ANALYSIS AND REDUCTION OF ESTIMATION BIAS FOR AN ITERATIVE FREQUENCY ESTIMATOR OF COMPLEX SINUSOID

Jan-Ray Liao[⋆†] *and Chun-Ming Chen*[†]

[⋆]Institute of Communication Engineering and [†]Department of Electrical Engineering
National Chung Hsing University, Taichung, Taiwan

ABSTRACT

Frequency estimation of a complex sinusoidal signal is a fundamental problem in signal processing. In this regard, Aboutanios and Mulgrew (A&M) proposed an iterative frequency estimator which can approach the theoretical bound in two iterations, thus, made it one of the best iterative estimators. In this paper, we theoretically analyze the two versions of the A&M estimator and show that the estimation biases of the two versions are not equivalent. The results of the theoretical analysis indicate that the bias of the first iteration can be accurately predicted by a polynomial equation. We then propose to use the roots of the polynomial equation to improve the estimation and reduce the bias. Experiments show that the proposed new estimator can significantly reduce the bias.

Index Terms— Frequency estimation, complex sinusoid, estimation bias, bias reduction.

1. INTRODUCTION

Frequency estimation of a single-tone complex sinusoidal signal is a fundamental problem in digital signal processing and has applications in areas such as radar signal processing and communication. The signal can be described as:

$$s[n] = A_0 e^{j(2\pi \frac{f_0}{f_s} n + \theta_0)} + w[n], n = 0, 1, \ldots, N-1 \quad (1)$$

where A_0, f_0, and θ_0 are the amplitude, frequency, and phase of the signal, respectively, the term $w[n]$ is an additive noise, f_s is the sampling frequency, and N is the number of samples.

When the noise $w[n]$ is a zero-mean white Gaussian noise, Rife et al. [1] showed that the maximum likelihood (ML) frequency estimation is the frequency that maximizes the magnitude of the periodogram of $s[n]$:

$$\hat{f}_{0,ML} = \arg\max_f \left\{ \left| \sum_{n=0}^{N-1} s[n] e^{-j2\pi n f} \right| \right\} \quad (2)$$

They also showed that the Cramér Rao lower bound (CRLB) of the mean squared error for the frequency estimation is $6f_s^2/(4\pi^2 N(N^2-1)\rho)$ with ρ being the SNR of $s[n]$.

This work was supported by National Science Council, Taiwan, Republic of China, under contract NSC 100-2221-E-005-001

The search for the frequency is usually divided into two steps: a coarse search and a fine search. The coarse search is to find an initial estimation from the peak magnitude of the discrete Fourier transform (DFT) of $s[n]$:

$$S[k] = \sum_{n=0}^{N-1} s[n] e^{-j\frac{2\pi}{N}nk} \quad (3)$$

Let k_p be the index of the peak magnitude in (3). The fine search is to find a frequency offset $\hat{\delta}$ around the initial estimation with the constraint $|\hat{\delta}| < 1/2$. The final estimated frequency is then:

$$\hat{f}_0 = (k_p + \hat{\delta}) \frac{f_s}{N} \quad (4)$$

The methods for fine search can be divided into two categories: direct approaches and iterative approaches. For the direct approaches, $\hat{\delta}$ is directly calculated from three or more spectrum lines [2–5]. On the contrary, iterative approaches repeatedly refine the solution either through dichotomous search [4, 6–8] or numerical method [1, 9–15].

Among the iterative estimators, the one proposed by Aboutanios and Mulgrew [11] can approach CRLB asymptotically in two iterations. This feature makes it one of the best iterative estimators. Since this estimator will be the focus of this paper, we will call it "A&M" estimator from now on. Based on the A&M estimator, several improvements have been proposed. Djurovic et. al. [12, 13] applied marginal median DFT and L-DFT for non-Gaussian noise. Minhas et. al. [14] replaced the original estimation equation to reduce computation cost. Liu et. al. [15] shifted the starting point in the first iteration to non-integer location to further improve its performance.

