ACCURATE ESTIMATION OF FREQUENCY OF A SINGLE SINUSOID BASED ON DOWNSAMPLING

Weize Sun, H.C. So and Yuan Chen

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

weizesun2@student.cityu.edu.hk, hcso@ee.cityu.edu.hk, yuanchen3@student.cityu.edu.hk

ABSTRACT

A new phase-based approach for frequency estimation of a single cisoid in the presence of additive white noise is proposed in this paper. The main idea is to divide the observed data into a number of segments by downsampling and exploit this new structure for parameter estimation. The maximum likelihood estimator for frequency is then developed, which is shown to be superior to conventional phase-based methods in terms of uniform performance. Computer simulations also illustrate that the mean square frequency error of the proposed scheme can attain Cramér-Rao lower bound for sufficiently high signal-to-noise ratio conditions.

***Index Terms*—** frequency estimation, phase unwrapping, maximum likelihood estimation

1. INTRODUCTION

Estimation of the frequency of a single sinusoid in additive white Gaussian noise is a long studied problem in communications and signal processing. Generally speaking, frequency estimation can be achieved by either nonparametric or parametric means [1]. In case of a single complex tone, a standard parametric methodology, namely, phase-based approach [2], is to utilize the angles of the observed data. Tretter [3] was the first of this proposal by noting that the phase of a cisoid is a linear function of frequency, which is modulated by 2π. That is, in order to achieve accurate estimation, the phase must be correctly unwrapped. A simple strategy is to compute the phase difference of adjacent received signals so that the resulting values follow a moving average process and the frequency can be retrieved by weighted phase average (WPA) [4]. This scheme in fact avoids the procedure of phase unwrapping, and has the advantages of performing well when the signal-to-noise ratio (SNR) is large. However, the WPA will fail when the frequency is close to $\pm\pi$ and is shown to be inconsistent [5]. Other approaches such as [5]-[7], which perform the unwrapping of the phases in different ways, can be regarded as time-domain processing techniques. Nevertheless, there are two major drawbacks of employing the phase unwrapping techniques, namely, the high threshold SNR and the fact that the estimation accuracy is not independent of the frequency. In this work, we contribute to derive a downsampling based approximate maximum likelihood (ML) approach to overcome the latter problem.

The rest of this paper is organized as follows. One general strategy for the phase unwrapping based algorithms is presented in Section 2, then the proposed estimator is derived in Section 3. In Section 4, simulation results are included to evaluate the performance of the developed approach by comparing with the WPA [4] and phase-based ML [7] estimator, as well as Cramér-Rao lower bound (CRLB). Finally, conclusions are drawn in Section 5.

2. A MODEL AND ALGORITHM

The signal model is:

$$y(n) = x(n) + q(n) \tag{1}$$

$$x(n) = Ae^{j(\omega n+\theta)}, \quad n = 0, 1, \cdots, N-1 \tag{2}$$

where ω, $\theta \in [-\pi, \pi)$ and A are the unknown frequency, phase and signal amplitude, while $q(n)$ is a zero mean complex white Gaussian noise with variance σ^2. Here we are interested in finding the frequency from the N samples of $y(n)$. We define $\angle y(n) = (\omega n+\theta+\angle q(n)) \mod (2\pi) \in [-\pi, \pi)$ as the phase of $y(n)$ before phase unwrapping while $\angle\bar{y}(n) = (\omega n + \theta + \angle q(n))$ being the phase with ideal phase unwrapping, where $\angle a$ represents the angle of a. In published works such as [5] [7], the authors propose different estimation functions, which can be generalized as f, to get the estimates from n samples of $\{y(i)\}$, $i = 0, 1, \cdots, n-1$:

$$\begin{bmatrix}\hat{\omega}^{(n-1)} & \hat{\theta}^{(n-1)}\end{bmatrix} = f(\{y(i), \angle\bar{y}(i)\}) \tag{3}$$

Once a new sample $y(n)$ is received, an easy way ro determine $\angle\bar{y}(n)$ is to choose an appropriate integer C [7] so that

$$\angle\bar{y}(n) = \angle y(n) + 2\pi C \tag{4}$$

lies in the interval $[n\hat{\omega}^{(n-1)}+\hat{\theta}^{(n-1)}-\pi, n\hat{\omega}^{(n-1)}+\hat{\theta}^{(n-1)}+\pi)$, and then an updated estimate of $\hat{\omega}^{(n)}$ and $\hat{\theta}^{(n)}$ can be obtained using (3) with the $(n+1)$ samples. By an initial choice of $\hat{\omega}^{(0)} = 0$ and $\hat{\theta}^{(0)} = \angle\bar{y}(0) = \angle y(0)$, as long as the SNR is sufficiently high and ω is not close to $\pm\pi$, (4) can unwrap the phase $\angle\bar{y}(n)$ correctly and then the algorithm goes on when a new sample comes in, therefore is suitable for real time computation.

