Structured Total Least Squares Approach for Efficient Frequency Estimation

F. K. W. Chan, H. C. So, W. H. Lau and C. F. Chan

ckwf@hkexperts.com, hcso@ee.cityu.edu.hk, itwhlau@cityu.edu.hk, itcfchan@cityu.edu.hk
Department of Electronic Engineering
City University of Hong Kong
Tat Chee Avenue, Kowloon, Hong Kong

Index Terms : Frequency estimation, structured total least squares, Newton's method

Abstract

A new structured total least squares (STLS) based frequency estimation algorithm for real sinusoids corrupted by white noise is devised. Numerical results are included to contrast the estimator performance with an existing STLS frequency estimation method as well as the Cramér-Rao lower bound in different signal-to-noise ratio conditions.

1 Introduction

Frequency estimation of sinusoids in noise is a frequently addressed problem in the signal processing literature because it has a wide variety of applications such as source localization, speech and audio signal analysis, biomedical engineering, communications, power delivery as well as instrumentation and measurement [1]–[5].

Well known frequency estimation algorithms include maximum likelihood (ML) estimator [6], subspace-based techniques [7], Yule-Walker method [8] and linear prediction (LP) [9] approach. The ML method is statistically efficient in the sense that its estimation performance can attain the Cramér-Rao lower bound (CRLB) asymptotically under additive white Gaussian noise (AWGN). However, the huge computation requirement prohibits its use in some time-critical applications. Although the other mentioned methods are suboptimal, they are computationally efficient compared with the ML method. In frequency estimation using the LP method, an over-determined set of linear equations $\mathbf{X}\mathbf{a} \approx \mathbf{b}$, where the matrix $\mathbf{X}$ and vector $\mathbf{b}$ are formed by the observed data, while $\mathbf{a}$, which contains the frequency information, is the parameter vector to be determined, is set up and solved. In order to obtain the vector $\mathbf{a}$ and hence the frequency values, the ordinary least squares (LS) can be used but it will provide inconsistent frequency estimates because both $\mathbf{X}$ and $\mathbf{b}$ are contaminated with noise in practice. In view of this problem, total least squares (TLS) [10]–[11] has been proposed to solve the over-determined system. Although consistent frequency estimation can be achieved by TLS, its estimates are inefficient because the singular value decomposition involved in TLS will perturb $\mathbf{X}$ and $\mathbf{b}$ least such that the over-determined system is satisfied, regardless the special structure of $[\mathbf{b} \quad \mathbf{X}]$, which may be Hankel or Toeplitz.

The structured total least squares (STLS) approach [12]–[16], which exploits the special structure involved in the over-determined system can provide efficient parameter estimates. In fact, earlier work of STLS can be found in [12] and amendments have been made to improve the rate of convergence and reduce the computational complexity [13]–[15]. For example, Philippe *et al.* have suggested an iterative method [16], namely, STLS2, based on Lagrange-Newton method and provided a fast implementation of the algorithm. However, the involvement of the Lagrange multipliers will inevitably increase the computations of the algorithm and restrict its applications. In this paper, an iterative frequency estimation algorithm that is computationally more attractive will be developed based on the framework of STLS.

The rest of the paper is organized as follows. In Section 2, the problem of frequency estimation will be formulated. In Section 3, frequency estimation will be cast into the STLS framework and an algorithm will then be developed. The convergence of the devised STLS method is studied in Section 4. Numerical examples are presented in Section 5 to evaluate the performance of the proposed algorithm by comparing with the STLS2 method and CRLB. Finally, conclusions are drawn in Section 6.

2 Problem Formulation

Frequency estimation of real sinusoids is considered and readers can find it fairly straightforward to apply the algorithm to the complex counterpart. The signal model for real tone frequency estimation is

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

where

$$s_n = \sum_{m=1}^{M} \mathcal{A}_m \cos(\omega_m n + \phi_m) \tag{2}$$

The $\mathcal{A}_m > 0$, $\omega_m \in (0, \pi)$ and $\phi_m \in [0, 2\pi)$ are unknown constants representing the amplitude, frequency and phase of the m-th sinusoidal component, respectively, while q_n is a zero-mean AWGN. It is assumed that the number of sinusoids, M, is known *a priori*. From the N samples of $\{x_n\}$, we are interested to find ω_m, $m = 1, 2, \cdots, M$. Based on the LP property of s_n, we have

$$\begin{aligned} s_n &= \sum_{l=1}^{2M} a_l s_{n-l}, \qquad a_l = a_{2M-l}, a_{2M} = -1 \\ \Rightarrow s_n + s_{n-2M} &= \sum_{l=1}^{M-1} a_l \left(s_{n-l} + s_{n-2M+l}\right) + a_M s_{n-M} \end{aligned} \tag{3}$$

where the symmetric $\{a_l\}_{l=1}^{M}$ are called the LP coefficients. The frequencies $\{\omega_m\}$ can be calculated from [9]:

$$\sum_{l=1}^{2M} a_l \exp\left(-j\omega_m l\right) = 1 \tag{4}$$

Exploiting the symmetric property again, the order of (4) can be reduced from $2M$ to M via employing the Chebyshev polynomial of the first kind [17]:

$$\cos(n\omega) = T_n\left(\cos(\omega)\right) \tag{5}$$

where

$$T_n(x) = \frac{n}{2} \sum_{r=0}^{\lfloor n/2 \rfloor} \frac{(-1)^r}{n-r} C_r^{n-r} (2x)^{n-2r}$$

and $\lfloor \varrho \rfloor$ denotes rounding ϱ to the nearest integer towards minus infinity.