There are two versions of the A&M estimator and their performance under the influence of noise are proved to be equivalent. However, a very important aspect of the estimator has not yet been studied, i.e., its estimation bias. The bias is an important limiting factor for the performance of an estimator because it becomes the major source of error as the level of noise becomes small. In this paper, we theoretically analyze the estimation biases of the two versions of the A&M estimator and show that the biases of the two versions are

Table 1. Aboutanios and Mulgrew Algorithm

$S = FFT(s)$
$k_p = \arg\max_k\{|S[k]|\}$
$\hat{\delta}_0 = 0$
for $i = 1$ to Q **do**
$X_p = \sum_{n=0}^{N-1} s[n]\exp(-j2\pi n(k_p+\hat{\delta}_{i-1}+p)/N), p = \pm 0.5$
$h(\hat{\delta}_{i-1}) = \begin{cases} \frac{1}{2}\text{Re}\left\{\frac{X_{0.5}+X_{-0.5}}{X_{0.5}-X_{-0.5}}\right\}, & \text{for version 1} \\ \frac{1}{2}\frac{|X_{0.5}|-|X_{-0.5}|}{|X_{0.5}|+|X_{-0.5}|}, & \text{for version 2} \end{cases}$
$\hat{\delta}_i = \hat{\delta}_{i-1} + h(\hat{\delta}_{i-1})$
end for
$\hat{f}_0 = (k_p + \hat{\delta}_Q)\frac{f_s}{N}$

not equivalent. Using the results of the theoretical analysis, we then propose a new estimator to significantly reduce the bias, thereby, extending the SNR range that the estimator approaches CRLB.

2. THEORETICAL ANALYSIS OF THE ESTIMATION BIAS

The original A&M algorithm is listed in Table 1. For the first iteration, the estimated frequency offsets of AM estimator are as follows:

$$\hat{\delta}_1^a = \frac{1}{2}\text{Re}\left\{\frac{S[k_p+\frac{1}{2}]+S[k_p-\frac{1}{2}]}{S[k_p+\frac{1}{2}]-S[k_p-\frac{1}{2}]}\right\} \tag{5}$$

$$\hat{\delta}_1^b = \frac{1}{2}\frac{\left|S[k_p+\frac{1}{2}]\right|+\left|S[k_p-\frac{1}{2}]\right|}{\left|S[k_p+\frac{1}{2}]\right|-\left|S[k_p-\frac{1}{2}]\right|} \tag{6}$$

where $\hat{\delta}_1^a$ and $\hat{\delta}_1^b$ are for the first and second version of the estimator respectively.

What we want to do is to theoretically analyze the above two equations. The first step is to derive the analytical expression for the two spectrum lines: $S[k_p-1/2]$ and $S[k_p+1/2]$. Since we will focus our discussion on the bias behavior, the noise $w[n]$ is set to zero. Plugging (1) and (4) into (3), we get:

$$S[k_p+k_d] = A_0 e^{j\theta_0}\sum_{n=0}^{N-1} e^{-j2\pi(\frac{k_d-\delta}{N})n} \tag{7}$$

This equation is the sum of a geometric series and the two spectrum lines can be expressed as:

$$S[k_p-\frac{1}{2}] = A_0 e^{j\theta_0} e^{j\frac{(N-1)\pi(\delta+\frac{1}{2})}{N}}\frac{\sin(\pi(\delta+\frac{1}{2}))}{\sin(\frac{\pi(\delta+\frac{1}{2})}{N})} \tag{8}$$

$$S[k_p+\frac{1}{2}] = A_0 e^{j\theta_0} e^{j\frac{(N-1)\pi(\delta-\frac{1}{2})}{N}}\frac{\sin(\pi(\delta-\frac{1}{2}))}{\sin(\frac{\pi(\delta-\frac{1}{2})}{N})} \tag{9}$$

Plugging (8) and (9) into the equations (5) and (6), the analytical expressions for the first estimator can be expressed as follows:

$$\hat{\delta}_1^a = \frac{1}{2}\text{Re}\left\{\frac{e^{\frac{j\pi}{2N}}\sin\left(\frac{\pi}{N}(\delta+\frac{1}{2})\right)+e^{\frac{-j\pi}{2N}}\sin\left(\frac{\pi}{N}(\delta-\frac{1}{2})\right)}{e^{\frac{j\pi}{2N}}\sin\left(\frac{\pi}{N}(\delta+\frac{1}{2})\right)-e^{\frac{-j\pi}{2N}}\sin\left(\frac{\pi}{N}(\delta-\frac{1}{2})\right)}\right\} \tag{10}$$

Let $s_1 = \sin(\pi(\delta+1/2)/N)$ and $s_2 = \sin(\pi(\delta-1/2)/N)$.

