EURASIP Journal on Applied Signal Processing 2004:12, 1791–1806

Adaptive Window Zero-Crossing-Based Instantaneous Frequency Estimation

S. Chandra Sekhar

Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560 012, India
Email: schash@protocol.ece.iisc.ernet.in

T. V. Sreenivas

Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560 012, India
Email: tvsree@ece.iisc.ernet.in

Received 2 September 2003; Revised 2 March 2004

We address the problem of estimating instantaneous frequency (IF) of a real-valued constant amplitude time-varying sinusoid. Estimation of polynomial IF is formulated using the zero-crossings of the signal. We propose an algorithm to estimate nonpolynomial IF by *local* approximation using a low-order polynomial, over a short segment of the signal. This involves the choice of window length to minimize the mean square error (MSE). The optimal window length found by directly minimizing the MSE is a function of the higher-order derivatives of the IF which are not available a priori. However, an optimum solution is formulated using an adaptive window technique based on the concept of *intersection of confidence intervals*. The adaptive algorithm enables minimum MSE-IF (MMSE-IF) estimation without requiring a priori information about the IF. Simulation results show that the adaptive window zero-crossing-based IF estimation method is superior to fixed window methods and is also better than adaptive spectrogram and adaptive Wigner-Ville distribution (WVD)-based IF estimators for different signal-to-noise ratio (SNR).

Keywords and phrases: zero-crossing, irregular sampling, instantaneous frequency, bias-variance tradeoff, confidence interval, adaptation.

1. INTRODUCTION

Almost all information carrying signals are time-varying in nature. The adjective "*time-varying*" is used to describe an "*attribute*" of the signal that is changing/evolving in time [1]. For most signals such as speech, audio, biomedical, or video signals, it is the spectral content that changes with time. These signals contain time-varying spectral attributes which are a direct consequence of the signal generation process. For example, continuous movements of the articulators, activated by time-varying excitation, is the cause of the time-varying spectral content in speech signals [2, 3]. In addition to these naturally occurring signals, man-made modulation signals, such as frequency-shift keyed (FSK) signals used for communication [4] carry information in their time-varying attributes. Estimating these attributes of a signal is important both for extracting their information content as well as synthesis in some applications.

Typical attributes of time-varying signals are amplitude modulation (AM), phase/frequency modulation (FM) of a sinusoid. Another time-varying signal model is the output of a linear system with time-varying impulse response. However, the simplest and fundamental signal processing model for time-varying signals is an AM-FM combination [5, 6, 7] of the type $s(t) = A(t)\sin(\phi(t))$. Further, if the amplitude does not vary with time, the signal is simplified to $s(t) = A\sin(\phi(t))$. Estimating the IF of such signals is a well-studied problem with limited performance for arbitrary IF laws and low SNR conditions [8, 9].

In [9], a novel auditory motivated level-crossing approach has been developed for estimating instantaneous frequency (IF) of a polynomial nature, that is, the instantaneous phase (IP), $\phi(t)$ is of the form $\phi(t) = \sum_{k=0}^{p} a_k t^k$ and the IF is given by $f(t) = (1/2\pi)(d\phi(t)/dt)$. In this paper, we address the estimation of IF of nonpolynomial nature of monocomponent phase signals, in the presence of noise, using zero-crossings of $s(t)$. We achieve this by performing *local* polynomial approximation to the IF using the zero-crossings (ZCs). This involves the choice of optimum window length to minimize the mean square error (MSE). The minimum MSE (MMSE) formulation gives rise to an optimum window length solution which requires a priori information about the IF. Also, the length of the window introduces a "*bias-variance tradeoff*" which is resolved using an adaptive approach [10, 11, 12] based on the *intersection of confidence intervals* of the zero-crossing-based IF (ZC-IF) estimator.

Fundamental contributions related to the ZCs of amplitude and frequency modulated signals were made in [13], wherein the factorization of an analytic signal in terms of real and complex time-domain zeros was proposed. A model based pole-zero product representation of an analytic signal was proposed in [14]. Recent contributions include the use of homomorphic signal processing techniques for factorization of real signals [15]. In contrast to these, we use the real ZCs of the signal $s(t)$ that can be directly estimated from its samples. The use of zero-crossing (ZC) information is a nonlinear approach to estimating IF; this has been reported earlier using either ZC rate information [8, 16] or ZC interval histogram information [17, 18] (in the context of speech signals). These earlier approaches are quasistationary and are inherently limited to estimating only mild frequency variations. The new approach developed in this paper fits a local, nonstationary model for the IF and uses the ZC *instant* information [19] for IF estimation.

This paper is organized as follows. In Section 2, we formulate the problem. In Section 3, we discuss ZC-based polynomial IF estimation and the need for local polynomial approximation for nonpolynomial IF estimation. Bias and variance of the ZC-IF estimator are derived in Section 4. The problem of optimal window length selection is addressed in Section 5 and an adaptive algorithm is discussed in Section 6. Simulation results are presented in Section 7. Section 8 concludes the paper.

2. IF ESTIMATION PROBLEM

Let $s(t) = A\sin(\phi(t))$ be the phase signal with constant amplitude and IF[1] is given by $f(t) = (1/2\pi)(d\phi(t)/dt)$. Let the frequency variation be bounded, but arbitrary and unknown. The signal $s(t)$ has strictly infinite bandwidth, but we assume that it is essentially bandlimited to $[-B\pi, B\pi]$. Let $s(t)$ be corrupted additively with Gaussian noise, $w(t)$, which has a flat power spectral density, $S_{ww}(\omega) = \sigma_w^2$ for $|\omega| \le B\pi$ and zero elsewhere. $w(t)$ is therefore bandlimited in nature. However, samples taken from this process at a rate of B samples/second are uncorrelated. Let the noisy signal be denoted by $y(t) = s(t) + w(t)$ for $t \in [0, T]$. The noisy signal when sampled at a rate of B samples/second yields the discrete-time observations $y[nT_s] = s[nT_s] + w[nT_s]$, where T_s is the sampling period. We normalize the sampling period to unity and write equivalently, $y[n] = s[n] + w[n]$ or $y[n] = A\sin(\phi[n]) + w[n]$, $0 \le n \le N-1$, where N is the number of discrete-time observations. The noise $w[n]$ is white Gaussian with a variance σ_w^2. The signal-to-noise ratio (SNR) is defined as $\text{SNR} = A^2/2\sigma_w^2$. The problem is to estimate the IF of the signal $s(t)$ using the samples $y[n]$ and estimating the ZCs of the signal, $y(t)$. Negative IF is only conceptual; naturally occurring IF is always positive and hence we confine our discussion to positive IF.

[1] It must be noted that this definition of IF is different from that obtained using the Hilbert transform.

3. ZC-IF ALGORITHM

Let the ZCs of the noise-free signal, $s(t) = A\sin(\phi(t))$, $t \in [0, T]$, be given by $\mathcal{Z} = \{t_j \mid s(t_j) = 0;\ j = 0, 1, 2, \ldots, Z\}$, where $Z+1$ is the number of ZCs of $s(t)$ in $[0, T]$. Correspondingly, the values of the phase function $\phi(t)$ are given by $\mathcal{P} = \{j\pi;\ j = 0, 1, 2, \ldots, Z\}$. The phase value corresponding to the first ZC over $[0, N-1]$ has been arbitrarily assigned to 0. This does not affect IF estimation because of the derivative operation. If the phase function $\phi(t)$ is a polynomial of order p, $0 < p < Z$, of the form, $\phi(t) = \sum_{k=0}^{p} a_k t^k$, then, up to an additive constant, it can be *uniquely* recovered from the set of ZC instants $\mathcal{Z}$. This property of uniqueness eliminates the need for a Hilbert transform based definition of IF. Corresponding to each ZC instant, t_j, we have an equation $j\pi = \sum_{k=0}^{p} a_k t_j^k$, $0 \le j \le Z$. The set of $(Z+1)$ equations, in general, is more than the number of unknowns, p, and in the absence of ZC estimation errors, they are consistent. Due to arbitrary assignment of $\phi(t_0)$ to 0, the coefficient estimate of a_0 will be in error; however, this does not affect the IF estimate and the IF can be recovered uniquely.

In practice, since ZC instants of $s(t)$ have to be estimated using $s[n]$, there is a small, nonzero error.[2] In such a case, the coefficient vector, $\mathbf{a} = \{a_k,\ k = 0, 1, 2, \ldots, p\}$ can be estimated by minimizing the cost function $\mathcal{C}_p(\mathbf{a})$ defined as

$$\mathcal{C}_p(\mathbf{a}) = \frac{1}{Z+1} \sum_{j=0}^{Z} \left(j\pi - \mathbf{a}^T \mathbf{e_j}\right)^2, \tag{1}$$

where

$$\mathbf{a} = \{a_k,\ k = 0, 1, 2, \ldots, p\}, \mathbf{e_j} = \begin{bmatrix} 1 & t_j & t_j^2 & \cdots & t_j^p \end{bmatrix}^T \tag{2}$$

(T stands for transpose operator). The optimum coefficient vector is obtained in a straightforward manner as

$$\hat{\mathbf{a}} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{\Phi}, \tag{3}$$

where $\mathbf{\Phi}$ is a column vector whose jth entry is $j\pi$ and $\mathbf{H}$ is a matrix whose jth row is $\mathbf{e_j}^T$. $\hat{\mathbf{a}} = [\hat{a}_0\ \hat{a}_1\ \hat{a}_2\ \cdots\ \hat{a}_p]$. At the sample instants, the IF is estimated as $\hat{f}[n] = (1/2\pi)\sum_{k=1}^{p} k\hat{a}_k n^{k-1}$, $0 \le n \le N-1$. We refer to this as the ZC-IF estimator.