Based on (3), an over-determined set of linear equations can be formed using $\{x_n\}$:

$$\mathbf{X}\mathbf{a} \approx \mathbf{b} \tag{6}$$

where

$$\mathbf{a} = \begin{bmatrix} a_1 & a_2 & \cdots & a_M \end{bmatrix}^T$$

$$\mathbf{b} = \begin{bmatrix} x_N + x_{N-2M} & x_{N-1} + x_{N-2M-1} & \cdots & x_{2M+1} + x_1 \end{bmatrix}^T$$

$$\mathbf{X} = \begin{bmatrix} x_{N-1} + x_{N-2M+1} & x_{N-2} + x_{N-2M+2} & \cdots & x_{N-M+1} + x_{N-M-1} & x_{N-M} \\ x_{N-2} + x_{N-2M} & x_{N-3} + x_{N-2M+1} & \cdots & x_{N-M} + x_{N-M-2} & x_{N-M-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{2M} + x_2 & x_{2M-1} + x_3 & \cdots & x_{M+2} + x_M & x_{M+1} \end{bmatrix}$$

It can be seen that $[\mathbf{b} \quad \mathbf{X}]$ is, except the last column, a sum of a Hankel and Toeplitz matrix and solving (6) by LS or TLS cannot give optimal frequency estimates since the perturbation by LS and TLS cannot maintain the Hankel-Toeplitz structure of $[\mathbf{b} \quad \mathbf{X}]$. On the other hand, the STLS approach which takes the matrix structure into account during minimization can provide a more accurate parameter estimate.

3 Algorithm Development

The STLS algorithm tries to perform the optimization:

$$\begin{aligned} & \min \|\Delta \mathbf{x}\|_2^2 \\ \text{subject to} \qquad & (\mathbf{X}+\mathbf{\Delta X})\,\hat{\mathbf{a}} = \mathbf{b}+\mathbf{\Delta b} \end{aligned} \tag{7}$$

where $\|.\|_2$ stands for l_2 norm, $\hat{\mathbf{a}}$ denotes the estimated value of $\mathbf{a}$ and the vector $\mathbf{\Delta x}$ represents the perturbation of $\mathbf{x} = [x_N \quad \cdots \quad x_1]^T$, which contains all the elements of the matrix $[\mathbf{b} \quad \mathbf{X}]$ with no repetition. The matrix $[\mathbf{\Delta b} \quad \mathbf{\Delta X}]$ is formed by elements of $\mathbf{\Delta x}$ in such a way that $[\mathbf{b} \quad \mathbf{X}]$ is constructed from $\mathbf{x}$ and thus it has the same structure as $[\mathbf{b} \quad \mathbf{X}]$. In this section, an iterative algorithm will be developed to solve (6) using the framework of (7).

Let $\mathbf{r}(\mathbf{y}) = \mathbf{b}+\mathbf{\Delta b} - (\mathbf{X}+\mathbf{\Delta X})\,\hat{\mathbf{a}}$ where $\mathbf{y} = \left[\hat{\mathbf{a}}^T \quad \mathbf{\Delta x}^T\right]^T$ and consider the Taylor series expansion of $\mathbf{r}(\mathbf{y}_{n+1})$ around $\mathbf{P}\mathbf{y}_n$ where $\mathbf{y}_{n+1} = \left[\hat{\mathbf{a}}_{n+1}^T \quad \mathbf{\Delta x}_{n+1}^T\right]^T, \mathbf{y}_n = \left[\hat{\mathbf{a}}_n^T \quad \mathbf{\Delta x}_n^T\right]^T$ and $\mathbf{P} = \mathrm{diag}(\underbrace{1,\cdots,1}_{M},\underbrace{0\cdots,0}_{N})$, we have

$$\begin{aligned} \mathbf{0}_{(N-2M)\times 1} &= \mathbf{r}(\mathbf{y}_{n+1}) \approx \mathbf{r}(\mathbf{P}\mathbf{y}_n) + \mathbf{J}(\mathbf{P}\mathbf{y}_n)(\mathbf{y}_{n+1} - \mathbf{P}\mathbf{y}_n) \\ \Rightarrow \mathbf{y}_{n+1} &= \mathbf{P}\mathbf{y}_n - \mathbf{J}(\mathbf{P}\mathbf{y}_n)^{\dagger}\,\mathbf{r}(\mathbf{P}\mathbf{y}_n) \end{aligned} \tag{8}$$

where $\mathbf{0}_{i\times j}$ represents the $i \times j$ zero matrix, $(.)^{\dagger}$ denotes the pseudo-inverse and $\mathbf{J}$ stands for the Jacobian matrix:

$$\mathbf{J}(\mathbf{y}) = \left[\ -(\mathbf{X}+\mathbf{\Delta X}) \quad \hat{\mathbf{A}}\ \right]$$

with $\hat{\mathbf{A}} \in \mathbb{R}^{(N-2M)\times N}$ of the form:

$$\hat{\mathbf{A}} = \mathrm{Toeplitz}\left(\left[1 \quad \mathbf{0}_{1\times(N-2M-1)}\right]^T, \left[1 \quad -\hat{\mathbf{a}}^T \quad -\hat{a}_{M-1} \quad \cdots \quad -\hat{a}_1 \quad 1 \quad \mathbf{0}_{1\times(N-2M)}\right]\right)$$

where Toeplitz$(\mathbf{u},\mathbf{v}^T)$ is the Toeplitz matrix with first column $\mathbf{u}$ and first row $\mathbf{v}^T$. Based on (8), the recursive algorithm is summarized as:

(i) Initialize $\hat{\mathbf{a}}$ with any consistent estimator and set $\mathbf{\Delta x}$ to $\mathbf{0}_{N\times 1}$.

(ii) While *stopping criterion* is not satisfied

$\mathbf{y}_{n+1} = \mathbf{P}\mathbf{y}_n - \mathbf{J}(\mathbf{P}\mathbf{y}_n)^{\dagger}\,\mathbf{r}(\mathbf{P}\mathbf{y}_n)$

end

The *stopping criterion* above can be the maximum number of allowable iterations or the difference of norm between successive $\hat{\mathbf{a}}$ is less than a small positive number. In each iteration, an under-determined system in Step (ii) is solved and thus the min-norm solution corresponding to $\mathbf{\Delta x} = \mathbf{0}_{N\times 1}$ is obtained, which means that the calculated $\hat{\mathbf{a}}$ is constrained by the minimum norm criterion of (7) and hence the constrained optimization is performed. It should be noted that the fast implementation technique [16] cannot be used in our situation since the matrix in [16] is either Hankel or Toeplitz but the matrix $\mathbf{X}$ in (6) is neither one of them. Furthermore, the STLS2 method needs to solve a kernel problem corresponding to Step (ii) of the proposed algorithm, which is the main computation of both algorithms in each iteration. However, the matrix $\mathbf{J}$ in (8) is a $(N-2M)\times(M+N)$ submatrix of the corresponding $(2N-M)\times(2N-M)$ matrix of the STLS2 method. Considering the computational complexity of solving a set of linear equations $\mathbf{\Xi}\boldsymbol{\beta} \approx \boldsymbol{\gamma}$ where $\mathbf{\Xi} \in \mathbb{R}^{m\times n}$, $\boldsymbol{\beta} \in \mathbb{R}^{n\times 1}$ and $\boldsymbol{\gamma} \in \mathbb{R}^{m\times 1}$ with $m \leq n$, by QR factorization [18], the whole LS process requires $2nm^2$ FLOPS for the QR factorization of $\mathbf{\Xi}^T = \mathbf{QR}$, m^2 FLOPS for the backward substitution of $\mathbf{R}^T\boldsymbol{\alpha} = \boldsymbol{\gamma}$ to solve $\boldsymbol{\alpha}$ and $2mn$ FLOPS for vector construction of $\boldsymbol{\gamma} = \mathbf{Q}\boldsymbol{\alpha}$. Therefore, the total numbers of FLOPS of the proposed and STLS2 algorithms are $8M^3-6MN-6MN^2+3N^2+2N^3$ and $3M^2-2M^3-12MN+12M^2N+12N^2-24MN^2+16N^3$, respectively. At each iteration, it is clear that the proposed scheme is more computationally attractive than the STLS2 algorithm, although both methods have a complexity of $\mathcal{O}(N^3)$.

4 Convergence Analysis

Regarding the local convergence, we cannot follow conventional STLS algorithms such as [14]–[16] to produce the proof for the proposed method which deals with an under-determined system of equations. Nevertheless, the local convergence of this kind of numerical algorithms which apply Newton's method to solve under-determined systems has been proved in [19]–[20]. We base on [19]–[20] to analyze the convergence of the proposed method as follows. If the initial value of $\mathbf{\Delta x}$ is set to $\mathbf{0}$ and only $\mathbf{a}$ of $\mathbf{y}$ is considered, the iteration of (8) is the same as

$$\mathbf{y}_{n+1} = \mathbf{y}_n - \mathbf{PJ}(\mathbf{y}_n)^{\dagger}\,\mathbf{r}(\mathbf{y}_n) \tag{9}$$

except that in the final iteration, (8) will provide a non-zero value of $\mathbf{\Delta x}$. Before proceeding to the convergence proof of the proposed algorithm, some useful lemmas will be given. We first define the following quantities:

$$\begin{aligned}
\gamma(\mathbf{y}) &= \sup_{n>1}\left\|\mathbf{r}^{(1)}(\mathbf{y})^{\dagger}\,\frac{\mathbf{r}^{(n)}(\mathbf{y})}{n!}\right\|^{\frac{1}{n-1}} \\
\beta(\mathbf{y}) &= \left\|\mathbf{r}^{(1)}(\mathbf{y})^{\dagger}\,\mathbf{r}(\mathbf{y})\right\| \\
\alpha(\mathbf{y}) &= \beta\gamma \\
\psi(u) &= 2u^2-4u+1
\end{aligned}$$