$$\begin{aligned}\hat{\delta}_1^a &= \frac{1}{2}\text{Re}\left\{\frac{\cos(\frac{\pi}{2N})(s_1+s_2)+j\sin(\frac{\pi}{2N})(s_1-s_2)}{\cos(\frac{\pi}{2N})(s_1-s_2)+j\sin(\frac{\pi}{2N})(s_1+s_2)}\right\} \\ &= \frac{1}{2}\frac{s_1^2-s_2^2}{s_1^2+s_2^2-2s_1s_2\cos(\frac{\pi}{N})}\end{aligned} \tag{11}$$

The analytical expressions for the second estimator is:

$$\hat{\delta}_1^b = \frac{1}{2}\frac{|s_1|+|s_2|}{|s_1|-|s_2|} \tag{12}$$

Plugging sinusoidal Taylor series into (11) and (12) and neglecting the terms with order higher than $1/N^4$, we get the following equation for the first estimator:

$$\hat{\delta}_1^a \approx \frac{1}{2}\frac{\frac{2\pi^2}{N^2}\delta-\frac{\pi^4}{3N^4}(\delta+4\delta^3)}{g_1-g_2-g_3} \tag{13}$$

where

$$g_1 = \frac{\pi^2}{2N^2}(1+4\delta^2) \tag{14}$$

$$g_2 = \frac{\pi^4}{24N^4}(1+24\delta^2+16\delta^4) \tag{15}$$

$$g_3 = \frac{\pi^2}{2N^2}(-1+4\delta^2)+\frac{\pi^4}{24N^4}(7-24\delta^2-16\delta^4) \tag{16}$$

It can then be further simplified to:

$$\begin{aligned}\hat{\delta}_1^a &\approx \frac{1}{2}\frac{\frac{2\pi^2}{N^2}\delta-\frac{\pi^4}{3N^4}(\delta+4\delta^3)}{\frac{\pi^2}{N^2}-\frac{\pi^4}{3N^4}} \\ &\approx \delta+\frac{\pi^2}{6N^2}(\delta-4\delta^3)\end{aligned} \tag{17}$$

For the second estimator, we get:

$$\begin{aligned}\hat{\delta}_1^b &\approx \frac{1}{2}\frac{\frac{2\pi\delta}{N}-\frac{\pi^3}{3N^3}(\frac{3}{4}\delta+\delta^3)}{\frac{\pi}{N}-\frac{\pi^3}{3N^3}(\frac{1}{8}+\frac{3}{4}\delta^2)} \\ &= \delta-\frac{\pi^2}{12N^2}(\delta-4\delta^3)\end{aligned} \tag{18}$$

From the above analysis, we can see that the two versions of the estimator are not equivalent in terms of estimation bias because the bias of the second version, $\pi^2(\delta-4\delta^3)/(12N^2)$, is half of the bias of the first version.

Applying the same analysis, the distance for the second iteration can be expressed as follows:

$$h(\hat{\delta}_1^a) \approx -\frac{\pi^2}{6N^2}(\delta - 4\delta^3) - \frac{\pi^4}{36N^4}(\delta - 4\delta^3) \quad (19)$$

$$h(\hat{\delta}_1^b) \approx \frac{\pi^2}{12N^2}(\delta - 4\delta^3) - \frac{\pi^4}{144N^4}(\delta - 4\delta^3) \quad (20)$$

and the estimation for the second iteration is:

$$\hat{\delta}_2^a \approx \delta - \frac{\pi^4}{36N^4}(\delta - 4\delta^3) \quad (21)$$

$$\hat{\delta}_2^b \approx \delta - \frac{\pi^4}{144N^4}(\delta - 4\delta^3) \quad (22)$$

From the above two equations, we can conclude that the bias of the second version is a quarter of the first version at the second iteration. A second conclusion can be deduced from the above analysis is that the bias is proportional to $1/N^{2i}$ at the i-th iteration for both versions. To put it more exactly, in each iteration, the bias is reduced by $6N^2$ and $12N^2$ for the first and second version respectively.

3. PROPOSED METHOD

Equations (17) and (18) express the relationship between the estimated offset $\hat{\delta}$ and the exact offset δ in polynomial form. Therefore, a more accurate estimate can be obtained by solving these two polynomial equations with $\hat{\delta}$ as the input. Using this new estimate as the input of the second iteration, we can reduce the bias and make the estimation more accurate.

