$	1
WPA	$5(N-1)$	$3N-4$	$N-1$
[10]	$4KN-2K(K+1)+4K$	$4KN-2K(K+1)-1$	K
[15]	$\frac{2NK(3K+1)}{2K+1}+5K+1$	$\frac{2NK(3K+1)}{2K+1}+K-2$	K

while the second and third components of (16) represent the higher order as well as zero-order weighted linear prediction terms. On the other hand, the estimate of the WLP is given by [9]

$$\hat{\omega}_{\mathrm{WLP}} = \angle\left(\sum_{m=1}^{N-1}\frac{6}{N^2-1}x_m^*\left[\mathbf{W}\right]_{m,m}x_{m+1}\right). \tag{18}$$

From (16)–(18), we see that the sum involving the main diagonal of $\mathbf{W}$, that is, the first component of (16), is exactly the WLP approach with the same parabolic weights up to a real scalar. While the summations correspond to other diagonals of $\mathbf{W}$ are parabolically weighted computations of the higher-order as well as zero-order autocorrelation lags. This clearly shows how the GWLP 2 generalizes the WLP. Furthermore, the expression of (9) is even more general than (15). Because of these, we call our approach as generalized weighted linear predictor. It is noteworthy to mention that (16) is also related to autocorrelation-based methods of [10], [13], and [15] which use different combinations of the higher-order autocorrelations, and in particular, [15] attempts to find an optimal set of autocorrelation lags for the purpose of fitting a line to their phases. However, one major difference between the GWLP 2 and these autocorrelation-based methods is that the latter are subject to phase unwrapping errors or have limited frequency operation range while the former is free of these demerits.

III. ANALYSIS OF GWLP 2

In this Section, we will investigate the computational complexity, convergence as well as mean and variance of the simplest frequency estimator, namely, GWLP 2.

A. Computational Complexity Analysis

Since the first term of (16) corresponds to (18) and remains a constant during the iterative process, using the WLP as initialization is advantageous in terms of computations in GWLP 2. For the first iteration, including the initialization, GWLP 2 requires two angle calculations, $(6N^2-8N-2)$ real-valued multiplications and $(5N^2-9N)$ real-valued additions. For each of the second or above iteration, an additional complexity of one angle calculation, $4(N-1)(N-2)$ real-valued multiplications and $4(N-1)(N-2)$ real-valued additions are needed. As a result, the computational complexity of the GWLP 2 has order N^2 or $O(N^2)$. To reduce its computational complexity, we can include the two-sided linear prediction terms up to the Kth order, where $K<N$, and ignore the rest, which is analogous to [10], [13], and [15], where only a set of autocorrelations is employed. In so doing, the computational complexity of the GWLP 2 can be reduced to $O(N)$. The computational requirements of the GWLP 2 as well as some conventional low-complexity frequency estimators [7], [9], [10], [15] are tabulated in Table I. Note that $K=N-3$ roughly refers to the GWLP 2 with no approximation. We can see that although GWLP 2 is more computationally demanding than the minimal order LP, WLP, and WPA, its complexity is comparable to those of [10] and [15].

B. Convergence Analysis

The following analysis is to show that if $\hat{\omega}_2$ is sufficiently close to ω, iterative application of (13) will converge to the true frequency as the number of samples goes to infinity. Basically, we have followed [34] to prove the convergence of the algorithm. Note that similar convergence analysis is also found in [35] which deals with iterative estimation of a single real tone. Let $g(\hat{\omega}_2)=\mathbf{X_2}^H\mathbf{W}\mathbf{X_1}$, where $\mathbf{W}$ in (12) is now a function of $\hat{\omega}_2$. To prove the convergence of the GWLP 2, it is sufficient [34] to show that the following two conditions are satisfied:

i) $|\angle g(\hat{\omega}_2)-\omega|<cN^{-\varepsilon}$

ii) $|\angle g(\hat{\omega}_a)-\angle g(\hat{\omega}_b))|<K|\hat{\omega}_a-\hat{\omega}_b|, \forall\hat{\omega}_a,\hat{\omega}_b\in\mathcal{A}_N$, and $0<K<1$

almost surely as $N\to\infty$, where $\mathcal{A}_N=\{\hat{\omega}_2:|\hat{\omega}_2-\omega|<cN^{-\varepsilon}\}$, $c>0$, and $\varepsilon>1$ are fixed. If these two conditions are satisfied, $\hat{\omega}_2$ will iterate to a unique fixed point, namely, ω, upon convergence. In our study, an alternative form for the second condition is employed:

ii) $|(d\angle g(\hat{\omega}_2))/(d\hat{\omega}_2)|<1\Rightarrow|\Im((g'(\hat{\omega}_2))/(g(\hat{\omega}_2)))|<1, \forall\hat{\omega}_2\in\mathcal{A}_N$

where $\Im$ denotes the imaginary part and $g'(\hat{\omega}_2)=(dg(\hat{\omega}_2))/(d\hat{\omega}_2)$.

In Appendix A, we have shown that

$$\angle g(\hat{\omega}_2) = \hat{\omega}_2 + O(N^{-\varepsilon}) \tag{19}$$

which implies that the first condition is satisfied. Moreover, $g(\hat{\omega}_2)$ and $g'(\hat{\omega}_2)$ have been derived as (see Appendix A)

$$g(\hat{\omega}_2) = O(N^3)e^{j(\hat{\omega}_2 + O(N^{-\varepsilon}))} \tag{20}$$

and

$$g'(\hat{\omega}_2) = O(N^4)e^{j(\hat{\omega}_2 + O(N^{-\varepsilon}))}. \tag{21}$$

Using (20) and (21), the second condition can be proved as follows:

$$\begin{aligned}\left|\Im\left(\frac{g'(\hat{\omega}_2)}{g(\hat{\omega}_2)}\right)\right| &= |O(N)\sin(O(N^{-\varepsilon}))| \\ &= |O(N^{1-\varepsilon})| < 1.\end{aligned} \tag{22}$$

As a result, $\hat{\omega}_2$ will approach ω upon convergence as N tends to infinity.

C. Mean and Variance Analysis

The bias and variance of the frequency estimate for the GWLP 2 are now derived. It is expected that the mean analysis will also hold for the GWLP 1 for sufficiently high SNRs and/or large N such that $\hat{\rho} \to e^{j\hat{\omega}_1}$ and $\angle(\mathbf{X_2}^H\mathbf{W}\mathbf{X_2}) \to 0$. To simplify the derivation, we assume that the ideal weighting matrix of (10) is used in (13). Note that this assumption becomes valid when the frequency estimate approaches the true value of ω, which is anticipated to occur at sufficiently large SNR and/or data length conditions. Taking the expected value of (13) with the use of the ideal weighting matrix gives (see Appendix B)

$$E\{\hat{\omega}_2\} = \omega + E\left\{\angle\left(e^{-j\omega}\mathbf{X_2}^H\mathbf{W}\mathbf{X_1}\right)\right\} \approx \omega \tag{23}$$

which indicates the approximately unbiasedness of the algorithm. In Appendix B, we have also derived the variance of $\hat{\omega}_2$, denoted by $\mathrm{var}(\hat{\omega}_2)$, as

$$\mathrm{var}(\hat{\omega}_2) \approx \frac{6}{\mathrm{SNR}\, N(N^2-1)} + \frac{18\sigma^2}{\mathrm{SNR}^2 N^2(N+1)^2} \tag{24}$$

where $\mathrm{SNR} = A^2/\sigma^2$. On the other hand, the CRLB for single frequency estimation in white Gaussian noise, denoted by $\mathrm{CRLB}(\omega)$, is given by [17]

$$\mathrm{CRLB}(\omega) = \frac{6}{\mathrm{SNR}\, N(N^2-1)}. \tag{25}$$

Comparing (24) and (25), we see that for a fixed SNR, the estimation accuracy of $\hat{\omega}_2$ in the presence of white Gaussian noise approaches the CRLB when the data length is sufficiently large. It is worthy to note that (24) can hold for noises with other probability density functions as well because its derivation assumes zero-mean white noise only.