Lemma 1

Let $\mathbf{A}, \mathbf{B} \in \mathcal{R}^{m\times n}$ *where* $m \leq n$. *If* $\left\|\mathbf{B}^{\dagger}\left(\mathbf{B}-\mathbf{A}\right)\right\| \leq \lambda < 1$, *then* $\left\|\mathbf{A}^{\dagger}\mathbf{B}\right\| < \dfrac{1}{1-\lambda}$

Proof

$$\begin{aligned}\left\|\mathbf{A}^{\dagger}\mathbf{B}\right\| &= \left\|\mathbf{A}^{\dagger}\mathbf{A}\left(\mathbf{I}-\mathbf{B}^{\dagger}\left(\mathbf{B}-\mathbf{A}\right)\right)^{-1}\right\| \\ &\leq \left\|\mathbf{A}^{\dagger}\mathbf{A}\right\|\left\|\left(\mathbf{I}-\mathbf{B}^{\dagger}\left(\mathbf{B}-\mathbf{A}\right)\right)^{-1}\right\| \\ &< \frac{1}{1-\lambda} \quad \blacksquare\end{aligned}$$

Lemma 2 *If* $\alpha\left(\mathbf{y}_{n-1}\right) < 1-\frac{\sqrt{2}}{2}$, *then*

a) $\mathbf{r}^{(1)}\left(\mathbf{y}_{n-1}\right)^{\dagger}$ *exists*

b) $\left\|\mathbf{r}^{(1)}\left(\mathbf{y}_{n}\right)^{\dagger}\mathbf{r}^{(1)}\left(\mathbf{y}_{n-1}\right)\right\| \leq \frac{\left(1-\alpha\left(\mathbf{y}_{n-1}\right)\right)^2}{\psi\left(\alpha\left(\mathbf{y}_{n-1}\right)\right)}$

c) $\gamma\left(\mathbf{y}_{n}\right) \leq \frac{\gamma\left(\mathbf{y}_{n-1}\right)}{\left(1-\alpha\left(\mathbf{y}_{n-1}\right)\right)\psi\left(\alpha\left(\mathbf{y}_{n-1}\right)\right)}$

Proof

$$\begin{aligned}\mathbf{r}^{(1)}\left(\mathbf{y}_{n}\right) &= \mathbf{r}^{(1)}\left(\mathbf{y}_{n-1}\right)+\sum_{k=2}^{\infty}k\frac{\mathbf{r}^{(k)}\left(\mathbf{y}_{n-1}\right)}{k!}\left(\mathbf{y}_{n}-\mathbf{y}_{n-1}\right)^{k-1} \\ \mathbf{r}^{(1)}\left(\mathbf{y}_{n-1}\right)^{\dagger}\left(\mathbf{r}^{(1)}\left(\mathbf{y}_{n}\right)-\mathbf{r}^{(1)}\left(\mathbf{y}_{n-1}\right)\right) &= \sum_{k=2}^{\infty}k\mathbf{r}^{(1)}\left(\mathbf{y}_{n-1}\right)^{\dagger}\frac{\mathbf{r}^{(k)}\left(\mathbf{y}_{n-1}\right)}{k!}\left(\mathbf{y}_{n}-\mathbf{y}_{n-1}\right)^{k-1}\end{aligned}$$

Taking norm of both sides, we get

$$\begin{aligned}\left\|\mathbf{r}^{(1)}\left(\mathbf{y}_{n-1}\right)^{\dagger}\left(\mathbf{r}^{(1)}\left(\mathbf{y}_{n}\right)-\mathbf{r}^{(1)}\left(\mathbf{y}_{n-1}\right)\right)\right\| &\leq \sum_{k=2}^{\infty}k\gamma\left(\mathbf{y}_{n-1}\right)^{k-1}\left\|\left(\mathbf{y}_{n}-\mathbf{y}_{n-1}\right)\right\|^{k-1} \\ &= \sum_{k=2}^{\infty}k\alpha\left(\mathbf{y}_{n-1}\right)^{k-1} \\ &= \frac{1}{\left(1-\alpha\left(\mathbf{y}_{n-1}\right)\right)^2}-1\end{aligned}$$

which is less than 1 since $\alpha\left(\mathbf{y}_{n-1}\right) < 1-\frac{\sqrt{2}}{2}$. By using Lemma 1, we can prove Lemmas 2a and 2b. To prove Lemma 2c, we consider:

$$\begin{aligned}\left\|\mathbf{r}^{(1)}\left(\mathbf{y}_{n}\right)^{\dagger}\frac{\mathbf{r}^{(k)}\left(\mathbf{y}_{n}\right)}{k!}\right\| &\leq \left\|\mathbf{r}^{(1)}\left(\mathbf{y}_{n}\right)^{\dagger}\mathbf{r}^{(1)}\left(\mathbf{y}_{n-1}\right)\right\|\left\|\sum_{l=0}^{\infty}\mathbf{r}^{(1)}\left(\mathbf{y}_{n-1}\right)^{\dagger}\frac{\mathbf{r}^{(k+l)}\left(\mathbf{y}_{n-1}\right)}{k!l!}\right\|\left\|\mathbf{y}_{n}-\mathbf{y}_{n-1}\right\|^{l} \\ &\leq \frac{\left(1-\alpha\left(\mathbf{y}_{n-1}\right)\right)^2}{\psi\left(\alpha\left(\mathbf{y}_{n-1}\right)\right)}\sum_{l=0}^{\infty}\left\|\frac{(k+l)!}{k!l!}\gamma\left(\mathbf{y}_{n-1}\right)^{k+l-1}\right\|\left\|\mathbf{y}_{n}-\mathbf{y}_{n-1}\right\|^{l} \\ &= \frac{\left(1-\alpha\left(\mathbf{y}_{n-1}\right)\right)^2}{\psi\left(\alpha\left(\mathbf{y}_{n-1}\right)\right)}\times\frac{\gamma\left(\mathbf{y}_{n-1}\right)^{k-1}}{\left(1-\alpha\left(\mathbf{y}_{n-1}\right)\right)^{k+1}}\end{aligned}$$

The last equality is obtained by the fact of $(1-x)^{-k} = \sum_{l=0}^{\infty} \frac{(k+l-1)x^l}{(k-1)!l!}$. Hence, the result follows by noting that $0 < \psi(u) < 1$ for $0 < u < 1 - \frac{\sqrt{2}}{2}$:

$$\begin{aligned}
\left\| \mathbf{r}^{(1)}(\mathbf{y}_n)^{\dagger} \frac{\mathbf{r}^{(k)}(\mathbf{y}_n)}{k!} \right\|^{\frac{1}{k-1}} &\leq \frac{\gamma(\mathbf{y}_{n-1})}{(1-\alpha(\mathbf{y}_{n-1}))\,\psi(\alpha(\mathbf{y}_{n-1}))^{k-1}} \\
&\leq \frac{\gamma(\mathbf{y}_{n-1})}{(1-\alpha(\mathbf{y}_{n-1}))\,\psi(\alpha(\mathbf{y}_{n-1}))} \\
\Rightarrow \gamma(\mathbf{y}_n) &\leq \frac{\gamma(\mathbf{y}_{n-1})}{(1-\alpha(\mathbf{y}_{n-1}))\,\psi(\alpha(\mathbf{y}_{n-1}))} \quad \blacksquare
\end{aligned}$$

Lemma 3 *Let* $\alpha(\mathbf{y}_{n-1}) < 1$*, then*

$$\left\| \mathbf{r}^{(1)}(\mathbf{y}_{n-1})^{\dagger} \mathbf{r}(\mathbf{y}_n) \right\| \leq \frac{\alpha(\mathbf{y}_{n-1})\,\beta(\mathbf{y}_{n-1})}{1-\alpha(\mathbf{y}_{n-1})}$$

Proof

$$\begin{aligned}
\left\| \mathbf{r}^{(1)}(\mathbf{y}_{n-1})^{\dagger} \mathbf{r}(\mathbf{y}_n) \right\| &\leq \sum_{k=2}^{\infty} \left(\left\| \mathbf{r}^{(1)}(\mathbf{y}_{n-1})^{\dagger} \frac{\mathbf{r}^{(k)}(\mathbf{y}_{n-1})}{k!} \right\| \|\mathbf{y}_n - \mathbf{y}_{n-1}\|^k \right) \\
&\leq \beta(\mathbf{y}_{n-1}) \sum_{k=2}^{\infty} (\gamma(\mathbf{y}_{n-1})\,\beta(\mathbf{y}_{n-1}))^{k-1} \\
&= \frac{\alpha(\mathbf{y}_{n-1})\,\beta(\mathbf{y}_{n-1})}{1-\alpha(\mathbf{y}_{n-1})} \quad \blacksquare
\end{aligned}$$

Lemma 4 *If* $\alpha(\mathbf{y}_{n-1}) < 1 - \frac{\sqrt{2}}{2}$*, then*

$$\beta(\mathbf{y}_n) \leq \beta(\mathbf{y}_{n-1}) \left(\frac{\alpha(\mathbf{y}_{n-1})\,(1-\alpha(\mathbf{y}_{n-1}))}{\psi(\alpha(\mathbf{y}_{n-1}))} \right)$$

Proof

$$\begin{aligned}
\beta(\mathbf{y}_n) &= \left\| \mathbf{r}^{(1)}(\mathbf{y}_n)^{\dagger} \mathbf{r}(\mathbf{y}_n) \right\| \\
&\leq \left\| \mathbf{r}^{(1)}(\mathbf{y}_n)^{\dagger} \mathbf{r}^{(1)}(\mathbf{y}_{n-1}) \right\| \left\| \mathbf{r}^{(1)}(\mathbf{y}_{n-1})^{\dagger} \mathbf{r}(\mathbf{y}_n) \right\| \\
&\leq \beta(\mathbf{y}_{n-1}) \left(\frac{\alpha(\mathbf{y}_{n-1})\,(1-\alpha(\mathbf{y}_{n-1}))}{\psi(\alpha(\mathbf{y}_{n-1}))} \right) \quad \blacksquare
\end{aligned}$$