3.1. Performance of the ZC-IF estimator

To illustrate the performance of the ZC-IF algorithm, 256 samples of a quadratic IF signal were generated. The ZC instants were estimated using 10 iterations, each through the root-finding approach. The actual and the estimated IF corresponding to $p = 3$ are shown in Figures 1a, 1b, respectively. For the IF estimates corresponding to orders $p = 1, 2, \ldots, 8$, the following error measures were computed: IP curve fitting error:

[2] If two successive samples, $s[m]$ and $s[m+1]$, are of opposite sign, then the corresponding continuous-time signal, $s(t)$, has a ZC in the interval $[m, m+1]$. The ZC instant is estimated using bandlimited interpolation [20] and a bisection approach, similar to root-finding problems in numerical analysis [21].

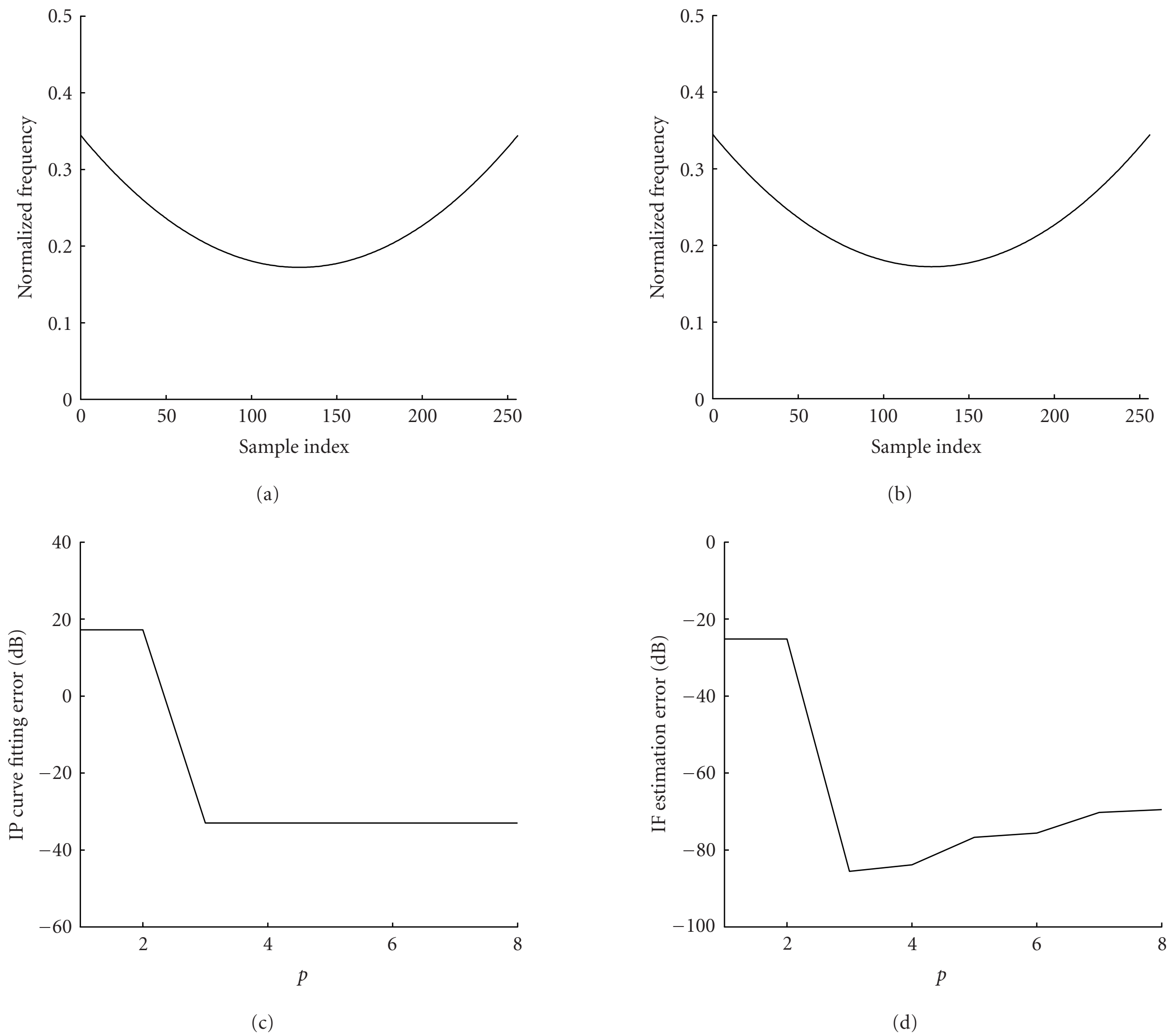

Figure 1: (a) Actual IF, (b) ZC-IF estimate (corresponding to $p = 3$), (c) IP curve fitting error (dB) versus p, and (d) IF estimation error (dB) versus p.

$$\mathcal{C}_p(\hat{\mathbf{a}}) = \frac{1}{Z+1}\sum_{j=0}^{Z}\left(j\pi - \hat{\mathbf{a}}^T\mathbf{e}_\mathbf{j}\right)^2, \tag{4}$$

IF estimation error:

$$\mathcal{J}_p(\hat{\mathbf{a}}) = \frac{1}{N}\sum_{n=0}^{N-1}\left(f[n] - \hat{f}[n]\right)^2. \tag{5}$$

It must be noted that $\mathcal{C}_p(\hat{\mathbf{a}})$ can be computed using the ZC information, whereas $\mathcal{J}_p(\hat{\mathbf{a}})$ can be computed only when the actual IF, $f[n]$, is known. Also, while $\mathcal{C}_p(\hat{\mathbf{a}})$ is a nonincreasing function of p, $\mathcal{J}_p(\hat{\mathbf{a}})$ need not be. These error measures are plotted in Figures 1c, 1d, respectively. From Figure 1c, it is clear that beyond $p = 3$ (cubic phase fitting or equivalently, quadratic IF), the error reduction is not appreciable. Thus, a measure of saturation of the IP fitting error can be used for order selection.

The algorithm works best when the actual IF and the assumed IF model are matched, that is, the underlying IF is a polynomial and the assumed IF model is also a polynomial of the same order. However, when there is a mismatch, that is, the underlying IF is not a polynomial but we approximate it using polynomials, the following problems arise.

(1) The choice of the order of the polynomial becomes crucial. A value of p that keeps the IP fitting error below a predetermined threshold does not necessarily yield the minimum IF estimation error. This problem occurs even when the underlying IF is a polynomial of unknown order as demonstrated in Figures 1c, 1d.

(2) Not all kinds of IF variations can be approximated by finite-order polynomials to a desired degree of accuracy.

(3) Fast IF variations in a given interval require very high polynomial orders and hence large amounts of data. However, this can often lead to numerically unstable set of equations in solving for the coefficients of the polynomial yielding erroneous and practically useless IF estimates.

A natural modification of the ZC-IF algorithm is to perform *local polynomial fitting*, that is, use lower-order polynomial functions to *locally* estimate the IF rather than use one large order polynomial over the entire observation window. If we always use a fixed low order polynomial, say $p = 3$, we are still faced with the question of window length selection; that is, *over what window length should a local polynomial approximation be performed*? An algorithm that helps us choose the appropriate window length should have the following features:

(1) require no a priori information about the IF,
(2) yield an IF estimate with the MMSE for all values of SNR.

The objective of this paper is to develop such an algorithm.

The relevant cost function is the MSE [22] of the estimate $\hat{f}$, of the quantity f, defined as $\text{MSE} = \mathcal{E}\{(\hat{f} - f)^2\}$, where $\mathcal{E}$ denotes the expectation operator. MSE can be rewritten as $\text{MSE} = (\mathcal{E}\{\hat{f}\} - f)^2 + \mathcal{E}\{(\hat{f} - \mathcal{E}\hat{f})^2\}$. The first term is the squared bias and the second term is the variance of the ZC-IF estimator. In the following sections, we obtain the bias and variance of the ZC-IF estimator and develop the algorithm.