IV. Simulation Results

Computer simulations had been carried out to evaluate the frequency estimation performance of the three proposed algorithms in the presence of complex white Gaussian noise by comparing with minimal order LP, WLP, and WPA as well as CRLB. We used 2 iterations in all proposed algorithms because no significant improvement was observed for more iterations. The signal power was unity, which corresponded to $A = 1$ and we scaled the noise sequence to produce different SNRs while we fixed the phase parameter as $\phi = 1$. All results provided were averages of 2000 independent runs.

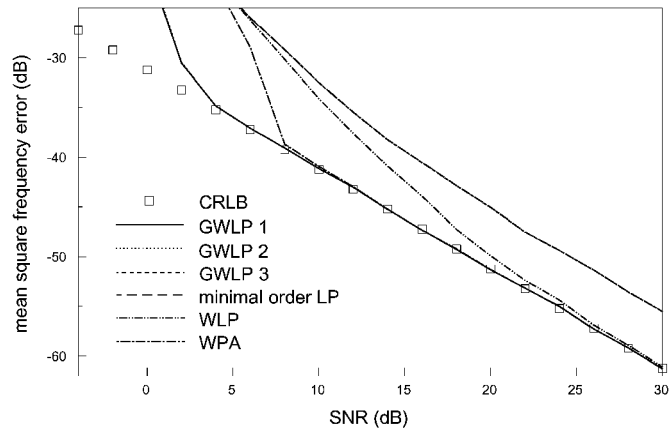

Fig. 1. Mean square frequency errors versus SNR at $N = 20$ and $\omega = 0.1\pi$.

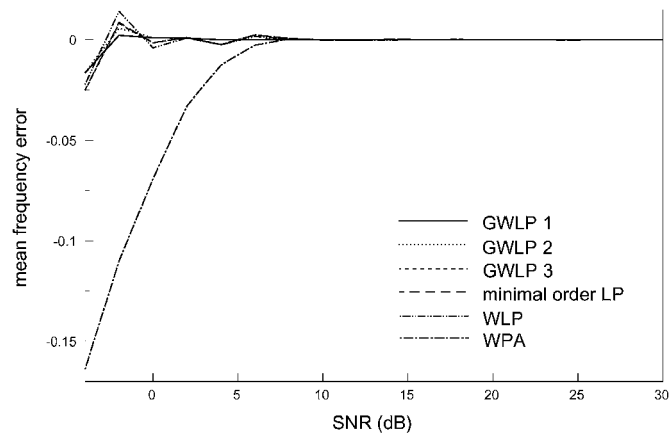

Fig. 2. Mean frequency errors versus SNR at $N = 20$ and $\omega = 0.1\pi$.

Fig. 1 shows the mean square frequency error (MSFE) performance versus SNR at $\omega = 1$ and $N = 20$, which corresponded to a small data length scenario. It is seen that the GWLP 1 and GWLP 2 performed almost identically and their MSFEs attained the CRLB for SNR ≥ 4 dB. On the other hand, the estimation performance of the GWLP 3 and minimal order LP was very similar, which had the largest MSFEs. We believe the inferiority of GWLP 3 was due to the discrepancy in the magnitude of both sides of (14) as pointed out in Section II-B. Although the WPA could attain the CRLB as well, its threshold SNR was higher than those of the GWLP 1 and GWLP 2, which implies a smaller SNR operation range. While it is seen that the WLP was optimum only for very high SNR conditions. The corresponding mean frequency errors, which were obtained by subtracting ω from the mean frequency estimates, are shown in Fig. 2. We observe that the biases in all the methods were negligible for sufficiently high SNRs, which demonstrates the approximately unbiasedness property of the proposed methods and indicates

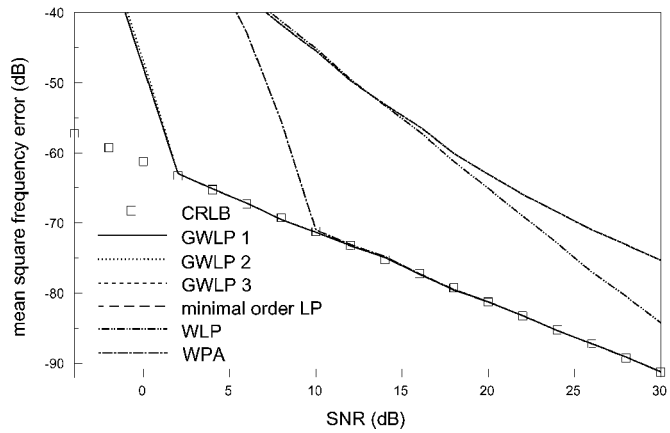

Fig. 3. Mean square frequency errors versus SNR at $N = 200$ and $\omega = 0.1\pi$.

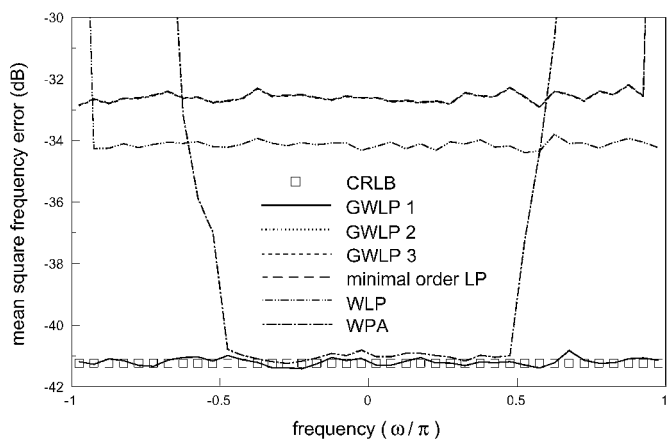

Fig. 4. Mean square frequency errors versus ω at SNR $= 10$ dB and $N = 20$.

that the MSFEs were mainly due to variances of the frequency estimates.

We repeated the first test for $N = 200$, which corresponded to a large data length scenario, and the MSFEs are plotted in Fig. 3. Similar findings were observed, in particular, the GWLP 1 and GWLP 2 could attain the CRLB with the largest SNR operation range. Note that we have not included the corresponding mean frequency error results because they were similar to those in Fig. 2.

Fig. 4 shows the MSFEs of different frequency estimators versus frequency at SNR $= 10$ dB and $N = 20$. We see that both GWLP 1 and GWLP 2 achieved optimum performance for the admissible frequency range while the optimality of the WPA only held for $\omega \in (-0.5\pi, 0.5\pi)$. Furthermore, the GWLP 3 and minimal order LP performed almost identically and were inferior to the suboptimal WLP. The above test was repeated for $N = 200$ and the results are shown in Fig. 5, and the findings were similar to those in Fig. 4.

Figs. 6–11 plot the frequency versus SNR contours of MSFE for the GWLP 1, GWLP 2, GWLP 3, minimal order LP, WLP, WPA, respectively, at $N = 20$ in order to investigate the threshold performance in more detail. We can see that GWLP 1 and GWLP 2 had the best threshold performance while that of WPA was the poorest. The corresponding contour plots at $N = 200$ were also produced which are shown in Figs. 12–17, and we had similar observations.

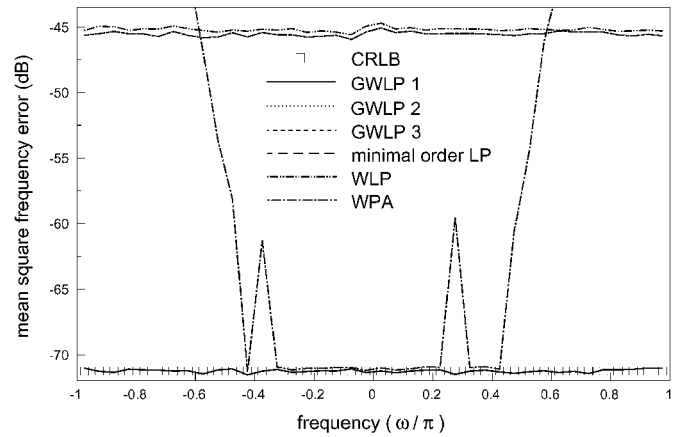

Fig. 5. Mean square frequency errors versus ω at SNR $= 10$ dB and $N = 200$.