One of the major problems of this kind of algorithms is that their performance is dependent on ω [7]. These approaches perform well when ω is close to 0, but might fail if the frequency is close to $\pm\pi$, as a small noise will make the actual value of $n\omega+\theta$ falling outside of the boundaries of the interval $[n\hat{\omega}^{(n-1)}+\hat{\theta}^{(n-1)}-\pi, n\hat{\omega}^{(n-1)}+\hat{\theta}^{(n-1)}+\pi)$. This further leads to an inaccurate phase unwrapping of $\angle\bar{y}(n)$ in (4), then a degradation of performance. In next section, we first propose a general idea to overcome this problem and then apply it to derive an approximate ML frequency estimator.

3. PROPOSED ALGORITHM

3.1. A General Solution

Consider the downsampled version of $y(n)$:

$$y_m(n_m) = x_m(n_m) + q_m(n_m), m = 0, 1, \cdots, M-1 \quad (5)$$

$$\begin{aligned} x_m(n_m) &= x(Mn_m + m) \\ &= Ae^{j(M\omega n_m + m\omega + \theta)} = Ae^{j(\mu n_m + \phi_m)} \end{aligned} \quad (6)$$

That is, we separate the whole data set into M subsets where

$$\mu = (M\omega) \mod (2\pi) = M\omega + 2\pi B \in [-\pi, \pi) \quad (7)$$

$$\phi_m = (\theta + m\omega) \mod (2\pi) \quad (8)$$

are the new common frequency and phase for the mth subset and B is an integer. By doing this segmentation, although the original frequency is fixed, we can choose an appropriate M to make μ not close to $\pm\pi$. Furthermore, instead of estimating $\hat{\omega}$ and $\hat{\theta}$ from $y(n)$, we retrieve $\hat{\mu}$ and the new phase vector $\widehat{\mathbf{\Phi}} = \begin{bmatrix}\hat{\phi}_0 & \hat{\phi}_1 & \cdots & \hat{\phi}_{M-1}\end{bmatrix}$ from M subsets of $\{y_m(n_m)\}$:

$$\begin{bmatrix}\hat{\mu} & \widehat{\mathbf{\Phi}}\end{bmatrix} = f(\{y_m(n_m), \angle\bar{y}_m(n_m)\}_{m=0}^{M-1}) \quad (9)$$

It will be shown that using the above different structure of $y(n)$, the derived estimator will not degrade, and the final estimate of $\hat{\omega}$ can be determined as

$$\hat{\omega} = \frac{\hat{\mu} - 2\pi B}{M} \quad (10)$$

Defining $\hat{\theta}_m = \hat{\phi}_m - m\hat{\omega} + 2\pi B_m \in [-\pi, \pi)$ where B_m, $m = 0, 1, \cdots, M-1$, are integers, according to (8), the final estimate of the phase becomes

$$\hat{\theta} = \frac{\sum_{m=0}^{M-1}\hat{\theta}_m}{M} \quad (11)$$

Now the problem is to select an M to make (5)–(11) work. Here we propose to apply the weighted linear prediction (WLP) technique [4] to the first R samples to get a rough estimatie of w, denoted by $\hat{\omega}^{\text{rough}}$, and set

$$\omega_r = (r\hat{\omega}^{\text{rough}}) \mod (2\pi), \quad r = 1, 2, \cdots, R \quad (12)$$

We further define

$$M = \arg\min_{r=1,2,\cdots,R} |\omega_r| \quad (13)$$

That is, choosing μ as the frequency closest to 0 among $\{\omega_r\}$. In Section 4, we show that even setting $R = 5$ in the following approximate ML methodology can achieve accurate and uniform frequency estimation performance.

3.2. Approximate ML Estimator

We now consider the ML criterion, which is widely used in parameter estimation, to derive the estimation function f in (9).