4. BIAS AND VARIANCE OF THE ZC-IF ESTIMATOR

Consider the ZCs $\{t_0, t_1, t_2, t_3, \ldots, t_Z\}$ and let $\{\phi(t_0), \phi(t_1), \phi(t_2), \phi(t_3), \ldots, \phi(t_Z)\}$ be the associated instantaneous phase values. In the presence of noise, the ZC instants get perturbed to $\{t_0 + \delta t_0, t_1 + \delta t_1, t_2 + \delta t_2, t_3 + \delta t_3, \ldots, t_Z + \delta t_Z\}$. We assume that the SNR is high enough that the ZC instants get perturbed by a small amount and no additional ZCs are introduced. Corresponding to these perturbed time instants is the set of IP values $\{\phi(t_0 + \delta t_0), \phi(t_1 + \delta t_1), \phi(t_2 + \delta t_2), \phi(t_3 + \delta t_3), \ldots, \phi(t_Z + \delta t_Z)\}$. Using a first-order Taylor series approximation, we can write $\phi(t_j + \delta t_j) \approx \phi(t_j) + \phi'(t_j)\delta t_j$ ($'$ denotes derivative), that is, the perturbation in t_j is mapped to $\phi(t_j)$. The distribution of $\phi'(t_j)\delta t_j$ can be found as follows.

At the ZCs, the noisy signal $y(t_j) = A\sin(\phi(t_j)) + w(t_j)$ may be approximated as $y(t_j) \approx A\sin(\phi(t_j + \delta t_j)) \approx A\sin(\phi(t_j)) + A\cos(\phi(t_j))\phi'(t_j)\delta t_j$. Therefore, $\phi'(t_j)\delta t_j \approx w(t_j)/A = \tilde{w}(t_j)$. Hence the perturbations in $\phi(t_j)$ are also Gaussian distributed with variance σ_w^2/A^2. Thus, under a high SNR assumption, one can approximate the effect of additive noise on the signal samples to have an additive *phase noise* effect [23].

Let $t \in [0, T]$ be the point where the IF estimate is desired. The basic principle in the new approach to IF estimation is to fit a polynomial, *locally*, to the ZCs and IP values within an interval L about the point t. The IF is obtained by the derivative operation. We use a rectangular window symmetric about t, that is, choose the window function, as $h(\tau) = 1/L$ for $\tau \in [-L/2, +L/2]$ and zero elsewhere. The window function is normalized to have unit area. Define the set $\mathcal{I}_{t,L} = \{\tau \mid t - L/2 \le \tau \le t + L/2\}$ which is the set of all points within the L-length window centered at $\tau = t$.

Consider the quadratic cost function

$$\mathcal{C}(t, \mathbf{a}) = \sum_{n=i \ni t_n \in \mathcal{I}_{t,L}}^{j} \left[\phi(t_n) + \tilde{w}(t_n) - \sum_{k=0}^{p} a_k t_n^k\right]^2 h(t - t_n). \quad (6)$$

The coefficients $\{a_0, a_1, a_2, \ldots, a_p\}$ are specific to the time instant t and can be obtained as the minimizers of the above quadratic cost function. The optimal coefficient estimates are denoted by $\{\hat{a}_0, \hat{a}_1, \hat{a}_2, \ldots, \hat{a}_p\}$ and defined as

$$\hat{a}_\ell = \arg\min_{a_\ell} \mathcal{C}, \quad 0 \le \ell \le p. \quad (7)$$

In other words, $\hat{a}_\ell$ is a solution to $\partial\mathcal{C}/\partial a_\ell = 0$ or, equivalently, $\partial\mathcal{C}/\partial a_\ell|_{a_\ell = \hat{a}_\ell} = 0$. We have

$$\frac{\partial\mathcal{C}}{\partial a_\ell} = -2 \sum_{n=i,\, t_n \in \mathcal{I}_{t,L}}^{j} \left[\phi(t_n) + \tilde{w}(t_n) - \sum_{k=0}^{p} a_k t_n^k\right] h(t - t_n) t_n^\ell, \quad 0 \le \ell \le p. \quad (8)$$

The estimation error, $\Delta a_\ell = \hat{a}_\ell - a_\ell$, is due to the following:

(1) error due to additive noise, $\delta_{\tilde{w}}$,
(2) error due to mismatch between the actual phase and the estimated phase using a local polynomial model (residual phase error), $\delta_{\Delta\phi}$.

The minimum of the cost function therefore is perturbed due to noise and residual phase effects. We can rewrite $\partial\mathcal{C}/\partial a_\ell$ as follows:

$$\frac{\partial\mathcal{C}}{\partial a_\ell} = \left.\frac{\partial\mathcal{C}}{\partial a_\ell}\right|_0 + \left.\frac{\partial^2\mathcal{C}}{\partial a_\ell^2}\right|_0 \Delta a_\ell + \left.\frac{\partial\mathcal{C}}{\partial a_\ell}\right|_0 \delta_{\Delta\phi} + \left.\frac{\partial\mathcal{C}}{\partial a_\ell}\right|_0 \delta_{\tilde{w}}, \quad (9)$$

where $|_0$ indicates that the quantities are those corresponding to zero-phase error and absence of noise, that is, $\Delta\phi = 0$ and $\tilde{w}(t) = 0$. Unlike the results in [10, 11, 24], where the derivative of the time-frequency distribution (TFD) is nonquadratic and approximate linearization of the derivative around the peak is done, here, the cost function is quadratic and hence its derivative is linear in the parameters to be estimated. Therefore, the above linear equation is exact and not approximate. The terms $\partial\mathcal{C}/\partial a_\ell|_0 \delta_{\Delta\phi}$ and $\partial\mathcal{C}/\partial a_\ell|_0 \delta_{\tilde{w}}$ indicate the perturbations in the derivative as a result of phase error and noise, respectively. Evaluation of these quantities, bias, and variance computation of the IF estimates is given in the appendix. The asymptotic expressions for bias, variance, and covariance (denoted by $\text{Bias}(\cdot)$, $\text{Var}(\cdot)$, and $\text{Cov}(\cdot)$, respectively) of the coefficient estimates are given by

$$\begin{aligned}
&\text{Bias}(\Delta a_\ell) \\
&\quad = \frac{(-1)^{p+1}\phi^{(p+1)}(t)}{(p+1)!}\left[\frac{\int_{-L/2}^{+L/2} s^{p+1}(t-s)^\ell ds}{\int_{-L/2}^{+L/2}(t-s)^{2\ell} ds}\right], \quad 0 \le \ell \le p, \\
&\text{Cov}(\Delta a_\ell, \Delta a_k) \\
&\quad = \frac{\sigma_w^2}{A^2}\left[\frac{\int_{-L/2}^{+L/2}(t-s)^{\ell+k} ds}{\int_{-L/2}^{+L/2}(t-s)^{2\ell} ds \int_{-L/2}^{+L/2}(t-s)^{2k} ds}\right], \quad 0 \le \ell,\ k \le p, \\
&\text{Var}(\Delta a_\ell) = \frac{\sigma_w^2/A^2}{\int_{-L/2}^{+L/2}(t-s)^{2\ell} ds}, \quad 0 \le \ell \le p.
\end{aligned} \quad (10)$$

Using these, the bias and variance of the IF estimator can be obtained as

$$\begin{aligned}\operatorname{Bias}\left(\hat{f}(t)\right) &= \frac{1}{2\pi}\sum_{\ell=1}^{p}\operatorname{Bias}\left(\Delta a_\ell\right)\ell t^{\ell-1},\\ \operatorname{Var}\left(\hat{f}(t)\right) &= \frac{1}{4\pi^2}\sum_{\ell=1}^{p}\sum_{m=1}^{p}\ell m t^{\ell+m-2}\operatorname{Cov}\left(\Delta a_\ell, \Delta a_m\right).\end{aligned} \tag{11}$$

Directly substituting the expressions for bias and covariance of the coefficient estimates gives rise to very complicated expressions for the bias and variance of the ZC-IF estimator. However, a considerable simplification can be achieved by using the idea of *data centering* about the origin, that is, without loss of generality, assume that the data is shifted to lie in the interval $[-L/2, +L/2]$ instead of $[t - L/2, t + L/2]$. Data centering is very useful in obtaining simplified expressions for the bias and variance of IF estimators [25]. It must be noted that data centering is an adjustment to yield simplified expressions and the IF estimate is unaffected in doing so because the estimates are computed using the centered data. This yields the following expressions for bias and variance of the coefficients:

$$\begin{aligned}&\operatorname{Bias}\left(\Delta a_\ell\right)\\ &\quad= \frac{\phi^{(p+1)}(0)(-1)^{p+\ell+1}}{(p+1)!}\left[\frac{\int_{-L/2}^{+L/2}\tau^{p+\ell+1}d\tau}{\int_{-L/2}^{+L/2}\tau^{2\ell}d\tau}\right], \quad 0 \le \ell \le p,\\ &\operatorname{Var}\left(\Delta a_\ell\right)\\ &\quad= \frac{\sigma_w^2}{A^2}\frac{(2\ell+1)2^{2\ell}}{L^{2\ell+1}}, \quad 0 \le \ell \le p.\end{aligned} \tag{12}$$

From the coefficient estimates, the expressions for bias and variance of the IF estimate at the center of the window ($t = 0$) are obtained as

$$\begin{aligned}\operatorname{Bias}\left(\hat{f}(0)\right) &= \frac{1}{2\pi}\frac{3\phi^{(p+1)}(0)}{2^p(p+1)!(p+3)}L^p,\\ \operatorname{Var}\left(\hat{f}(0)\right) &= \frac{3\sigma_w^2}{\pi^2A^2L^3}.\end{aligned} \tag{13}$$

It may be noted that these are approximate asymptotic expressions for bias and variance of the ZC-IF estimator.