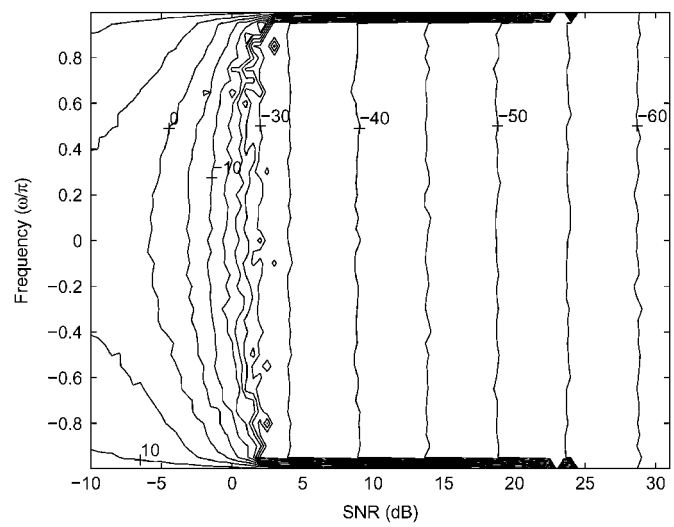

Fig. 6. Contour plot of GWLP 1 at $N = 20$.

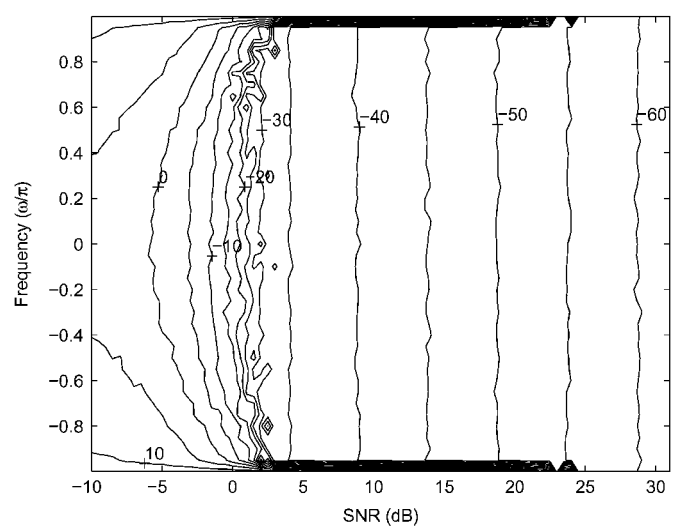

Fig. 7. Contour plot of GWLP 2 at $N = 20$.

In Figs. 18–21, the performance of different approximations of GWLP 2 was evaluated, that is, we only employed the linear prediction terms up to the Kth order. The results of $K = N/4$,

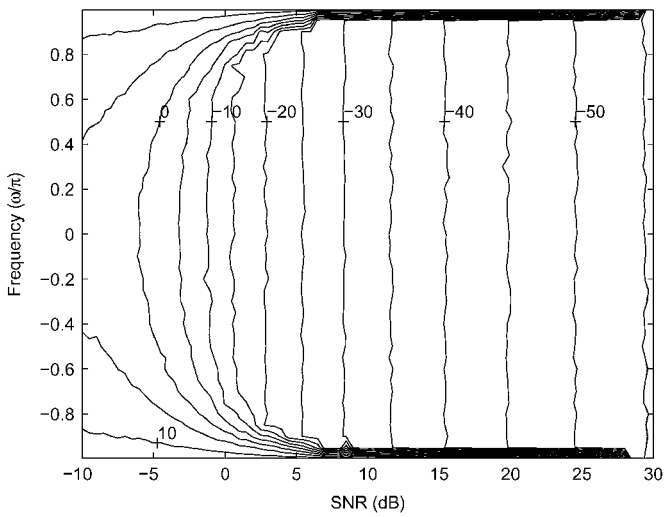

Fig. 8. Contour plot of GWLP 3 at $N = 20$.

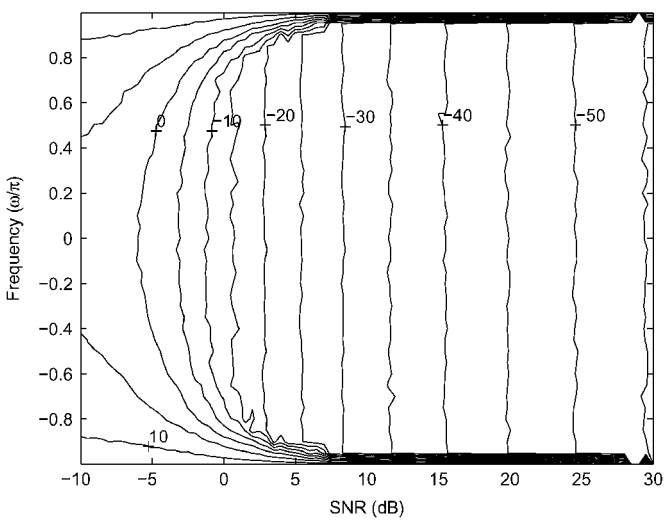

Fig. 9. Contour plot of minimal order LP at $N = 20$.

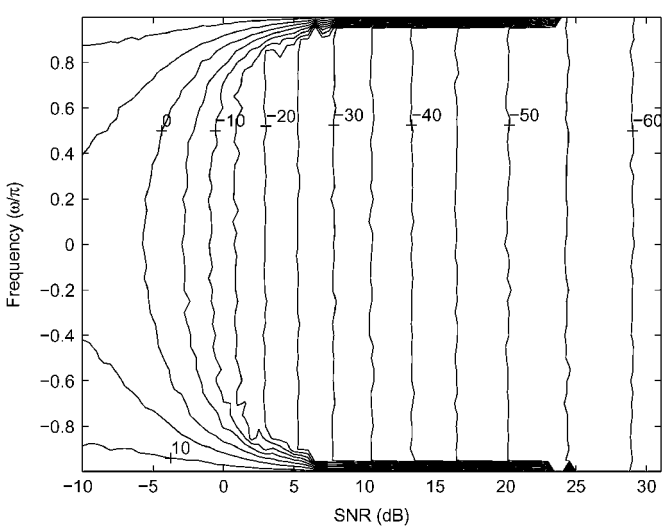

Fig. 10. Contour plot of WLP at $N = 20$.

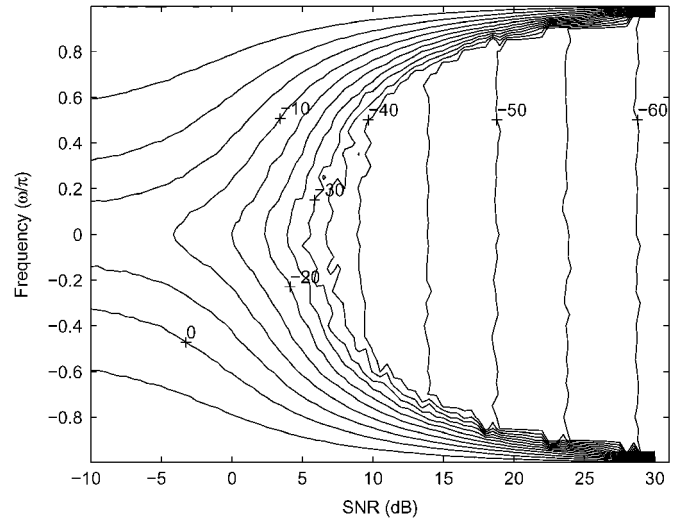

Fig. 11. Contour plot of WPA at $N = 20$.