At each time point $n = N-1$, we want to estimate μ and $\mathbf{\Phi}$ based on all the N received samples $y(n)$ up to time $N-1$, which is, from all the subsets $y_m(n_m)$ defined in (5). Decomposing $N-1 = ML+m$ where $m \in [0, M-1]$ and L are integers. We notice that each subset $y_m(n_m)$ has a length L_m where $L_0 = L_1 = \cdots = L_m = L+1$, $L_{m+1} = L_{m+2} = \cdots = L_{M-1} = L$ and $\sum_{m=0}^{M-1} L_m = N$. Writing

$$\mathbf{y} = \begin{bmatrix}\mathbf{y}_0^T & \mathbf{y}_1^T & \cdots & \mathbf{y}_{M-1}^T\end{bmatrix} \quad (14)$$

$$\mathbf{y}_m = \begin{bmatrix}y_m(0) & y_m(1) & \cdots & y_m(L_m-1)\end{bmatrix}^T \quad (15)$$

and the signal and noise parts being $\mathbf{x}$ and $\mathbf{q}$, respectively. As long as the noises are independent and identically distributed, the ML estimates, namely, $\hat{\mu}$ and $\widehat{\mathbf{\Phi}}$ correspond to maximizing

$$\begin{aligned} &p(\mathbf{x}|\mu, \mathbf{\Phi}) \\ =&\frac{1}{(\pi\sigma^2)^N}\exp(-\frac{\sum_{m=0}^{M-1}\sum_{n_m=0}^{L_m-1}|y_m(n_m) - Ae^{j(\mu n_m+\phi_m)}|^2}{\sigma^2}) \\ =&D\exp(\frac{2A}{\sigma^2}\sum_{m=0}^{M-1}\sum_{n_m=0}^{L_m-1}\cos(\angle\bar{y}_m(n_m) - (\mu n_m + \phi_m))) \end{aligned} \quad (16)$$

where $D = 1/(\pi\sigma^2)^N e^{-\sum_{m=0}^{M-1}\sum_{n_m=0}^{L_m-1}(|y_m(n_m)|^2+A^2)/\sigma^2}$ is a constant. Taking the natural logarithm of both sides of (16) yields

$$\begin{aligned} \Lambda(\mu, \mathbf{\Phi}) &= \ln p(\mathbf{x}|\mu, \mathbf{\Phi}) \\ &= \ln D + \frac{2A}{\sigma^2}\sum_{m=0}^{M-1}\sum_{n_m=0}^{L_m-1}|y_m(n_m)| \\ &\quad \cos(\angle\bar{y}_m(n_m) - (\mu n_m + \phi_m)) \end{aligned} \quad (17)$$

Differentiating $\Lambda(\mu, \mathbf{\Phi})$ with respect to μ and ϕ_m and equating the resultant expressions to zero result in

$$\begin{aligned} \frac{\partial\Lambda(\mu, \mathbf{\Phi})}{\partial\mu} =& \frac{2A}{\sigma^2}\sum_{m=0}^{M-1}\sum_{n_m=0}^{L_m-1} n_m|y_m(n_m)| \\ & \sin(\angle\bar{y}_m(n_m) - (\mu n_m + \phi_m)) = 0 \end{aligned} \quad (18)$$

$$\frac{\partial \Lambda(\mu, \mathbf{\Phi})}{\partial \phi_m} = \frac{2A}{\sigma^2} \sum_{n_m=0}^{L_m-1} |y_m(n_m)| \sin(\angle \bar{y}_m(n_m) - (\mu n_m + \phi_m)) = 0 \quad (19)$$

for $m = 0, 1, \cdots, M-1$. The sine function in (18) and (19) makes the equations highly nonlinear, therefore a closed form solution does not exist. To solve this problem, we assume that the SNR, defined as A^2/σ^2, is sufficiently high, then the phase error $\angle \bar{y}_m(n_m) - (\mu n_m + \phi_m) \to 0$. By approximating $\sin x \approx x$ for small x, (18) and (19) become

$$\frac{2A}{\sigma^2} \sum_{m=0}^{M-1} (P_{m,1} - P_{m,2}\mu - P_{m,3}\phi_m) = 0 \quad (20)$$

$$\frac{2A}{\sigma^2} (P_{m,4} - P_{m,3}\mu - P_{m,5}\phi_m) = 0 \quad (21)$$

$$P_{m,1} = \sum_{n_m=0}^{L_m-1} n_m |y_m(n_m)| \angle \bar{y}_m(n_m) \quad (22)$$

$$P_{m,2} = \sum_{n_m=0}^{L_m-1} n_m^2 |y_m(n_m)| \quad (23)$$

$$P_{m,3} = \sum_{n_m=0}^{L_m-1} n_m |y_m(n_m)| \quad (24)$$

$$P_{m,4} = \sum_{n_m=0}^{L_m-1} |y_m(n_m)| \angle \bar{y}_m(n_m) \quad (25)$$

$$P_{m,5} = \sum_{n_m=0}^{L_m-1} |y_m(n_m)| \quad (26)$$

for $m = 0, 1, \cdots, M-1$. Solving all the equations yields

$$\hat{\mu} = \frac{\sum_{m=0}^{M-1} Q_{m,1}}{\sum_{m=0}^{M-1} Q_{m,2}} \quad (27)$$