5. OPTIMUM WINDOW LENGTH SELECTION

Substituting the expressions for bias and variance obtained above, we can write the expression for MSE, $\operatorname{MSE}(\hat{f}(0))$ as follows:

$$\operatorname{MSE}\left(\hat{f}(0)\right) = \left[\frac{1}{2\pi}\frac{3\phi^{(p+1)}(0)}{2^p(p+1)!(p+3)}L^p\right]^2 + \frac{3\sigma_w^2}{A^2\pi^2L^3}. \tag{14}$$

The MSE is a function of the window length L. In Figure 2, we illustrate the variation of bias, variance, and MSE, as a function of window length. Since the bias, variance, and MSE characterize an estimator, the y-axis is commonly labelled

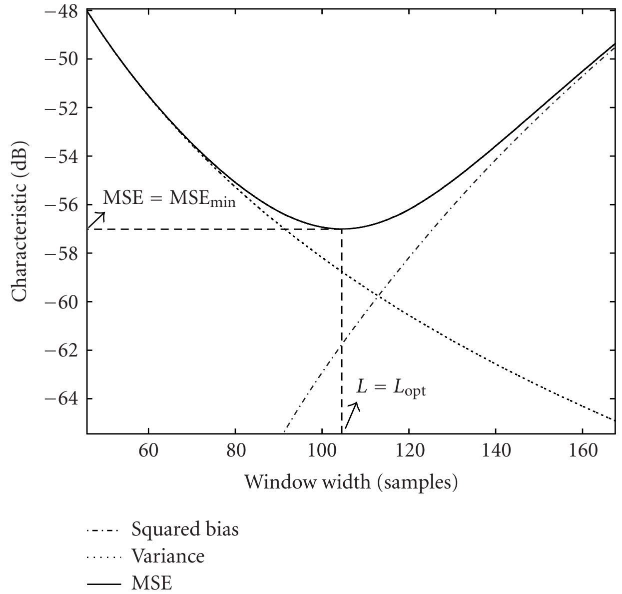

FIGURE 2: Asymptotic squared bias, variance, and mean square error as a function of the window length.

as *characteristic* and plotted in decibel (dB) scale. From the figure, we infer that the MSE has a minimum with respect to window length. The optimal window length, L_{opt} corresponding to MMSE is given as

$$\begin{aligned}L_{\text{opt}} &= \arg\min_L \operatorname{MSE}\\ &= \left[\frac{\sigma_w^2\left[(p+1)!\right]^2 2^{2p}(p+3)^2}{2\pi^2A^2p\left[\phi^{(p+1)}(0)\right]^2}\right]^{1/(2p+3)}.\end{aligned} \tag{15}$$

All the mathematically valid minimizers of the MSE are not practically meaningful. Only the real solution, L_{opt} above, is relevant. The optimum window length is a function of the higher-order derivatives of the IF which are not known a priori, because the IF itself is not known and it has to be estimated. The above expression for the optimal window length is mainly of theoretical interest. The analysis, however, throws light on the issues and tradeoff involved in window length selection for MMSE-ZC-IF estimation. Unlike the expression for bias, the expression for variance does not require any a priori knowledge of the IF, but depends only on the SNR which can be estimated. The expression for variance can be used to devise an adaptive window algorithm to solve the bias-variance tradeoff for MMSE ZC-IF estimation.

5.1. *Bias-variance tradeoff*

The expressions for squared bias and variance can be restated as follows:

$$\begin{aligned}\operatorname{Bias}^2\left(\hat{f}(0)\right) &= \mathcal{B}L^{2p},\\ \operatorname{Var}\left(\hat{f}(0)\right) &= \frac{\mathcal{V}(\text{SNR})}{L^3},\end{aligned} \tag{16}$$

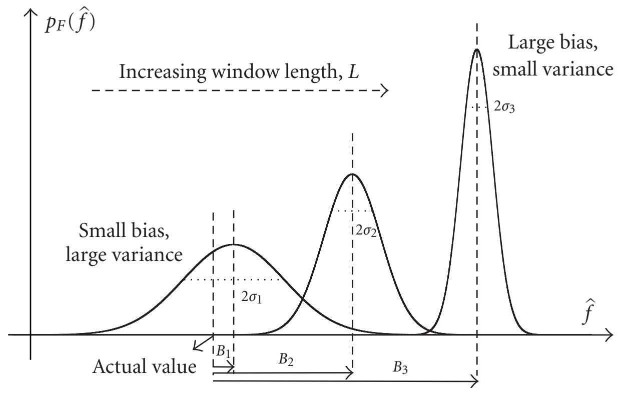

Figure 3: Asymptotic distribution of the ZC-IF estimator for different window lengths.

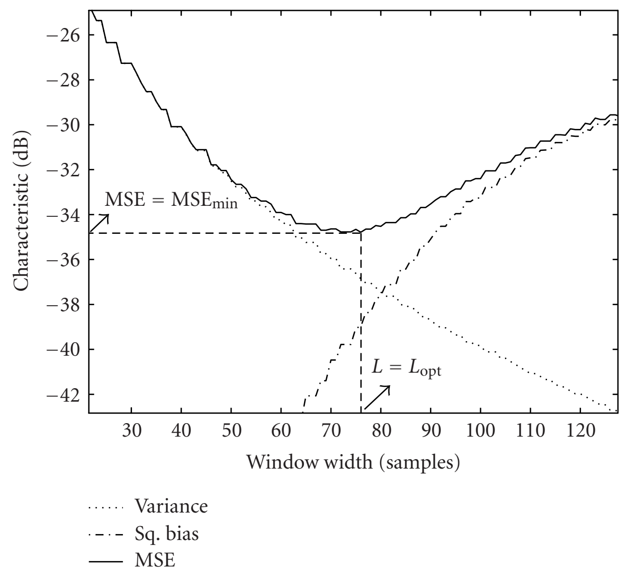

Figure 4: Bias-variance tradeoff in the irregular sampling scenario relevant to ZC-IF estimator.

where

$$\mathcal{B} = \left[\frac{1}{2\pi}\frac{3\phi^{(p+1)}(0)}{2^p(p+1)!(p+3)}\right]^2,$$
$$\mathcal{V}(\text{SNR}) = \frac{3\sigma_w^2}{A^2\pi^2}, \quad \text{SNR} = \frac{A^2}{2\sigma_w^2}. \tag{17}$$

At $L = L_{\text{opt}}$,

$$\text{Bias}^2\left(\hat{f}(0)\right) = \left(\frac{3}{2p}\right)^{2p/(2p+3)} \mathcal{B}^{3/(2p+3)}\mathcal{V}^{2p/(2p+3)},$$
$$\text{Var}\left(\hat{f}(0)\right) = \left(\frac{2p}{3}\right)^{3/(2p+3)} \mathcal{B}^{3/(2p+3)}\mathcal{V}^{2p/(2p+3)}, \tag{18}$$
$$\text{Bias}\left(\hat{f}(0), L_{\text{opt}}\right) = \sqrt{\frac{3}{2p}\text{Var}\left(\hat{f}(0)\right)}.$$

From the expressions above, it is clear that the squared bias is directly proportional to L^{2p} and the variance is inversely proportional to L^3, clearly indicating *bias-variance tradeoff* frequently encountered in devising estimators operating on windowed data [12, 26]. The increased smoothing of the estimate for a long window decreases variance but increases bias; conversely, reduced smoothing with a short window increases variance but bias decreases. The asymptotic distribution of the estimator is shown in Figure 3.

We need to emphasize an important aspect specific to the ZC-IF estimator. Unlike regular sampling, in an irregular sampling scenario (ZC data belongs to this class), the distribution of data is not uniform. In the case of uniform sampling, as the window length is increased, in multiples of the sampling period, the window encompasses more data and hence the associated bias and variance change monotonically. However, in the irregular sampling case, as the window length is increased in multiples of the sampling period, the window may or may not encompass more data, depending on the data distribution. Thus, the associated bias and variance do not vary smoothly. This is illustrated through an example.