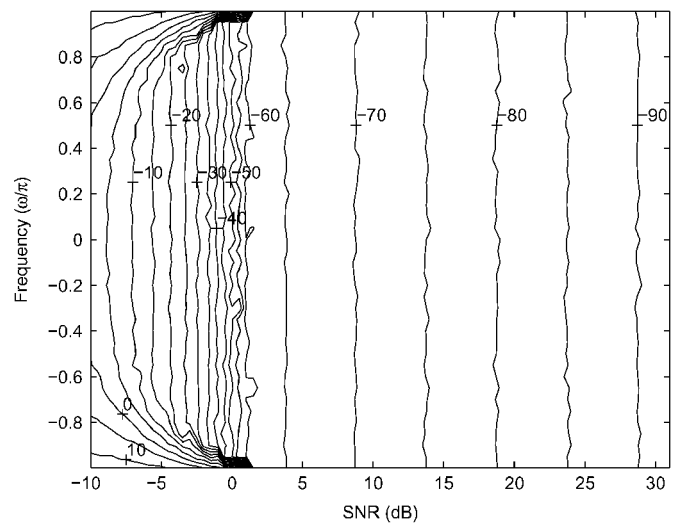

Fig. 12. Contour plot of GWLP 1 at $N = 200$.

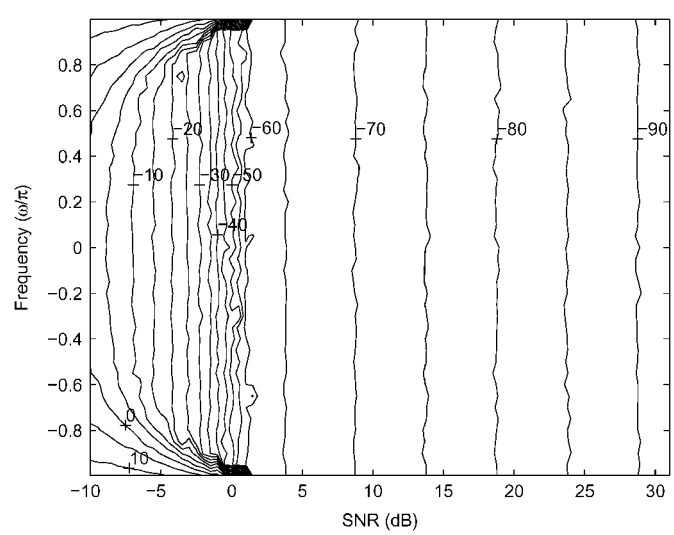

Fig. 13. Contour plot of GWLP 2 at $N = 200$.

$K = N/2$, $K = 3N/4$, and $K = N-3$, which corresponded to no approximation, and the baseline algorithm of WLP were given. The simulation settings of Figs. 18–21 were identical to those of Figs. 1, 3, 4, and 5, respectively. From the figures, we see that using $K = N/2$ and $K = 3N/4$ had comparable MSFEs with those of the exact version, except that the GWLP 2

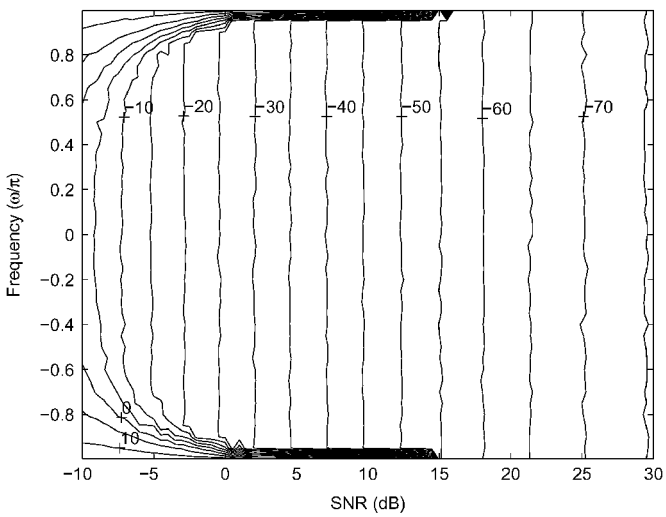

Fig. 14. Contour plot of GWLP 3 at $N = 200$.

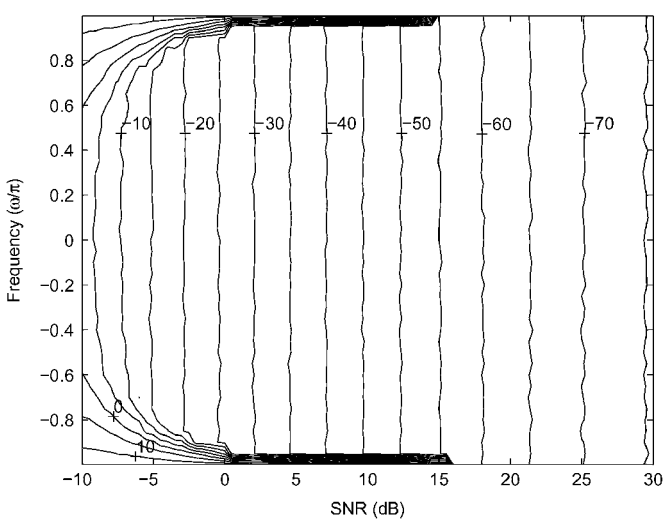

Fig. 15. Contour plot of minimal order LP at $N = 200$.

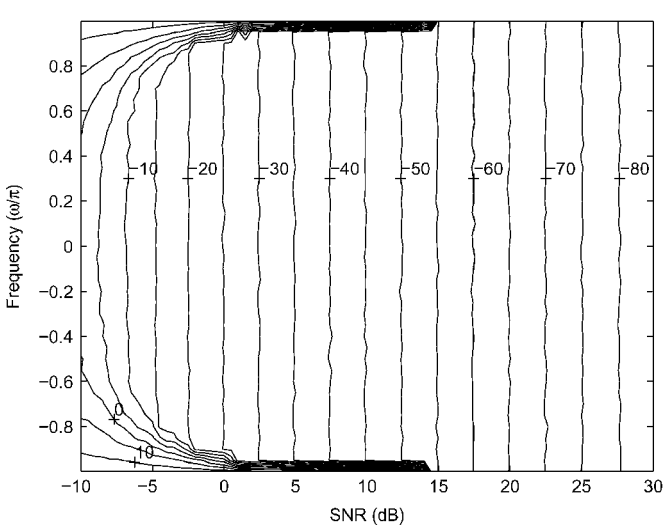

Fig. 16. Contour plot of WLP at $N = 200$.

with $K = N/2$ had a larger threshold SNR. It is also observed that the estimation accuracy increased with K.

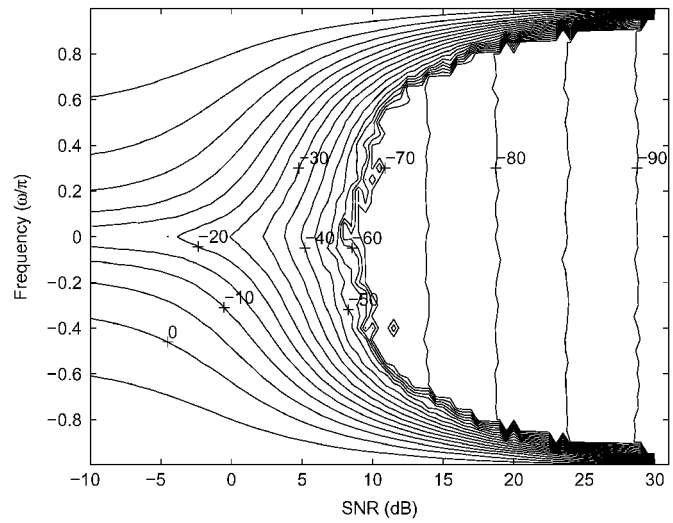

Fig. 17. Contour plot of WPA at $N = 200$.