Lemma 5 *If* $\alpha(\mathbf{y}_{n-1}) < 1 - \frac{\sqrt{2}}{2}$*, then*

$$\alpha(\mathbf{y}_n) \leq \left(\frac{\alpha(\mathbf{y}_{n-1})}{\psi(\alpha(\mathbf{y}_{n-1}))} \right)^2$$

Proof

$$\begin{aligned}
\alpha(\mathbf{y}_n) &= \beta(\mathbf{y}_n)\,\gamma(\mathbf{y}_n) \\
&\leq \beta(\mathbf{y}_{n-1}) \left(\frac{\alpha(\mathbf{y}_{n-1})\,(1-\alpha(\mathbf{y}_{n-1}))}{\psi(\alpha(\mathbf{y}_{n-1}))} \right) \left(\frac{\gamma(\mathbf{y}_{n-1})}{(1-\alpha(\mathbf{y}_{n-1}))\,\psi(\alpha(\mathbf{y}_{n-1}))} \right) \\
&= \left(\frac{\alpha(\mathbf{y}_{n-1})}{\psi(\alpha(\mathbf{y}_{n-1}))} \right)^2 \quad \blacksquare
\end{aligned}$$

Theorem 1

$$\|\mathbf{y}_n - \mathbf{y}_{n-1}\| \leq \mu^{2^{n-1}-1} \|\mathbf{y}_1 - \mathbf{y}_0\| \; \forall n \textit{ where } \mu = \frac{\alpha(\mathbf{y}_0)}{\psi(\alpha(\mathbf{y}_0))^2} < 1$$

Proof The case of $n = 1$ is trivial. For $n > 1$, we have:

$$\begin{aligned}
\|\mathbf{y}_n - \mathbf{y}_{n-1}\| &\leq \|\mathbf{y}_{n-1} - \mathbf{y}_{n-2}\| \frac{\alpha(\mathbf{y}_{n-2})(1-\alpha(\mathbf{y}_{n-2}))}{\psi(\alpha(\mathbf{y}_{n-2}))} \\
&\leq \|\mathbf{y}_1 - \mathbf{y}_0\| \prod_{i=0}^{n-2} \frac{\alpha(\mathbf{y}_i)(1-\alpha(\mathbf{y}_i))}{\psi(\alpha(\mathbf{y}_i))} \\
&\leq \|\mathbf{y}_1 - \mathbf{y}_0\| \prod_{i=0}^{n-2} \frac{\alpha(\mathbf{y}_i)}{\psi(\alpha(\mathbf{y}_0))} \\
&\leq \|\mathbf{y}_1 - \mathbf{y}_0\| \prod_{i=0}^{n-2} \left(\frac{\alpha(\mathbf{y}_0)}{\psi(\alpha(\mathbf{y}_0))^2}\right)^{2^n-1} \frac{\alpha(\mathbf{y}_0)}{\psi(\alpha(\mathbf{y}_0))} \\
&\leq \|\mathbf{y}_1 - \mathbf{y}_0\| \prod_{i=0}^{n-2} \left(\frac{\alpha(\mathbf{y}_0)}{\psi(\alpha(\mathbf{y}_0))^2}\right)^{2^n} \psi(\alpha(\mathbf{y}_0)) \\
&\leq \|\mathbf{y}_1 - \mathbf{y}_0\| \left(\frac{\alpha(\mathbf{y}_0)}{\psi(\alpha(\mathbf{y}_0))^2}\right)^{2^{n-1}-1} \\
&= \mu^{2^{n-1}-1} \|\mathbf{y}_1 - \mathbf{y}_0\| \quad \blacksquare
\end{aligned}$$

5 Numerical Examples

Simulation tests have been conducted to evaluate the performance of the proposed algorithm by comparing with the STLS2 method of [16] as well as CRLB in a closely-spaced sinusoids scenario. The signal s_n is composed of two sinusoids with amplitudes $\alpha_1 = \alpha_2 = \sqrt{2}$, frequencies $\omega_1 = 0.3\pi$ and $\omega_2 = 0.38\pi$ and phases $\phi_1 = 1$ and $\phi_2 = 2$ at $N = 20$. All results provided are averages of 1000 independent runs using a computer with Pentium Dual Core 2 GHz processors and 1GB RAM. Both the STLS2 and proposed algorithms terminate when the difference of norm of successive $\hat{\mathbf{a}}$ is less than 10^{-6} or the number of iterations has reached the maximum allowable value, which is set to 10. The initial value of $\hat{\mathbf{x}}$ is obtained by a simple consistent method [22] and that of $\boldsymbol{\Delta}\hat{\mathbf{x}}$ is set to $\mathbf{0}_{N\times 1}$ for both algorithms.