A noise sequence, white and Gaussian distributed, 256 samples long, was lowpass filtered (filter's normalized cutoff frequency arbitrarily chosen as 0.05 Hz). The filtered signal was rescaled and adjusted to have amplitude excursions limited to $[0, 0.45]$. This was used as the IF to simulate a constant amplitude, frequency modulated sinusoid. Additive white Gaussian noise was added to achieve an SNR of 25 dB. Since this is a synthesized example, the underlying IF is known and hence bias can be computed directly. Using ZCs to perform a third-order polynomial phase fitting, the IF was estimated at the center of the observation window for different window lengths. The experiment was repeated 100 times and the bias and variance were computed and plotted in Figure 4. The figure clearly illustrates the bias-variance tradeoff for the ZC-IF estimator using noisy signal data.

6. ADAPTIVE WINDOW ZC-IF TECHNIQUE (AZC-IF)

Asymptotically, the IF estimate[3] $\hat{f}_L$ (the subscript L denotes the window length) can be considered as a Gaussian random variable distributed around the actual value, f, with bias, $b(\hat{f}_L)$ and standard deviation, $\sigma(\hat{f}_L)$. Thus, we can write the following relation:

$$\left| f - \hat{f}_L - b(\hat{f}_L) \right| \le \kappa\sigma(\hat{f}_L) \tag{19}$$

for a given SNR. This inequality holds with probability $P(|f - \hat{f}_L - b(\hat{f}_L)| \le \kappa\sigma(\hat{f}_L))$. In terms of the standard normal

[3]We simplify the notation used. Assuming data centering, the time instant of IF estimate is dropped. The IF estimate obtained using window length L is indicated as $\hat{f}_L$. Bias and standard deviation are denoted by $b(\hat{f}_L)$ and $\sigma(\hat{f}_L)$, respectively.

distribution, $\mathcal{N}(0,1)$, this probability is given as $P(\kappa)$ and tends to unity as κ tends to infinity. We can rewrite this inequality as

$$|f - \hat{f}_L| \le |b(\hat{f}_L)| + \kappa\sigma(\hat{f}_L) \tag{20}$$

which holds with probability $P(|f - \hat{f}_L| \le |b(\hat{f}_L)| + \kappa\sigma(\hat{f}_L))$. Now, if $|b(\hat{f}_L)| \le \Delta\kappa\sigma(\hat{f}_L)$, we can rewrite the inequality as

$$|f - \hat{f}_L| \le (\Delta\kappa + \kappa)\sigma(\hat{f}_L) \tag{21}$$

which holds with probability $P(|f - \hat{f}_L| \le (\Delta\kappa + \kappa)\sigma(\hat{f}_L))$. Therefore, we can define a confidence interval for the IF estimate (using window length L) as

$$D = \left[\hat{f}_L - (\Delta\kappa + \kappa)\sigma(\hat{f}_L), \hat{f}_L + (\Delta\kappa + \kappa)\sigma(\hat{f}_L)\right]. \tag{22}$$

We define a set of discrete-window lengths, $\mathcal{H} = \{L_s \mid L_s = a^s L_0,\ s = 0,1,2,\dots,s_{\max};\ a > 1\}$.[4] If $a = 2$, this set is *dyadic* in nature. Likewise, if $a = 3$, it is a *triadic* window set. We choose $a = 2$. At this point, we recall a theorem from [10] using which we can show that, for the present case,

$$\begin{aligned} \Delta\kappa &= \sqrt{\frac{3}{2p}2^{3/2}\frac{2^p - 1}{2^{3/2}+1}}, \\ \kappa &< \sqrt{\frac{3}{2p}2^{1/2}\frac{2^p - 1}{2^{3/2}+1}\left(2^{(3+2p)/2} - 1\right)}. \end{aligned} \tag{23}$$

For a third-order fit, that is, $p = 3$, we have $\Delta\kappa = 3.6569$ and $\kappa < 43.2013$. Together, we have $\kappa + \Delta\kappa < 46.8582$. This is only an upper bound obtained using the approximate asymptotic analysis in Section 4. For simulations reported in this paper, a 5σ confidence interval, that is, $\kappa + \Delta\kappa = 2.5$ was used. For this value of $\kappa + \Delta\kappa$, the coverage probability is nearly 0.99. For a detailed discussion on the choice of $\kappa + \Delta\kappa$, see [27].

We can also define $\mathcal{H}$ as $\mathcal{H} = \{L_s \mid L_s = (s+1)L_0,\ s = 0,1,2,\dots,\tilde{s}_{\max}\}$, that is, the window lengths are in arithmetic progression. The consequence of such a choice is studied in Section 7.2.

6.1. Algorithm

The algorithm for AZC-IF estimation at a point t is summarized as follows.

(1) Initialization. Choose $\mathcal{H} = \{L_s \mid L_s = a^s L_0,\ s = 0,1,2,\dots,s_{\max};\ a > 1\}$, $\kappa + \Delta\kappa = 2.5$. Set $s = 0$. L_0 is chosen as the window length encompassing $p + 1$ farthest ZCs. This ensures that at any stage of the algorithm, there is sufficient data to perform a pth-order fit. $s_{\max}$ is chosen such that $L_{s_{\max}+1}$ just exceeds the observation window length. The IF estimate $\hat{f}_{L_s}$ is obtained using the window length L_s, that is, the ZC data (after data centering) within the window is used to perform a pth-order fit to obtain the IP and the IF. Let $\hat{f}_{L_s}$ be the corresponding AZC-IF estimate.

[4] The choice of $s_{\max}$ and L_0 is discussed in Sections 6.1 and 7.2.

(2) Confidence interval computation. The limits of the confidence interval are computed as follows:

$$\begin{aligned} P_s &= \hat{f}_{L_s} - (\kappa + \Delta\kappa)\sigma(\hat{f}_{L_s}), \\ Q_s &= \hat{f}_{L_s} + (\kappa + \Delta\kappa)\sigma(\hat{f}_{L_s}). \end{aligned} \tag{24}$$

(3) Estimation. Obtain $\hat{f}_{L_{s+1}}$ using the next window length, $L_{s+1} = 2L_s$, from the set $\mathcal{H}$. Compute the confidence interval limits as follows:

$$\begin{aligned} P_{s+1} &= \hat{f}_{L_{s+1}} - (\kappa + \Delta\kappa)\sigma(\hat{f}_{L_{s+1}}), \\ Q_{s+1} &= \hat{f}_{L_{s+1}} + (\kappa + \Delta\kappa)\sigma(\hat{f}_{L_{s+1}}). \end{aligned} \tag{25}$$

(4) Check. Is $D_s \cap D_{s+1} = \emptyset$? ($D_s = [P_s, Q_s]$, $D_{s+1} = [P_{s+1}, Q_{s+1}]$ and $\emptyset$ denotes the empty set). In other words, the following condition is checked:

$$|\hat{f}_{L_{s+1}} - \hat{f}_{L_s}| \le 2(\kappa + \Delta\kappa)\left[\sigma(\hat{f}_{L_s}) + \sigma(\hat{f}_{L_{s+1}})\right]. \tag{26}$$

The smallest value of s for which the condition is satisfied yields the optimum window length, that is, if s^* is the smallest value of s for which the condition is satisfied, then $L_{\text{opt}} = L_{s^*}$; else $s \leftarrow s + 1$ and steps 3 and 4 are repeated.

Since the bias varies as L^{2p}, large values of p imply a fast-varying bias. This results in an MSE that is steep about the optimum. With a discretized search space of window lengths, small changes in the window length about the optimum can cause steep rise in the MSE. Also, large values of p can give rise to numerically unstable set of equations. On the other hand, small values of p, that is, $p = 1, 2$, correspond to a not-so-clearly defined minimum; $p = 3$ was found to be a satisfactory choice and is used in the simulations reported in this paper.

For implementing the algorithm, the computation of variance requires an estimate of the SNR. The SNR estimator suggested in [10, 11] requires oversampling of the signal. Though robust at very low SNRs, in general, it was found to yield poor estimates of the SNR even with considerably large oversampling factors. Therefore, an alternative *method of moments* estimator is proposed for estimating SNR. A detailed study of its properties and improved adaptive TFD-based IF estimation is reported separately. For the sake of completeness, the SNR estimator is given below (the hat is used to denote an estimate):

$$\frac{\hat{A}^2}{2\sigma_w^2} = \frac{3\left[(1/N)\sum_{n=0}^{N-1}|h_y[n]|^2\right]^2 - (1/N)\sum_{n=0}^{N-1}|h_y[n]|^4}{(1/N)\sum_{n=0}^{N-1}|h_y[n]|^4 - \left[(1/N)\sum_{n=0}^{N-1}|h_y[n]|^2\right]^2}, \tag{27}$$

where $h_y[n]$ is the analytic signal [20] of $y[n]$.

7. SIMULATIONS

We present here simulation results evaluating the performance of the AZC-IF technique and also compare it with the fixed window approaches.