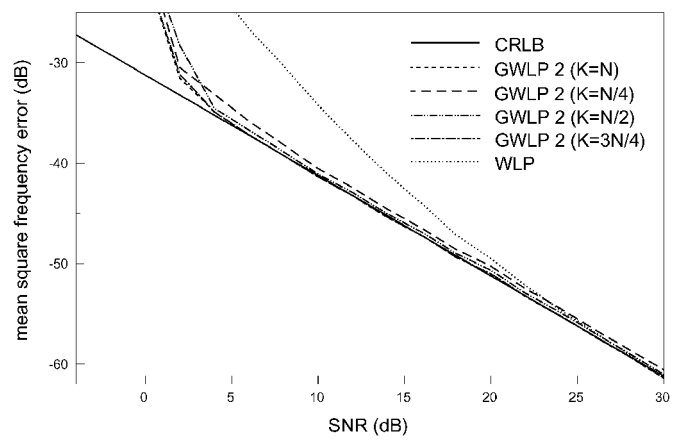

Fig. 18. Mean square frequency errors of different approximations of GWLP 2 versus SNR at $N = 20$ and $\omega = 0.1\pi$.

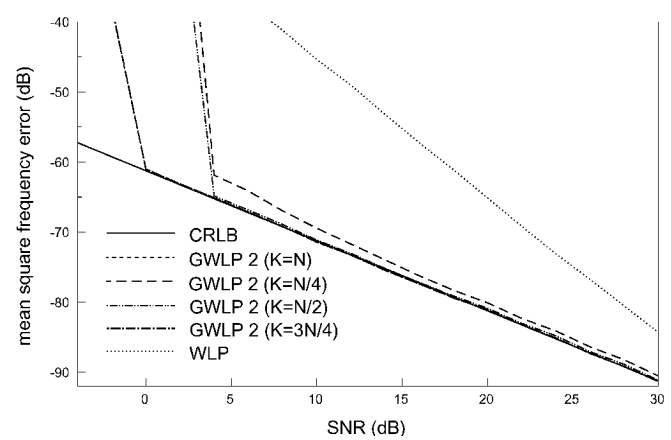

Fig. 19. Mean square frequency errors of different approximations of GWLP 2 versus SNR at $N = 200$ and $\omega = 0.1\pi$.

V. Conclusion

Three computationally simple frequency estimation algorithms, viz. GWLP 1, GWLP 2, and GWLP 3, have been developed for a complex sinusoid embedded in white noise. The GWLP 1 is the fundamental algorithm which is derived straightforwardly using the ideas of linear prediction and

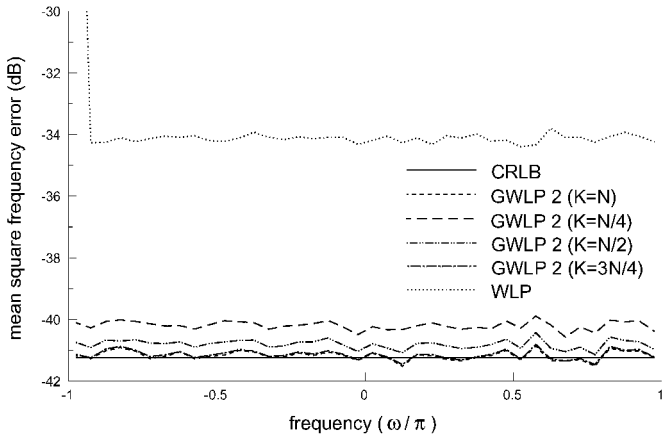

Fig. 20. Mean square frequency errors of different approximations of GWLP 2 versus ω at SNR $= 10$ dB and $N = 20$.

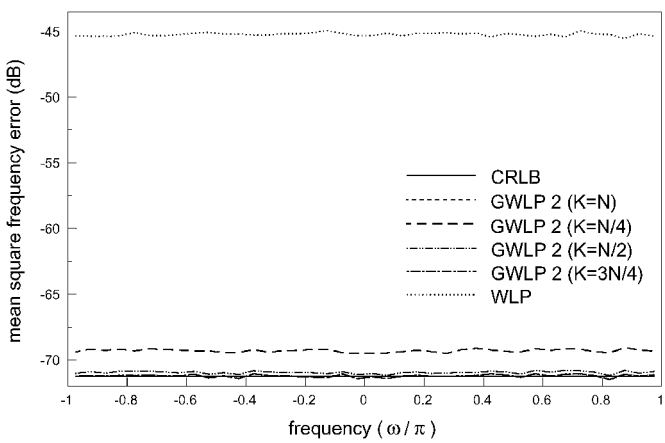

Fig. 21. Mean square frequency errors of different approximations of GWLP 2 versus ω at SNR $= 10$ dB and $N = 200$.

weighted least squares. The other two algorithms are the approximate realizations of the GWLP 1 and they involve fewer computations. The proposed approach can be considered as a generalized version of Kay's weighted linear predictor frequency estimator. In particular, computational requirement, convergence as well as mean and variance analysis of the GWLP 2 are studied. It is shown that the GWLP 1 and GWLP 2 can provide optimum estimation accuracy while the GWLP 3 is a suboptimum estimator. As a result, the GWLP 2 is the best among the three estimators in terms of estimation performance and implementation complexity.

APPENDIX A

In this Appendix, we prove that if $|\hat{\omega}_2 - \omega| < cN^{-\varepsilon}$, the frequency estimate of GWLP 2 will converge to the true frequency for infinite data samples. Expanding $g(\hat{\omega}_2) = \mathbf{X_2}^H \mathbf{W} \mathbf{X_1}$ yields

$$g(\hat{\omega}_2) = \mathbf{S_2}^H \mathbf{W} \mathbf{S_1} + \mathbf{S_2}^H \mathbf{W} \mathbf{Q_1} + \mathbf{Q_2}^H \mathbf{W} \mathbf{S_1} + \mathbf{Q_2}^H \mathbf{W} \mathbf{Q_1} \quad \text{(A1)}$$

where $\mathbf{S_1}=[s_N, s_{N-1}, \ldots, s_2]^T$, $\mathbf{S_2} = [s_{N-1}, s_{N-2}, \cdots, s_1]^T$, $\mathbf{Q_1} = [q_N, q_{N-1}, \ldots, q_2]^T$, and $\mathbf{Q_2} = [q_{N-1}, q_{N-2}, \ldots, q_1]^T$ such that $\mathbf{X}_1 = \mathbf{S}_1 + \mathbf{Q}_1$ and $\mathbf{X}_2 = \mathbf{S}_2 + \mathbf{Q}_2$. The terms in (A1) are analyzed as follows. The first term of (A1) is

$$\begin{aligned}
\mathbf{S}_2^H \mathbf{W} \mathbf{S}_1 &= \frac{A^2}{N} \sum_{m=1}^{N-1} \sum_{n=1}^{N-1} \Big(e^{j(m-n+1)\omega} (N - \max(m,n)) \\
&\qquad \times \min(m,n) e^{j(n-m)\hat{\omega}_2} \Big) \\
&= \frac{A^2 e^{j\omega}}{N} \sum_{m=1}^{N-1} \sum_{n=1}^{N-1} \Big(e^{j(m-n)(\omega-\hat{\omega}_2)} (N - \max(m,n)) \\
&\qquad \times \min(m,n) \Big) \\
&= O(N^3) e^{j\omega} \\
&= O(N^3) e^{j(\hat{\omega}_2 + O(N^{-\varepsilon}))}
\end{aligned} \quad \text{(A2)}$$

where $\max(m,n) = m$ if $m > n$ and it is equal to n otherwise. Considering the second and third terms of (A1) together, we have

$$\begin{aligned}
&\mathbf{Q}_2^H \mathbf{W} \mathbf{S}_1 + \mathbf{S}_2^H \mathbf{W} \mathbf{Q}_1 \\
&\quad = \frac{A}{N} \sum_{m=1}^{N-1} \sum_{n=1}^{N-1} \Big(q_{N-m}^* e^{j((N-n+1)\omega+\phi)} \\
&\qquad + q_{N-n+1} e^{-j((N-m)\omega+\phi)} \Big) \\
&\qquad \times (N - \max(m,n)) \min(m,n) e^{j(n-m)\hat{\omega}_2} \\
&\quad = z_{2,1} + z_{2,2} + z_{2,3}
\end{aligned} \quad \text{(A3)}$$