$$\hat{\phi}_m = \frac{P_{m,4} - P_{m,3}\hat{\mu}}{P_{m,5}} \quad (28)$$

where

$$Q_{m,1} = P_{m,1} - \frac{P_{m,3}P_{m,4}}{P_{m,5}} \quad (29)$$

$$Q_{m,2} = P_{m,2} - \frac{P_{m,3}^2}{P_{m,5}} \quad (30)$$

indicating that when a new sample at time $n = N-1 = ML + m$ comes in, only the corresponding subset $\mathbf{y}_m$ is changed, and we only need to update P_{m,i_1} and Q_{m,i_2}, $i_1 = 1, 2, \cdots, 5$ and $i_2 = 1, 2$, to compute the new estimates $\hat{\mu}$, $\hat{\phi}_m$, then $\hat{\omega}$ can be solved using (10). It is also worthy to notice that this approach is a generalization of [7]. That is, when $M = 1$, it is reduced to [7]. Finally, the algorithm of the proposed estimator is summarized in Table 1.

4. SIMULATION RESULTS

Computer simulations have been carried out to evaluate the performance of the proposed algorithm in the presence of white Gaussian noise. The mean square frequency error (MSFE) is assigned to evaluate the algorithm performance. All results provided are averages of 2000 independent runs.

In the first test, we study the performance of the proposed algorithm comparing with the WPA [4] and phase-based ML estimator [7]. The tone amplitude and frequency are $A = 1$ and $\omega = 0.7\pi$ while the phase varies from $-\pi$ to π in each independent run. We choose $R = 5$ in the proposed algorithm and three data lengths, namely, $N = 10$, $N = 30$ and $N = 100$ are tested, while the MSFEs versus SNR are plotted in Figure 1. The proposed approach performs the best in the sense of achieving the threshold SNR of 10dB in all the three cases. It is also shown that the threshold performance of the proposed estimator is independent of the length of the data, indicating that it in fact only depends on selection of M and the accuracy of the phase unwrapping procedure of (4) under the new frequency μ, which in turn depends on the SNR. This result also indicates that the idea of dividing the whole data set into several subsets will not degrade the performance as long as the number of subsets, M, is chosen to make μ not close to $\pm\pi$.

Figures 2 to 4 plot the frequency versus SNR contours of MSFE for the WPA, phase-based ML and proposed methods under $N = 30$. It is observed that proposed scheme has a desirable merit of uniform performance. It also has the best threshold SNR of 10-14dB when ω varies from -0.9π to 0.9π, among all the methods, which agrees with the first test. Notice that when the true frequency is close to 0, the proposed method performs almost the same as the phase-based ML algorithm [7]. It is because in that case, the system will choose $M = 1$ and it is reduced to [7].

---

1. Use the first R samples to determine M by (12) and (13). Set the initial estimates $\hat{\omega} = \hat{\mu} = 0$ and $\hat{\theta}_m = \angle y_m(n_m)$ for $m = 0, 1, \cdots, M-1$.
2. For time point $n = N-1 > M-1$, stack the new coming sample $y(n)$ in the end of the corresponding subset $\{y_m(n_m)\}$ according to (6), get $\angle \bar{y}_m(n_m)$ from (4) by substituting $\hat{\omega}$ and $\hat{\theta}$ with $\hat{\mu}$ and $\hat{\phi}_m$.
3. Calculate the new P_{m,i_1} and Q_{m,i_2}, then get updated $\hat{\mu}$ and $\hat{\phi}_m$ using (22)–(30).
4. Update $\hat{\omega}$ by (10).
5. Repeat 2-4 every time a new sample comes in.

Table 1. Approximate ML Frequency Estimation Algorithm

5. CONCLUSION

In this paper, we first discuss a general procedure of the phase-based frequency estimation algorithms, and then an idea of dividing the whole data set into several subsets is introduced. The maximum likelihood criterion is also applied to derived a new estimator based on this downsampling technique. Computer simulation shows that the segmentation of the data does not degrade the performance of the estimator, and the proposed scheme performs outstandingly in terms of estimation accuracy under a wide range of frequency.

6. REFERENCES

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

[2] M.L. Fowler, “Phase-based frequency estimation: A review,” *Digital Signal Processing*, vol. 12, no. 4, pp. 590-615, Oct. 2002.

[3] S.A. Tretter, “Estimating the frequency of a noisy sinusoid by linear regression,” *IEEE Transactions on Information Theory*, vol. 31, no. 6, pp. 832-835, Nov. 1985.