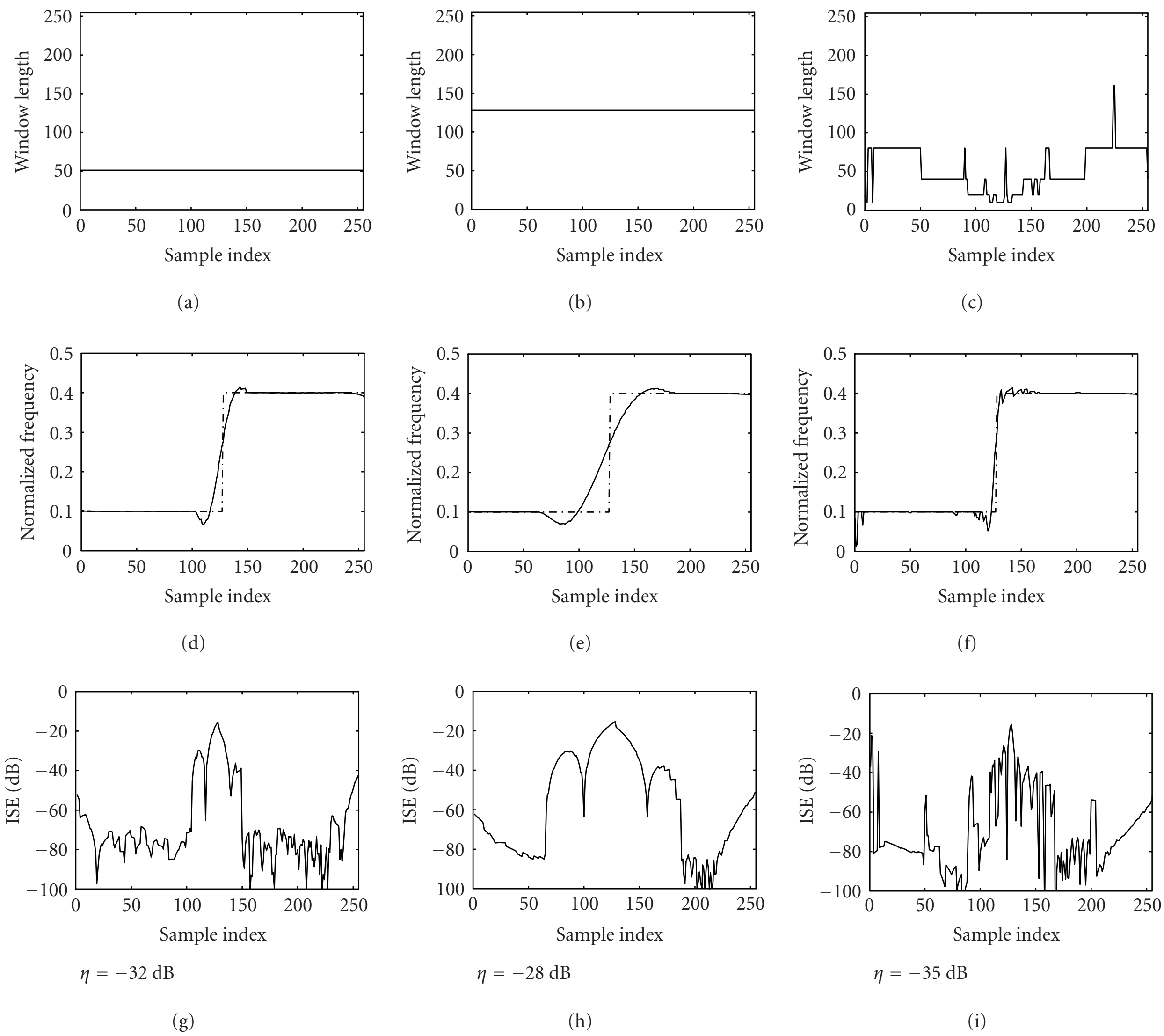

Figure 5: ZC technique using fixed and adaptive windows for step IF estimation (SNR = 25 dB). The columns correspond to medium window (51 samples), long window (129 samples), and adaptive window, respectively. In (a), (b), and (c), the corresponding window length is shown as a function of the sample index. In (d), (e), and (f), the actual IF is shown in dashed-dotted style and the estimated IF is shown in solid line style. In (g), (h), and (i), ISE stands for instantaneous squared error. η is average error.

7.1. Fixed window versus adaptive window ZC-IF estimator

To illustrate the adaptation of window length, we consider the following IF laws.

(1) Step IF.

$$f[n] = \begin{cases} 0.1 & \text{for } 0 \le n \le 127, \\ 0.4 & \text{for } 128 \le n \le 255. \end{cases} \tag{28}$$

(2) "Sum of sinusoids" IF.

$$\begin{aligned} f[n] &= 0.1092\sin(0.128n) + 0.0595\sin(0.1n) \\ &\quad + 0.2338 \quad \text{for } 0 \le n \le 255. \end{aligned} \tag{29}$$

The coefficients and frequencies of the sinusoids were chosen arbitrarily. The coefficients were rescaled and a suitable constant added to bring the IF within the normalized frequency range $[0, 0.5]$.

(3) Triangular IF.

$$f[n] = \begin{cases} 0.2 + \dfrac{0.2n}{127} & \text{for } 0 \le n \le 127, \\ 0.4 - \dfrac{0.2(n-127)}{128} & \text{for } 127 \le n \le 255. \end{cases} \tag{30}$$

For each of the IF above, the following experiments were conducted:

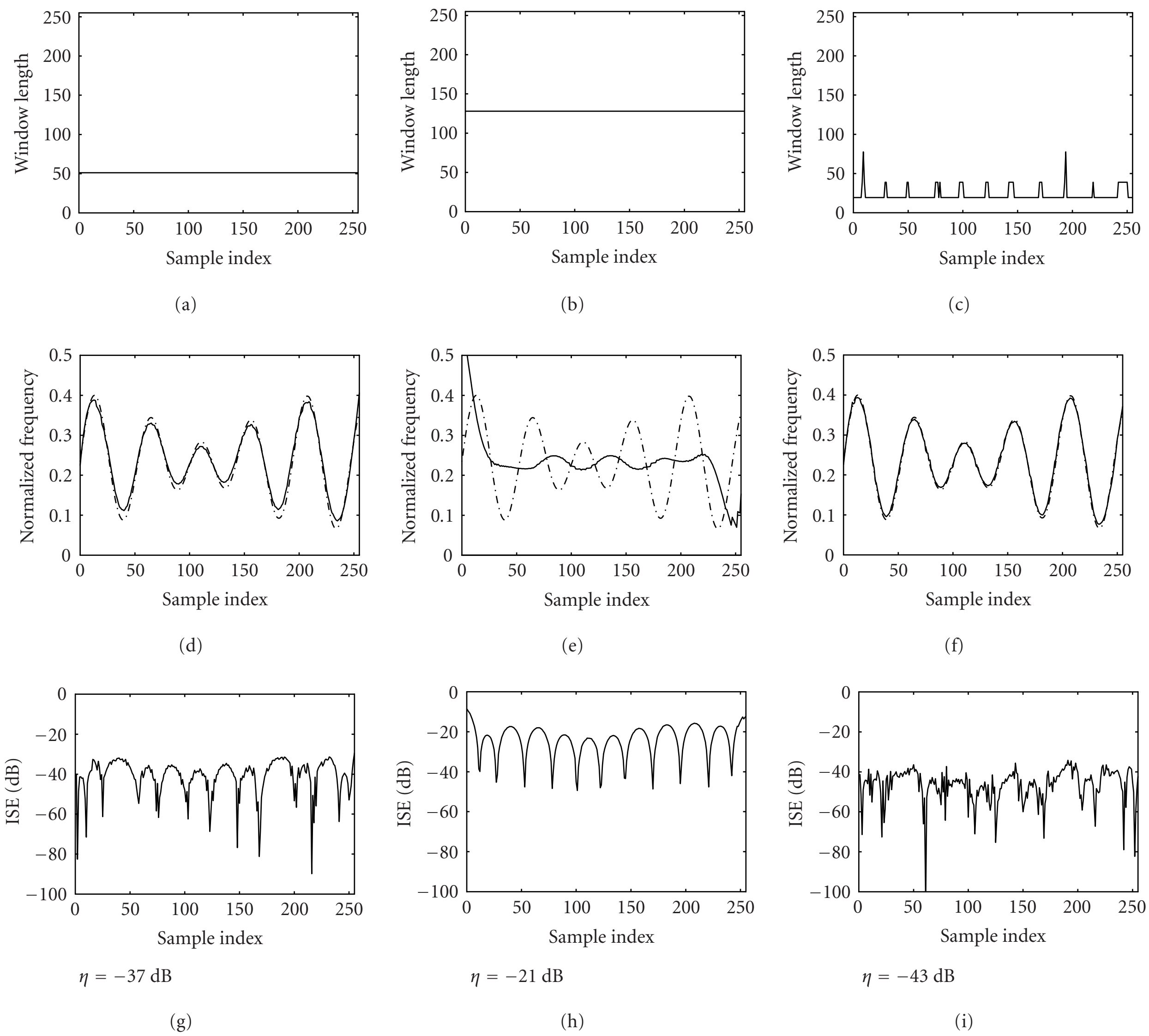

Figure 6: ZC technique using fixed and adaptive windows for "sum of sinusoids" IF estimation (SNR = 25 dB). The columns correspond to medium window (51 samples), long window (129 samples), and adaptive window, respectively. In (a), (b), and (c), the corresponding window length is shown as a function of the sample index. In (d), (e), and (f), the actual IF is shown in dashed-dotted style and the estimated IF is shown in solid line style. In (g), (h), and (i), ISE stands for instantaneous squared error. η is average error.