where $z_{2,1}$ is given by

$$\begin{aligned}
z_{2,1} &= \frac{A}{N} \sum_{m=1}^{N-2} \sum_{n=1}^{N-2} \Big(q_{N-m}^* e^{j((N-n)\omega+\phi)} \\
&\qquad + q_{N-n} e^{-j((N-m)\omega+\phi)} \Big) \\
&\qquad \times (N - \max(m,n+1)) \min(m,n+1) e^{j(n+1-m)\hat{\omega}_2} \\
&= e^{j\hat{\omega}_2} z_{2,1,1} + z_{2,1,2} - z_{2,1,3}
\end{aligned}$$

with

$$\begin{aligned}
z_{2,1,1} &= \frac{A}{N} \sum_{m=1}^{N-2} \sum_{n=1}^{N-2} \Big(q_{N-m}^* e^{j((N-n)\omega+\phi)} \\
&\qquad + q_{N-n} e^{-j((N-m)\omega+\phi)} \Big) \\
&\qquad \times (N - \max(m,n)) \min(m,n) e^{j(n-m)\hat{\omega}_2} \\
z_{2,1,2} &= \frac{A}{N} \sum_{m=1}^{N-2} \sum_{n=1}^{m-1} \Big(q_{N-m}^* e^{j((N-n)\omega+\phi)} \\
&\qquad + q_{N-n} e^{-j((N-m)\omega+\phi)} \Big) \\
&\qquad \times (N - m) e^{j(n+1-m)\hat{\omega}_2}
\end{aligned}$$

and

$$z_{2,1,3} = \frac{A}{N} \sum_{m=1}^{N-2} \sum_{n=m}^{N-2} \Big(q_{N-m}^* e^{j((N-n)\omega+\phi)} + q_{N-n} e^{-j((N-m)\omega+\phi)} \Big) m e^{j(n+1-m)\hat{\omega}_2}$$

while the terms $z_{2,2}$ and $z_{2,3}$ are

$$z_{2,2} = \frac{A}{N} \sum_{n=1}^{N-1} \Big(q_1^* e^{j((N-n+1)\omega+\phi)} + q_{N-n+1} e^{-j(\omega+\phi)} \Big) n e^{j(n-N+1)\hat{\omega}_2}$$

and

$$z_{2,3} = \frac{A}{N}\sum_{m=1}^{N-2}\Big(q_{N-m}^{*}e^{j(N\omega+\phi)} + q_N e^{-j((N-m)\omega+\phi)}\Big)(N-m)e^{j(1-m)\hat{\omega}_2}.$$

We notice that $z_{2,1,1}$ is real and has order of $N^{2.5}(\log N)^{0.5}$. Furthermore, $|z_{2,1,2}-z_{2,1,3}+z_{2,2}+z_{2,3}|$ is $O(N^{1.5}(\log N)^{0.5})$. Combining these results, (A3) is simplified as

$$\mathbf{Q}_2^H\mathbf{W}\mathbf{S}_1 + \mathbf{S}_2^H\mathbf{W}\mathbf{Q}_1 = O(N^{2.5}(\log N)^{0.5})e^{j(\hat{\omega}_2+O(N^{-1}))}. \tag{A4}$$

The last term of (A1) is

$$\begin{aligned}\mathbf{Q}_2^H\mathbf{W}\mathbf{Q}_1 &= \sum_{m=1}^{N-1}\sum_{n=1}^{N-1}\Big(q_{N-m}^{*}(N-\max(m,n)) \times \min(m,n)q_{N-n+1}e^{j(n-m)\hat{\omega}_2}\Big)\\ &= e^{j\hat{\omega}_2}z_{3,1} + z_{3,2} + z_{3,3}\end{aligned} \tag{A5}$$

where $z_{3,1}$ is

$$\begin{aligned}z_{3,1} &= \frac{1}{N}\sum_{m=1}^{N-2}\sum_{n=1}^{N-2}\Big(q_{N-m}^{*}(N-\max(m,n+1)) \times \min(m,n+1)q_{N-n}e^{j(n-m)\hat{\omega}_2}\Big)\\ &= z_{3,1,1} + z_{3,1,2} + z_{3,1,3}\end{aligned}$$

with

$$z_{3,1,1} = \frac{1}{N}\sum_{m=1}^{N-2}\sum_{n=1}^{N-2}\Big((N-\max(m,n))\min(m,n) \times q_{N-m}^{*}q_{N-n}e^{j(n-m)\hat{\omega}_2}\Big)$$

$$z_{3,1,2} = \frac{1}{N}\sum_{m=1}^{N-2}\sum_{n=1}^{m-1}(N-m)q_{N-m}^{*}q_{N-n}e^{j(n-m)\hat{\omega}_2}$$

and

$$z_{3,1,3} = \frac{1}{N}\sum_{m=1}^{N-2}\sum_{n=m}^{N-2}mq_{N-m}^{*}q_{N-n}e^{j(n-m)\hat{\omega}_2}$$

while the terms $z_{3,2}$ and $z_{3,3}$ are

$$z_{3,2} = \frac{q_1}{N}^{*}\sum_{n=1}^{N-1}nq_{N-n+1}e^{j(n-N+1)\hat{\omega}_2}$$

and

$$z_{3,3} = \frac{q_N}{N}\sum_{m=1}^{N-1}(N-m)q_{N-m}^{*}e^{j(1-m)\hat{\omega}_2}.$$

We note that $z_{3,1,1}$ is real and applying the following result [35]:

$$\left|\sum_{k=1}^{N}q_k\exp(-j\omega k)\right| = O((N\log N)^{0.5})$$

$z_{3,1,1}$ is $O(N^2\log N)$. Similarly, both $|z_{3,1,2}|$ and $|z_{3,1,3}|$ are $O(N\log N)$ while $|z_{3,2}|$ and $|z_{3,3}|$ are $O(N\log N)^{0.5}$. Therefore, $|(z_{3,1,2}+z_{3,1,3})e^{j\hat{\omega}_2}+z_{3,2}+z_{3,3}|$ is $O(N\log N)$. Combining these results, (A5) is simplified as

$$\mathbf{Q}_2^H\mathbf{W}\mathbf{Q}_1 = O(N^2\log N)e^{j(\hat{\omega}_2+O(N^{-1}))}. \tag{A6}$$

With the use of (A2), (A4), and (A6), the magnitude and phase angle of $g(\hat{\omega}_2)$ can be calculated as (A7) and (A8), shown at the bottom of the page. By applying the following formulas with $\alpha \to 0$:

$$\begin{aligned}\sin(\alpha) &= \alpha + O(\alpha^3)\\ \cos(\alpha) &= 1 + O(\alpha^2)\\ \tan^{-1}(\alpha) &= \alpha + O(\alpha^3)\end{aligned}$$

the second component in (A8) can be simplified to $O(N^{-\varepsilon})$, and this implies

$$\angle g(\hat{\omega}_2) = \hat{\omega}_2 + O(N^{-\varepsilon}) \tag{A9}$$

which is (19). From (A7) and (A9), we get

$$g(\hat{\omega}_2) = O(N^3)e^{j(\hat{\omega}_2+O(N^{-\varepsilon}))}. \tag{A10}$$

In a similar manner, it can be shown that

$$g'(\hat{\omega}_2) = O(N^4)e^{j(\hat{\omega}_2+O(N^{-\varepsilon}))}. \tag{A11}$$

Utilizing (A9)–(A11), the two conditions for convergence of GWLP 2 are proved.