For this purpose, (17) can be rewritten as:

$$-\frac{2\pi^2}{3N^2}(\hat{\delta}_p^a)^3 + \left(1 + \frac{\pi^2}{6N^2}\right)\hat{\delta}_p^a - \hat{\delta}_1^a = 0 \quad (23)$$

and (18) can be rewritten as:

$$\frac{\pi^2}{3N^2}(\hat{\delta}_p^b)^3 + \left(1 - \frac{\pi^2}{12N^2}\right)\hat{\delta}_p^b - \hat{\delta}_1^b = 0 \quad (24)$$

We replace δ in (17) and (18) with $\hat{\delta}_p^a$ and $\hat{\delta}_p^b$ because the roots of the polynomials are now used as a new estimation of δ.

To be a valid solution, the root must be both real and in the range between -0.5 and 0.5. For the two polynomials, both have only one root that satisfies these conditions. We find that the solution for (23) of version 1 estimator is:

$$\hat{\delta}_p^a = \frac{(1 - j\sqrt{3})(1 + \frac{\pi^2}{6N^2})}{2^{\frac{2}{3}} D_a^{\frac{1}{3}}} - \frac{(1 + j\sqrt{3}) D_a^{\frac{1}{3}}}{2^{\frac{1}{3}} \frac{4\pi^2}{N^2}} \quad (25)$$

where D_a is

$$D_a = \frac{12\pi^4 \hat{\delta}_1^a}{N^4} + \sqrt{\frac{144\pi^8 (\hat{\delta}_1^a)^2}{N^8} - \frac{32\pi^6}{N^6}(1 + \frac{\pi^2}{6N^2})^3} \quad (26)$$

The solution for (24) of version 2 estimator is:

$$\hat{\delta}_p^b = \frac{D_b^{\frac{1}{3}}}{2^{\frac{1}{3}} \frac{\pi^2}{N^2}} - \frac{2^{\frac{1}{3}}(1 - \frac{\pi^2}{12N^2})}{D_b^{\frac{1}{3}}} \quad (27)$$

where D_b is

$$D_b = \frac{3\pi^4 \hat{\delta}_1^b}{N^4} + \sqrt{\frac{9\pi^8 (\hat{\delta}_1^b)^2}{N^8} + \frac{4\pi^6}{N^6}(1 - \frac{\pi^2}{12N^2})^3} \quad (28)$$

In summary, the first iteration of our proposed estimation algorithm can be accomplished in two steps. The first step is to obtain an initial estimation from (5) or (6). Then, plugging the initial estimation to (25) or (27) yields the second estimation of the first iteration. The second estimation is then used as the input for the second iteration.

4. EXPERIMENTAL RESULTS

All the experiments are conducted with $k_p = 2$, and $f_s = 1$. Fig. 1 and Fig. 2 show the behavior of the estimation bias for $N = 8$. The estimation bias is calculated as $(\delta - \hat{\delta})$ where δ is the true frequency offset and $\hat{\delta}$ is the estimated frequency offset. For Fig. 1, δ varies from -0.49 to 0.49 with step size 0.01. Since the bias is symmetric, the figures only show the part for positive δ.

In terms of the accuracy of theoretical prediction, Fig. 1(a) shows the true biases and the predicted biases in (17) and (18). We can see that the prediction is very accurate. To quantify the accuracy of the prediction, we calculate $|\text{bias}_{\text{predicted}} - \text{bias}_{\text{true}}|/|\delta|$. The maximum percentage is 0.047% for version 1 and 0.0068% for version 2.

In terms of improvements from the proposed method, both Fig. 1 and Fig. 2 show that the estimation from the proposed method is much more accurate than the A&M method in both the first and the second iterations. In Fig. 1 , the maximum distance between the proposed method and the A&M method is $10^{1.9519}$ for version 1 and $10^{2.5829}$ for version 2 in both iterations. The minimum distance is $10^{1.7568}$ for version 1 and $10^{1.7711}$ for version 2 in both iterations. Fig. 2 shows how the estimation bias varies with N at $\delta = 0.25$. We can see that as N gets larger, the gap between the proposed method and the A&M method also gets larger. At $N = 8$ and in the second iteration, the distance between the proposed method and the A&M method is $10^{1.7971}$ for version 1 and $10^{2.1911}$ for version 2. At $N = 256$, the distance becomes $10^{4.7839}$ for version 1 and $10^{5.1818}$ for version 2 in the second iteration. In summary, the proposed method does indeed provide significant reduction on estimation bias as compared to the A&M method.