[4] S. Kay, “A fast and accurate single frequency estimator,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol.37, no.12, pp.1987-1990, Dec. 1989

[5] R.G. McKilliam, B.G. Quinn, I.V.L. Clarkson, and B. Moran, “Frequency estimation by phase unwrapping,” *IEEE Transactions on Signal Processing*, vol.58, no.6, pp.2953-2963, Jun. 2010

[6] P.Y. Kam, “Maximum likelihood carrier phase recovery for linear suppressed-carrier digital data modulations,” *IEEE Transactions on Communications*, vol. 34, no. 6, pp. 522-527, Jun. 1986.

[7] H. Fu and P.Y. Kam, “MAP/ML estimation of the frequency and phase of a single sinusoid in noise,” *IEEE Transactions on Signal Processing*, vol.55, no.3, pp.834-845, Mar. 2007

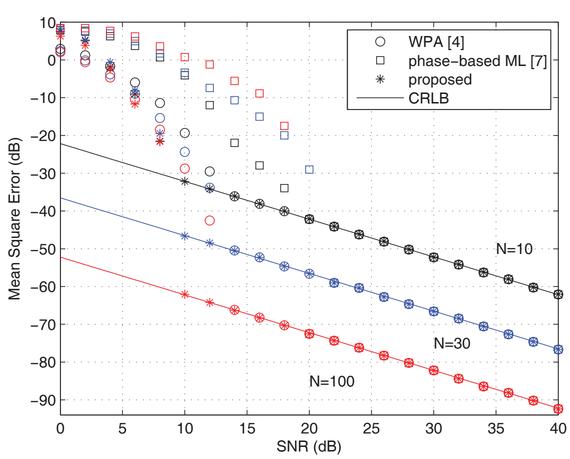

Fig. 1. Mean square frequency error versus SNR for different N.

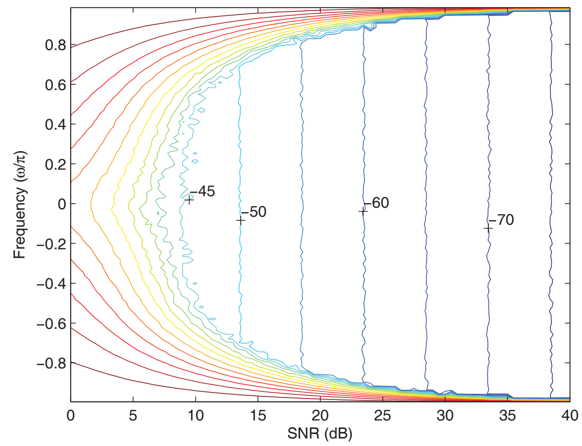

Fig. 2. Contour plot of WPA.

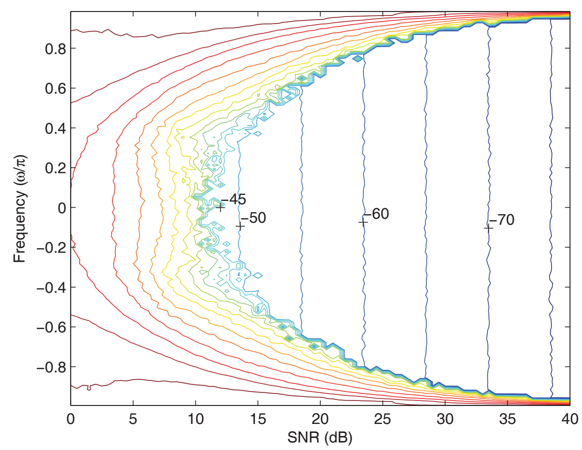

Fig. 3. Contour plot of phase-based ML method.

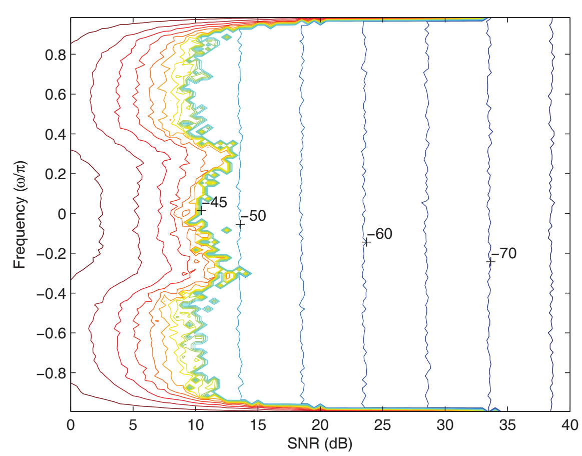

Fig. 4. Contour plot of proposed estimator