APPENDIX B

In this Appendix, we will prove that the frequency estimate of the GWLP 2 is approximately unbiased and produce its variance expression. With the use of (13) and (A1), the frequency estimate of the GWLP 2 is expressed as

$$\hat{\omega}_2 = \omega + \Delta\omega \tag{B1}$$

$$\begin{aligned}|g(\hat{\omega}_2)| &= \sqrt{\left|O(N^3)e^{j(\hat{\omega}_2+O(N^{-\varepsilon}))} + O(N^{2.5}(\log N)^{0.5})e^{j(\hat{\omega}_2+O(N^{-1}))} + O(N^2\log N)e^{j(\hat{\omega}_2+O(N^{-1}))}\right|^2}\\ &= O(N^3)\end{aligned} \tag{A7}$$

and

$$\begin{aligned}\angle g(\hat{\omega}_2) &= \hat{\omega}_2\\ &+ \tan^{-1}\Big(\frac{O(N^3)\sin(O(N^{-\varepsilon})) + O(N^{2.5}(\log N)^{0.5})\sin(O(N^{-1})) + O(N^{1.5}\log N)\sin(O(N^{-1}))}{O(N^3)\cos(O(N^{-\varepsilon})) + O(N^{2.5}(\log N)^{0.5})\cos(O(N^{-1})) + O(N^{1.5}\log N)\cos(O(N^{-1}))}\Big).\end{aligned} \tag{A8}$$

where

$$\Delta\omega = \angle\left(e^{-j\omega}\left(\mathbf{S}_2^H\mathbf{W}\mathbf{S}_1 + \mathbf{S}_2^H\mathbf{W}\mathbf{Q}_1 + \mathbf{Q}_2^H\mathbf{W}\mathbf{S}_1 + \mathbf{Q}_2^H\mathbf{W}\mathbf{Q}_1\right)\right) \quad \text{(B2)}$$

is the error in the frequency estimate.

We first notice that $e^{-j\omega}\mathbf{S}_2^H\mathbf{W}\mathbf{S}_1 = \mathbf{S}_2^H\mathbf{W}\mathbf{S}_2$ is real. Let the real and imaginary parts of $e^{-j\omega}(\mathbf{S}_2^H\mathbf{W}\mathbf{Q}_1 + \mathbf{Q}_2^H\mathbf{W}\mathbf{S}_1 + \mathbf{Q}_2^H\mathbf{W}\mathbf{Q}_1)$ be λ_r and λ_i, respectively, and noting that $\mathbf{W}$ here corresponds to the ideal weighting matrix of (10). Assuming that the estimation error is sufficiently small, we use Taylor's series to expand $\Delta\omega$ around 0 up to the first-order term to obtain [7]

$$\Delta\omega = \tan^{-1}\left(\frac{\lambda_i}{e^{-j\omega}\mathbf{S}_2^H\mathbf{W}\mathbf{S}_1 + \lambda_r}\right) \approx \frac{\lambda_i}{e^{-j\omega}\mathbf{S}_2^H\mathbf{W}\mathbf{S}_1} = \frac{\lambda_i}{\mathbf{S}_2^H\mathbf{W}\mathbf{S}_2}. \quad \text{(B3)}$$

The denominator of (B3) can also be expressed as

$$\mathbf{S_2}^H\mathbf{W}\mathbf{S_2} = \mathrm{tr}(\mathbf{W}\mathbf{T}) \quad \text{(B4)}$$

where

$$\mathbf{T} = \mathbf{S_2}\mathbf{S_2}^H \quad \text{(B5)}$$

and tr represents the trace operation. The (m,n) entry of $\mathbf{T}$ is evaluated as $[\mathbf{T}]_{m,n} = A^2e^{j(n-m)\omega}$. With the use of (12) and $\mathrm{tr}(\mathbf{A}^T\mathbf{B}) = \sum_m\sum_n[\mathbf{A}]_{m,n}[\mathbf{B}]_{m,n}$ [36], $\mathbf{S_2}^H\mathbf{W}\mathbf{S_2}$ is calculated as

$$\begin{aligned}\mathbf{S_2}^H\mathbf{W}\mathbf{S_2} =& \mathrm{tr}(\mathbf{W}\mathbf{T})\\ =& \sum_{m=1}^{N-1}\sum_{n=1,m\neq n}^{N-1}[\mathbf{W}]_{m,n}[\mathbf{T}]_{n,m}\\ &+ \sum_{m=1}^{N-1}[\mathbf{W}]_{m,m}[\mathbf{T}]_{m,m}\\ =& \frac{2A^2\sum_{m=1}^{N-1}\sum_{n=m+1}^{N-1}(Nm-mn)}{N} + A^2\mathrm{tr}(\mathbf{W})\\ =& \frac{2A^2\sum_{m=1}^{N-1}\sum_{n=m+1}^{N-1}(Nm-mn)}{N}\\ &+ A^2\sum_{m=1}^{N-1}\frac{(Nm-m^2)}{N}\\ =& \frac{(N^2-1)NA^2}{12}. \end{aligned} \quad \text{(B6)}$$

To investigate the numerator of (B3), we decompose λ_i into $\lambda_i = \lambda_{i,1}+\lambda_{i,2}$ where $\lambda_{i,1} = \Im\{e^{-j\omega}(\mathbf{S}_2^H\mathbf{W}\mathbf{Q}_1+\mathbf{Q}_2^H\mathbf{W}\mathbf{S}_1)\}$, and $\lambda_{i,2} = \Im\{e^{-j\omega}(\mathbf{Q}_2^H\mathbf{W}\mathbf{Q}_1)\}$. Expanding q_n as $q_n = Au_n + jAv_n$, $\lambda_{i,1}$, and $\lambda_{i,2}$ are computed as

$$\lambda_{i,1} = A^2\sum_{i=0}^{N-1}(v_{N-i}\cos((N-i)\omega+\phi) - u_{N-i}\sin((N-i)\omega+\phi))\frac{N-2i-1}{2} \quad \text{(B7)}$$

and

$$\lambda_{i,2} = A^2\left(\sum_{m=1}^{N-2}\sum_{n=m+1}^{N-1}\frac{m(N-n)(t_1+t_2+t_3+t_4)}{N} + t_5\right) \quad \text{(B8)}$$

where (see the equations at the bottom of the page). Since $\{u_n\}$ and $\{v_n\}$ are uncorrelated, it is easily seen from (B7) and (B8) that $E\{\lambda_i\} = E\{\lambda_{i,1}\} = E\{\lambda_{i,2}\} = 0$, which implies $E\{\Delta\omega\} \approx 0$ or the approximately unbiasedness of $\hat{\omega}_2$.

To compute the variance of $\hat{\omega}_2$, we use (B3) again, as follows:

$$\mathrm{var}(\hat{\omega}_2) = E\{(\Delta\omega)^2\} \approx \frac{E\{\lambda_i^2\}}{\left(\mathbf{S}_2^H\mathbf{W}\mathbf{S}_2\right)^2}. \quad \text{(B9)}$$

Since $\lambda_{i,1}$ and $\lambda_{i,2}$ are uncorrelated, we have

$$E\{\lambda_i^2\} = E\{\lambda_{i,1}^2\} + E\{\lambda_{i,2}^2\}. \quad \text{(B10)}$$

With the use of (B7) and (B8), $E\{\lambda_{i,1}^2\}$ and $E\{\lambda_{i,2}^2\}$ are calculated as

$$E\{\lambda_{i,1}^2\} = A^2\sigma^2\sum_{i=1}^{N-1}\left(\frac{N-2i-1}{2}\right)^2 = \frac{A^2N(N^2-1)}{24}\sigma^2 \quad \text{(B11)}$$

and

$$\begin{aligned}E\{\lambda_{i,2}^2\} =& A^2E\left\{\left(\sum_{m=1}^{N-2}\sum_{n=m+1}^{N-1}\frac{m(N-n)(t_1+t_2+t_3+t_4)}{N}\right)^2\right\}\\ &+ A^2E\left\{2t_5\sum_{m=1}^{N-2}\sum_{n=m+1}^{N-1}\frac{m(N-n)(t_1+t_2+t_3+t_4)}{N} + t_5^2\right\}\\ =& A^2\sigma^4\left(\frac{N^4-1}{60N} + \frac{(N-2)(N-1)(2N^2+N-11)}{120N} - \frac{(N-1)(N-2)(N-3)(N+1)}{30N}\right)\\ =& \frac{A^2(N-1)^2}{8}\sigma^4. \end{aligned} \quad \text{(B12)}$$