Fig. 3 shows the performance under noise. White Gaussian noise is added to the complex sinusoidal signal with $\delta = 0.1$ and 0.4. The error of the estimated frequency is calculated as $(f_0 - \hat{f}_0)$ where $\hat{f}_0$ is calculated as in (4) and f_0 is

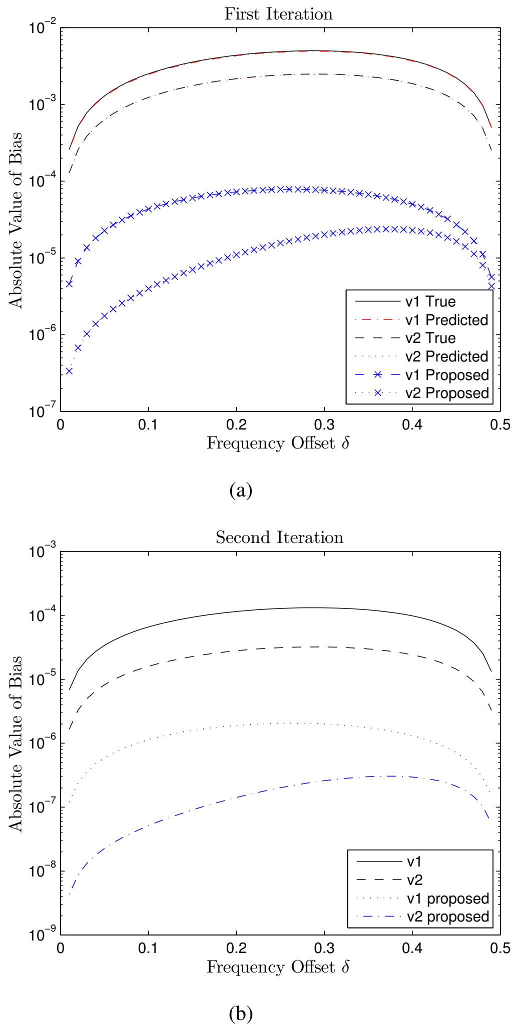

Fig. 1. Absolute value of bias vs. frequency offset δ: (a) first iteration and (b) second iteration.

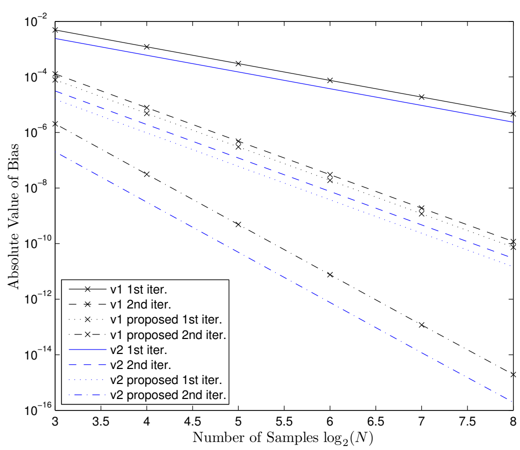

Fig. 2. Absolute value of bias vs. number of samples $log_2(N)$.

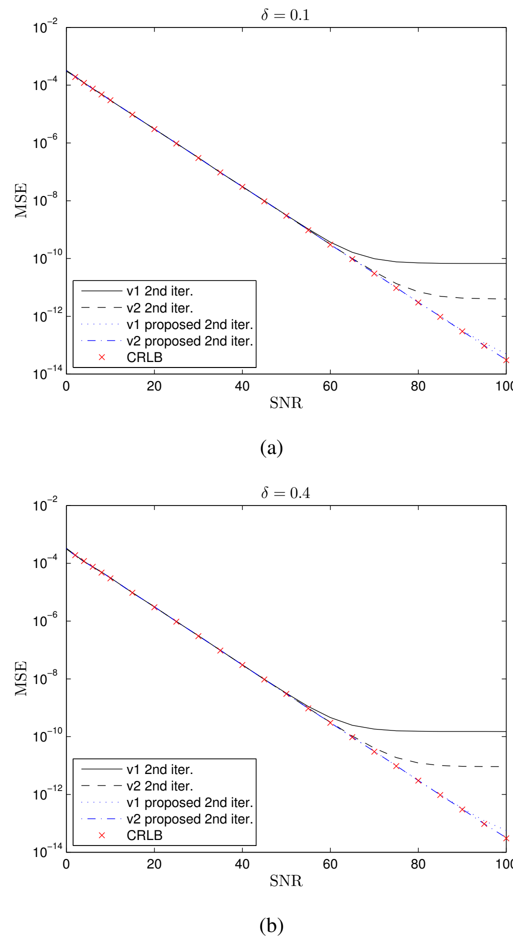

Fig. 3. MSE of estimated frequency vs. SNR: (a) $\delta = 0.1$ and (b) $\delta = 0.4$.