In Figure 1, the mean square frequency errors of $\hat{\omega}_1$ are plotted against the signal-to-noise ratio (SNR). It is observed that the performance of the proposed method attains CRLB when SNR ≥ 12 dB while the STLS2 algorithm is optimum only when SNR ≥ 18 dB, which indicates the former has a better threshold performance. Figure 2, which corresponds to ω_2, shows that the threshold SNR of both STLS2 and proposed methods is 16 dB. Figure 3 shows the average iteration numbers of both algorithms for parameter convergence. It is observed that the average number of the proposed method is less than that of the STLS2 method when SNR ≤ 22 dB, which demonstrates the superiority of the proposed method over the STLS2 in terms of rate of convergence. Furthermore, the computational times of the proposed and STLS2 algorithms for one iteration are measured as 1.56×10^{-4} s and 4.25×10^{-4} s, respectively, which agree with the complexity analysis in Section 3.

6 Conclusion

An iterative algorithm has been developed for frequency estimation based on the framework of STLS. Based on computer simulations, the efficient statistical performance of the proposed method is demonstrated by comparing with the STLS2 algorithm and CRLB. Furthermore, the proposed method is more computationally attractive than the STLS2 scheme.

References

[1] M. G. Christensen, P. Stoica, A. Jakobsson and S. H. Jensen, "Multi-pitch estimation," *Signal Processing*, vol.88, no.4, pp.972-983, Apr. 2008.

[2] J. He and Z. Liu, "Two-dimensional direction finding of acoustic sources by a vector sensor array using the propagator method," *Signal Processing*, vol.88, no.10, pp.2492-2499, Oct. 2008.

[3] S.M. Sameer and R. V. Raja Kumar, "An efficient technique for the integer frequency offset estimation in OFDM systems," *Signal Processing*, vol.89, no.2, pp.252-256, Feb. 2009.

[4] R. Punchalard, J. Koseeyaporn and P. Wardkein, "Indirect frequency estimation based on second-order adaptive FIR notch filter," *Signal Processing*, vol.89, no.7, pp.1428-1435, Jul. 2009.

[5] R. Chudamani, K. Vasudevan and C. S. Ramalingam, "Real-time estimation of power system frequency using nonlinear least squares," *IEEE Transactions on Power Delivery*, vol.24, no.3, pp.1021-1028, Jul. 2009.

[6] P. Stoica, R. L. Moses, B. Friedlander and T. Soderstrom, "Maximum likelihood estimation of the parameters of multiple sinusoids from noisy measurements, " *IEEE Trans. Acoust., Speech, Signal Processing*, vol.37, no.3, pp.378-392, Mar. 1989.

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

[8] P. Stoica, R. L. Moses, T. Soderstrom and J. Li, "Optimal high-order Yule-Walker estimation of sinusoidal frequencies," *IEEE Trans. Acoust. Speech Signal Process.,* vol. 39, no. 6, pp. 1360-1368, Jun. 1991.

[9] Y. T. Chan, J. M. M. Lavoie and J. B. Plant, "A parameter estimation approach to estimation of frquencies of sinusoids," *IEEE Trans. Acoust. Speech, Signal Processing*, vol.29, pp.214-219, 1981.

[10] S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*," SIAM Philadelphia, 1991.

[11] I. Markovsky and S. Van Huffel, "Overview of total least-squares methods," *Signal Processing*, vol.87, no.10, pp.2283-2302, Oct. 2007.

[12] T. J. Abatzoglou, J. M. Mendel and G. A. Harada, "The constrained total least squares technique and its applications to harmonic superresolution," *IEEE Trans. Signal Processing*, vol.39, pp.1070-1087, 1991.

[13] B. De Moor, "Total least squares for affinely structured matrices and the noisy realization problem," *IEEE Trans. Signal Processing*, vol.42, pp.3104-3113, 1994.

[14] J. B. Rosen, H. Park and J. Glick, "Total least norm formulation and solution for structured problems," *SIAM J. Matrix Anal. Appl.*, vol.17, pp.110-126, 1996.

[15] S. Van Huffel, H. Park and J. B. Rosen, "Formulation and solution of structured total least norm problems for parameter estimation," *IEEE Trans. Signal Processing*, vol.44, pp.2464-2474, 1996.

[16] P. Lemmerling, N. Mastronardi and S. Van Huffel, "Fast algorithm for solving the Hankel/Toeplitz structured total least squares problem," *Numerical Algorithms*, vol.23, pp.371-392, 2000.

[17] T. J. Rivlin, *Chebyshev Polynomials from Approximation Theory to Algebra and Number Theory*, Wiley-Interscience, 1980.

[18] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The John Hopkins University Press, 1996.

[19] S. Smale, "Newton's method estimates from data at one point," *The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics*, Springer-Verlag, pp.185-196, 1986.

[20] M. Shub and S. Smale, "Complexity of Bézout's theorem IV: probability of success, extensions," *SIAM J. Numer. Anal.*, vol.33, pp.128-148, 1996.

[21] J. P. Dedieu and M. Shub, "Multihomogeneous Newton methods," *Mathematics of Computations*, vol.69, pp.1071-1098, 1999.

[22] H. C. So, K. W. Chan, Y. T. Chan and K. C. Ho, "Linear prediction approach for efficient freqeuncy estimation of multiple real sinusoids: algorithms and analyses," *IEEE Trans. Signal Processing*, vol.53, pp.2290-2305, 2005.