Substituting (B6) and (B10)–(B12) into (B9) yields (24).

$$\begin{aligned}t_1 &= (u_{N-m}v_{N-n+1} - v_{N-m}u_{N-n+1})\cos((n-m-1)\omega)\\ t_2 &= (u_{N-m}u_{N-n+1} - v_{N-m}v_{N-n+1})\sin((n-m-1)\omega)\\ t_3 &= (u_mv_{n+1} - v_mu_{n+1})\cos((n-m+1)\omega)\\ t_4 &= (u_mu_{n+1} - v_mv_{n+1})\sin((n-m+1)\omega)\\ t_5 &= \sum_{i=1}^{N-1}\frac{i(N-i)[(u_{N-i}v_{N-i+1} - v_{N-i}u_{N-i+1})\cos(\omega) - (u_{N-i}u_{N-i+1} + v_{N-i}v_{N-i+1})\sin(\omega)]}{N}.\end{aligned}$$

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their careful reading and insightful comments, which significantly enhanced the quality of this paper.

REFERENCES

[1] R. Prony, "Essa: Experimentale et analytique," in *J. Ecole Polytechnique* Paris, France, 1795, pp. 24–76.
[2] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Upper Saddle River, NJ: Prentice-Hall, 1997.
[3] G. Zhu and Y. Hua, "Quantitative NMR signal analysis by an iterative quadratic maximum likelihood method," *Chem. Phys. Lett.*, vol. 264, pp. 424–428, 1997.
[4] S. L. Marple, *Digital Spectral Analysis with Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
[5] S. M. Kay, *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
[6] B. G. Quinn and E. J. Hannan, *The Estimation and Tracking of Frequency*. Cambridge, U.K.: Cambridge Univ. Press, 2001.
[7] G. W. Lank, I. S. Reed, and G. E. Pollon, "A semicoherent detection and Doppler estimation statistic," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 9, no. 2, pp. 151–165, Mar. 1973.
[8] S. A. Tretter, "Estimating the frequency of a noisy sinusoid by linear regression," *IEEE Trans. Inf. Theory*, vol. 31, no. 6, pp. 832–835, Nov. 1985.
[9] S. Kay, "A fast and accurate single frequency estimator," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 1987–1990, Dec. 1989.
[10] M. P. Fitz, "Further results in the fast estimation of a single frequency," *IEEE Trans. Commun.*, vol. 42, no. 234, pp. 862–864, Feb. 1994.
[11] D. Kim, M. Narasimha, and D. Cox, "An improved single frequency estimator," *IEEE Signal Process. Lett.*, vol. 3, no. 7, pp. 212–214, Jul. 1996.
[12] M. D. Macleod, "Fast nearly ML estimation of the parameters of real or complex single tones or resolved multiple tones," *IEEE Trans. Signal Process.*, vol. 46, no. 1, pp. 141–148, Jan. 1998.
[13] P. Händel, "Markov-based single-tone frequency estimation," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 45, no. 2, pp. 230–232, Feb. 1998.
[14] M. L. Fowler and J. A. Johnson, "Extending the threshold and frequency range for phase-based frequency estimation," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2857–2863, Oct. 1999.
[15] B. Völcker and P. Händel, "Frequency estimation from proper sets of correlations," *IEEE Trans. Signal Process.*, vol. 50, no. 4, pp. 791–802, Apr. 2002.
[16] T. Brown and M. M. Wang, "An iterative algorithm for single-frequency estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2671–2682, Nov. 2002.
[17] D. C. Rife and R. R. Boorstyn, "Single tone parameter estimation from discrete-time observations," *IEEE Trans. Inf. Theory*, vol. 20, no. 5, pp. 591–598, Sep. 1974.
[18] V. Clarkson, P. J. Kootsookos, and B. G. Quinn, "Analysis of the variance threshold of Kay's weighted linear frequency estimator," *IEEE Trans. Signal Process.*, vol. 42, no. 9, pp. 2370–2379, Sep. 1994.
[19] T. Soderstrom and P. Stoica, *System Identification*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
[20] G. C. Goodwin and R. L. Payne, *Dynamic System Identification: Experiment Design and Data Analysis*. New York, NY: Academic, 1977.
[21] L. C. Palmer, "Coarse frequency estimation using the discrete Fourier transform," *IEEE Trans. Inf. Theory*, vol. 20, no. 1, pp. 104–109, Jan. 1974.
[22] M. Marcus, "Basic theorems in matrix theory," in *Appl. Math. Series*: National Bureau of Standards, 1960.
[23] F. A. Graybill, *Matrices with Applications in Statistics*. Belmont, CA: Wadsworth Int. Group, 1983.
[24] Y. Hua, "The most efficient implementation of the IQML algorithm," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 2203–2204, Aug. 1994.
[25] D. M. Young, *Iterative Solution of Large Linear Systems*. New York: Academic, 1971.
[26] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: The Johns Hopkins Univ. Press, 1983.
[27] M. Aoki and P. C. Yue, "On *a priori* error estimates of some identification methods," *IEEE Trans. Autom. Control*, vol. 15, no. 5, pp. 541–548, Oct. 1970.
[28] R. Kumaresan, L. L. Scharf, and A. K. Shaw, "An algorithm for pole-zero modeling and spectral analysis," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 3, pp. 637–640, Jun. 1986.
[29] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 5, pp. 1081–1089, Oct. 1986.
[30] J. Li, P. Stoica, and Z.-S. Liu, "Comparative study of IQML and MODE direction-of-arrival estimators," *IEEE Trans. Signal Process.*, vol. 46, no. 1, pp. 149–160, Jan. 1998.
[31] J. H. McClellan and D. Lee, "Exact equivalence of the Steiglitz–McBride iteration and IQML," *IEEE Trans. Signal Process.*, vol. 39, no. 2, pp. 509–512, Feb. 1991.
[32] K. Steiglitz and L. E. McBride, "A technique for the identification of linear systems," *IEEE Trans. Autom. Control*, vol. 10, no. 4, pp. 461–464, Oct. 1965.
[33] S. M. Kay, "Accurate frequency estimation at low signal-to-noise ratio," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 3, pp. 540–547, Jun. 1984.
[34] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, 3rd ed. New York: Springer-Verlag, 2002.
[35] B. G. Quinn and J. M. Fernandes, "A fast efficient technique for the estimation of frequency," *Biometrika*, vol. 78, pp. 489–498, 1991.
[36] H. Lutkepohl, *Handbook of Matrices*. New York: Wiley, 1996.

H. C. So (S'90–M'90) was born in Hong Kong. He received the B.Eng. degree from the City University of Hong Kong, Hong Kong, and the Ph.D. degree from The Chinese University of Hong Kong, Hong Kong, both in electronic engineering, in 1990 and 1995, respectively.

From 1990 to 1991, he was an Electronic Engineer with the Research and Development Division of Everex Systems Engineering, Ltd., Hong Kong. From 1995 to 1996, he worked as a Postdoctoral Fellow at The Chinese University of Hong Kong. From 1996 to 1999, he was a Research Assistant Professor with the Department of Electronic Engineering, City University of Hong Kong. Currently, he is an Associate Professor in the Department of Electronic Engineering, City University of Hong Kong. His research interests include adaptive filter theory, detection and estimation, wavelet transform, and signal processing for communications and multimedia.

Frankie Kit Wing Chan received the B.Eng. degree in computer engineering and the M.Phil. degree both from the City University of Hong Kong, Hong Kong, in 2002 and 2005, respectively. Currently, he is working toward the Ph.D. degree at the Department of Electronic Engineering, City University of Hong Kong.

His research interests include statistical signal processing and their applications, with particular attention to frequency estimation and related mathematics.