the true frequency. Each experiment is repeated 10,000 times and MSE of the estimated frequency is calculated. From the figure, we can see that the two versions of the proposed estimator follow CRLB closely in the entire range of SNR while the two versions of the A&M estimator become flat after a threshold.

5. CONCLUSION

In this paper, we theoretically analyze the two versions of the A&M estimator and show that the estimation biases of the two versions are not equivalent. The results of the theoretical analysis indicate that the bias of the first iteration can be expressed as a third-order polynomial equation. We then propose to solve the third-order polynomial equation to reduce the bias of the first iteration and, thereby, reduce the biases of the later iterations. Experiments show that the proposed new estimator can significantly reduce the bias and make the estimator follow CRLB in a much broader range of SNR than the A&M estimator.

6. REFERENCES

[1] D. Rife, R. Boorstyn, "Single tone parameter estimation from discrete-time observations," IEEE Trans. Inf. Theory, vol. 20, no. 5, pp. 591-598, Sep. 1974.

[2] B. G. Quinn, "Estimating frequency by interpolation using Fourier coefficients," IEEE Trans. Signal Process., vol. 42, no. 5, pp. 1264-1268, May 1994.

[3] 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.

[4] Y. V. Zakharov, V. M. Baronkin, T. C. Tozer, "DFT-based frequency estimators with narrow acquisition range," IEE Proc. Commun., vol. 148, no. 1, pp. 1-7, Feb. 2001.

[5] C. Candan, "A method for fine resolution frequency estimation from three DFT samples," IEEE Signal Process. Lett., vol. 18, no. 6, pp. 351-354, Jun. 2011.

[6] Y. V. Zakharov, T. C. Tozer, "Frequency estimator with dichotomous search of periodogram peak," Electron. Lett., vol. 35, no. 19, pp. 1608-1609, Sep. 1999.

[7] E. Aboutanios, "A modified dichotomous search frequency estimator," IEEE Signal Process. Lett., vol. 11, no. 2, pp. 186-188, Feb. 2004.

[8] S. Djukanovic, I. Djurovic, "Robust M-periodogram with dichotomous search," Signal Process., vol. 91, no. 10, pp. 2410-2414, Oct. 2011.

[9] T. Abatzoglou, "A fast maximum likelihood algorithm for frequency estimation of a sinusoid based on Newton's method," IEEE Trans. Acoust. Speech Signal Process., vol. 33, no. 1, pp. 77-89, Feb. 1985.

[10] S. Reisenfeld, E. Aboutanios, "A new algorithm for the estimation of the frequency of a complex exponential in additive Gaussian noise," IEEE Commun. Lett., vol. 7, no. 11, pp. 549-551, Nov. 2003.

[11] E. Aboutanios, B. Mulgrew, "Iterative frequency estimation by interpolation on Fourier coefficients," IEEE Trans. Signal Process., vol. 53, no. 4, pp. 1237-1242, Apr. 2005.

[12] I. Djurovic, "Estimation of the Sinusoidal Signal Frequency Based on the Marginal Median DFT," IEEE Trans. Signal Process., vol. 55, no. 5, pp. 2043-2051, May 2007.

[13] I. Djurovic, V. V. Lukin, "Estimation of single-tone signal frequency by using the L-DFT," Signal Process., vo. 87, no. 6, pp. 1537-1544, Jun. 2007.

[14] S. Minhas, E. Aboutanios, "Estimation of the frequency of a complex exponential," Proc. IEEE ICASSP, 2008, pp. 3693-3696.

[15] Y. Liu, Z. Nie, Z. Zhao, Q. H. Liu, "Generalization of iterative Fourier interpolation algorithms for single frequency estimation," Digital Signal Process., vol. 21, no. 1, pp. 141-149, Jan. 2011.