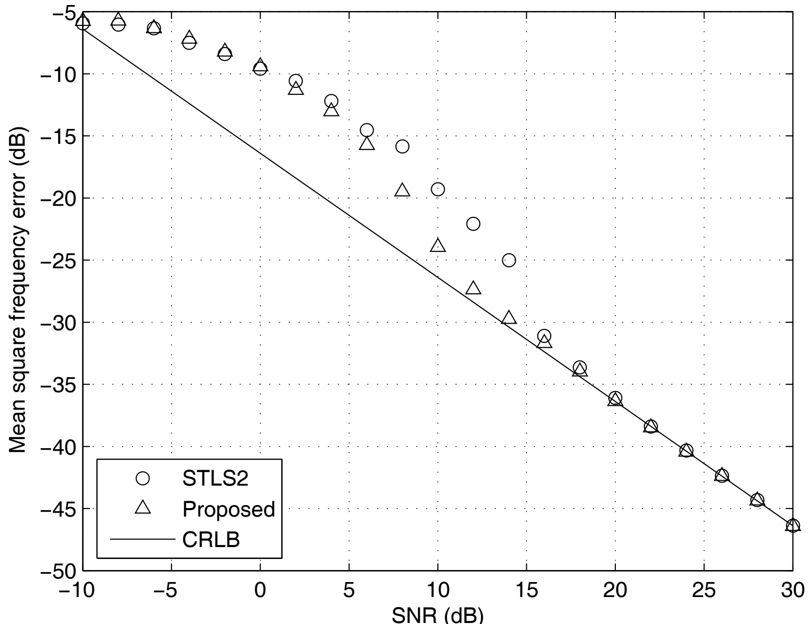

Figure 1: Mean square frequency error for ω_1 versus SNR

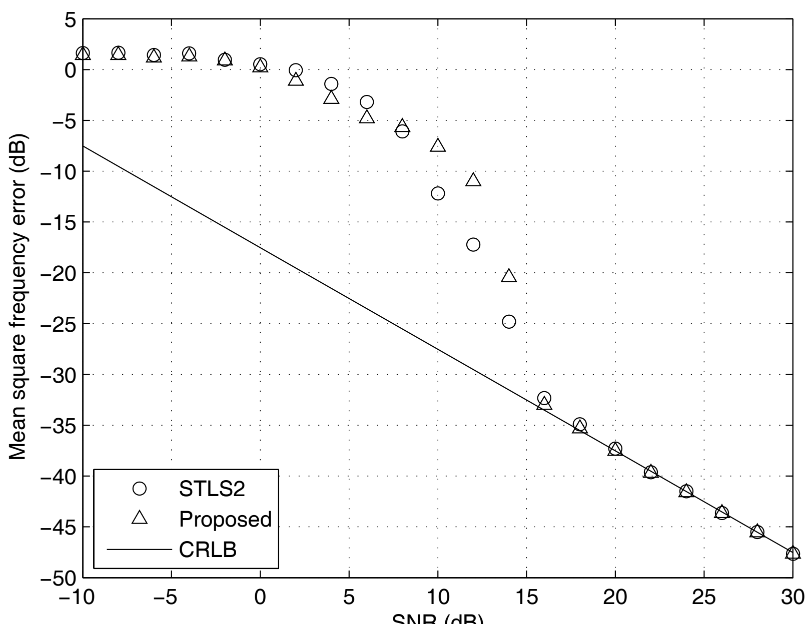

Figure 2: Mean square frequency error for ω_2 versus SNR

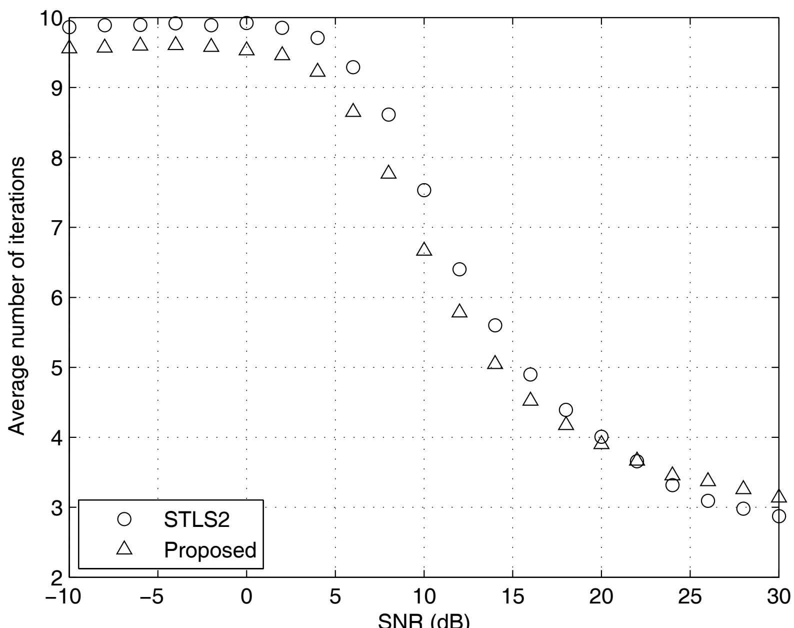

Figure 3: Average number of iterations versus SNR