(1) ZC-IF estimation using a fixed medium window (51-samples long),
(2) ZC-IF estimation using a fixed long window (129 samples long),
(3) adaptive window ZC-IF estimation.

$p = 3$ was used in all the simulations. The window lengths of 51 and 129 samples are arbitrary. The ZC-IF estimates were obtained for each IF. The following IF error measures are computed:

(1) *instantaneous squared error*, $\text{ISE}[n] = (f[n] - \hat{f}[n])^2$,
(2) *average error*, $\eta = (1/(N-20))\sum_{n=11}^{N-10}(f[n] - \hat{f}[n])^2$,

10 samples[5] at the extremes of the signal window are excluded to eliminate errors due to boundary effects, because for most methods, the errors at the edges are large giving rise to unreasonable estimates. The results are shown in Figures 5, 6, 7. From these figures, the following observations can be made.

(1) For relatively stationary regions of the IF, the adaptive algorithm chooses larger window lengths thereby reducing variance via increased data smoothing.

[5]The number 10 was arrived at by comparing AZC-IF, adaptive spectrogram and WVD peak-based IF estimation errors (reported in Section 7.3) for different window lengths.

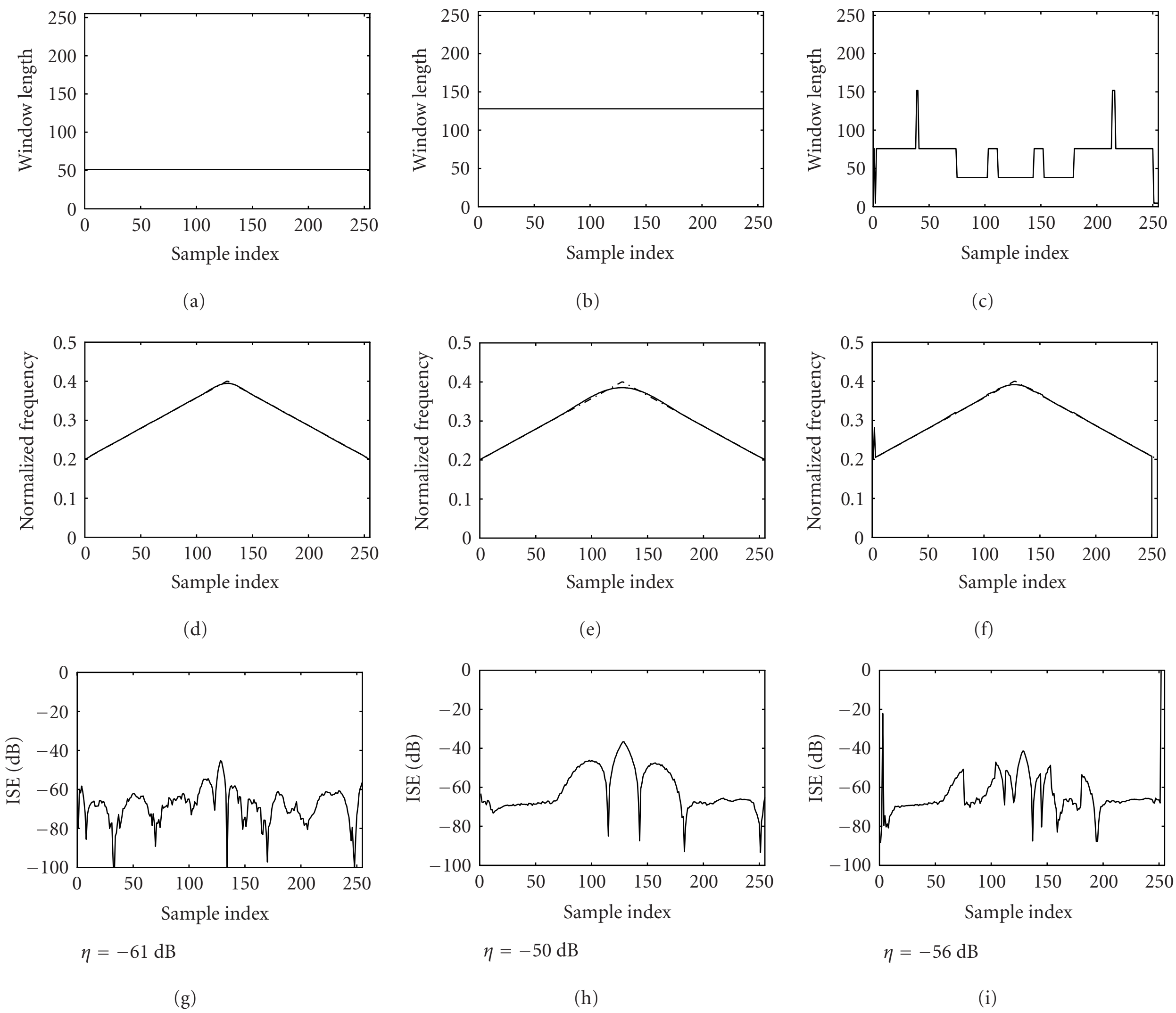

Figure 7: ZC technique using fixed and adaptive windows for triangular IF estimation (SNR = 25 dB). The columns correspond to medium window (51 samples), long window (129 samples), and adaptive window, respectively. In (a), (b), and (c), the corresponding window length is shown as a function of the sample index. In (d), (e), and (f), the actual IF is shown in dashed-dotted style and the estimated IF is shown in solid line style. In (g), (h), and (i), ISE stands for instantaneous squared error. η is average error.

(2) In the vicinity of a fast change in IF (like the discontinuity in the case of step IF), the algorithm chooses shorter window length thereby reducing bias and hence capturing "*events in time*." This improves *time resolution* but at the expense of large variance. At the discontinuity, the local polynomial approximation does not hold; the corresponding window length chosen by the algorithm is large.

(3) Fixed window ZC-IF estimate obtained with a longer window length is smeared/oversmoothed than that obtained with a shorter window length.

(4) The average error, η, is the best with adaptive window length estimator for the case of the step IF and also "sum of sinusoids" IF. However, with triangular IF, we find that the AZC-IF estimate obtained with the adaptive window length has a few dB higher error compared to that obtained with a medium window length. Such a behaviour, which appears to be counterintuitive at first, is possible with any kind of IF, as simulations later will show. However, it must be noted that this is because of the choice of the set of window lengths. The set of window lengths chosen are dyadic in nature and hence the optimum MSE search is very coarse. It is possible that, in such a case, the adaptive algorithm determines a window length quite away from the optimum, which yields poorer performance than the medium window length. This may be overcome by finely searching the space of window lengths. This is discussed in the following section.

7.2. Coarse search versus fine search

We study the effect of discretizing the search space of windows on the performance of the algorithm. We consider the following window lengths:

(1) *medium window length*: $L = 51$ samples,

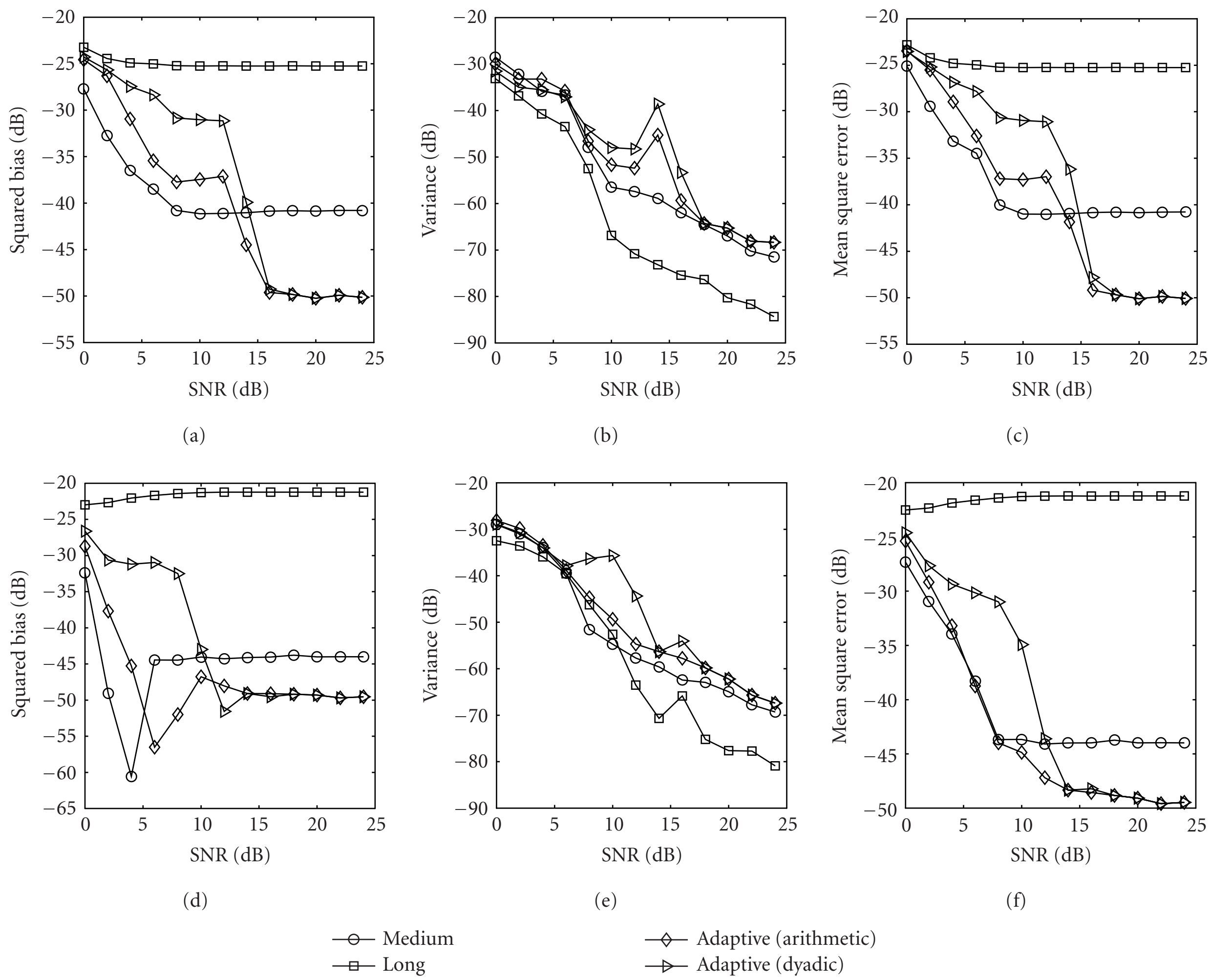

Figure 8: Performance of fixed and adaptive ZC techniques as a function of SNR (dB) for different window choices, for "sum of sinusoids" IF estimation. Top row corresponds to $n = 128$ and bottom row corresponds to $n = 200$.

(2) *long window*: $L = 129$ samples,

(3) *arithmetic set of window lengths*: window lengths in arithmetic progression, that is, $L_s = (s+1)L_0$, $s = 0, 1, 2, \ldots, \tilde{s}_{\max}$,

(4) *dyadic set of window lengths*: window lengths given by $L_s = 2^s L_0$, $s = 0, 1, 2, \ldots, s_{\max}$.

In the arithmetic and dyadic cases, L_0 is the initial window length encompassing $(p+1)$ farthest ZCs; the values of $s_{\max}$ and $\tilde{s}_{\max}$ are chosen such that $L_{s_{\max}+1}$ and $L_{\tilde{s}_{\max}+1}$, respectively, just exceed the given observation window length.

For illustration, we chose the "sum of sinusoids" IF. The ZC-IF estimates corresponding to each of the above windows are obtained. The SNR was varied from 0 dB to 24 dB in steps of 2 dB. For each SNR, 100 Monte-Carlo realizations were run to obtain the statistics of the ZC-IF estimator. The squared bias, variance, and MSE corresponding to the sample instants $n = 128$ and $n = 200$ are plotted in Figure 8. From the figures, we observe that ZC-IF estimates obtained with medium window size (whose choice cannot be made a priori anyway) show "slightly" better performance than the adaptive window algorithms at low SNR. This is because the adaptation algorithm uses asymptotic expressions for bias and variance which are derived under a high SNR assumption and hence less appropriate at low SNR. At high SNR, however, the adaptive estimates are consistently better than fixed window estimates. Also of interest is the comparison between the arithmetic and dyadic window choices for the adaptive algorithm. Consistently, the AZC-IF estimates obtained with arithmetic window set (corresponding to fine search) outperform those obtained with the dyadic set (corresponding to coarse search) for low to moderate SNR. Sometimes, the improvement in performance can be as high as 12 dB (bottom row in Figure 8, corresponding to SNR = 8 dB). However, for high SNR, there seems to be no distinct advantage in doing a fine search. One can protract this observation to state that if the SNR estimate indicates a high value (roughly beyond 18 dB or so), a considerable computational saving can be achieved by using the dyadic window set.

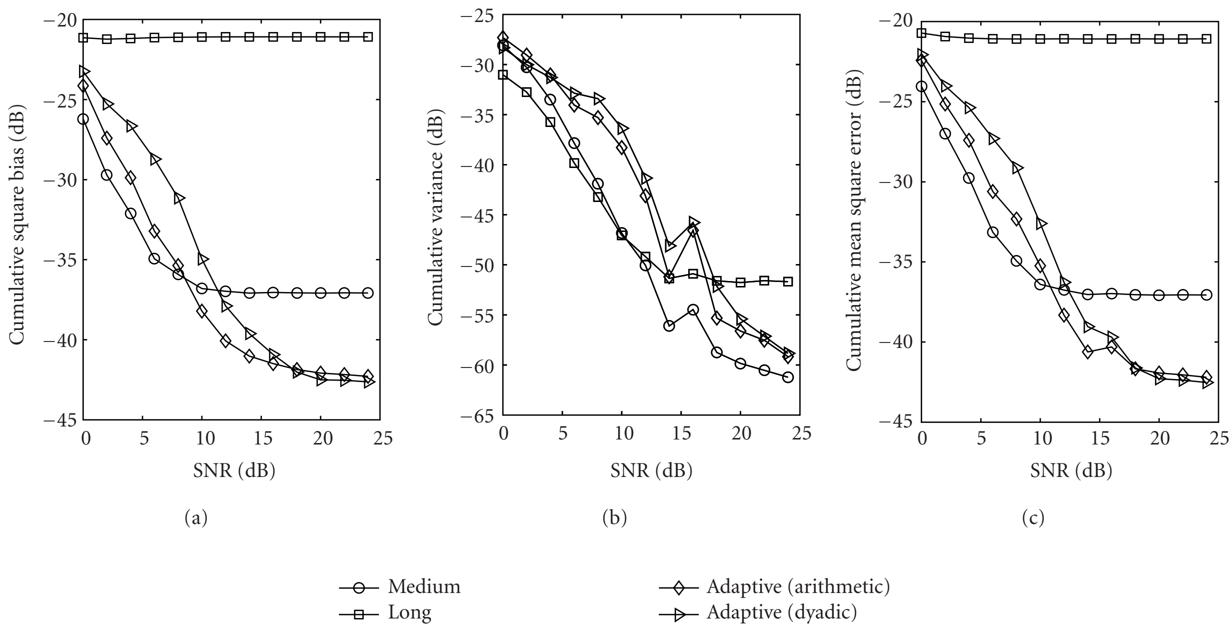

FIGURE 9: Performance of fixed and adaptive ZC techniques as a function of SNR (dB) for different window choices, for "sum of sinusoids" IF estimation.

Since we are interested in tracking IF, all the pointwise errors can be combined to give cumulative error measures which we introduce as

$$\begin{aligned} \text{cumulative squared bias} &= \left(\frac{1}{N-20}\right)\sum_{n=11}^{N-10}\left(\text{Bias}\left(\hat{f}[n]\right)\right)^2, \\ \text{cumulative variance} &= \left(\frac{1}{N-20}\right)\sum_{n=11}^{N-10}\text{Var}\left(\hat{f}[n]\right), \\ \text{cumulative MSE} &= \text{cumulative squared bias} \\ &\quad + \text{cumulative variance}. \end{aligned} \tag{31}$$

These are time averages of instantaneous squared bias, variance, and MSE, respectively. Their importance is due to the fact that they give a single error measure taking into account pointwise errors. These can be useful indicators of the performance of IF estimators for a given data. These are shown in Figure 9. The significant improvement in performance offered by adaptive estimates compared to fixed window ZC-IF estimates stresses the significance of appropriate window length choice in ZC-IF estimation.

7.3. Comparison with other techniques

We compare the performance of the AZC-IF algorithm with adaptive spectrogram (ASPEC) and Wigner-Ville distribution (AWVD)-based IF estimates. For a discussion on adaptive spectrogram and adaptive WVD-based IF estimation, refer [10, 11, 24].

A dyadic window set was used for the three methods. For ASPEC and AWVD, the initial window size was chosen to be 2 samples, whereas for AZC, it is chosen as the window length encompassing $(p+1)$ farthest consecutive ZCs. This varies from realization to realization and is definitely greater than twice the sampling period and hence relatively, the search for AZC is more coarse than ASPEC and AWVD. The spectrogram and WVD are computed by zero-padding to compute a $4N$-point discrete Fourier transform (DFT) (N is the length of the entire data available, $N = 256$ samples in the present examples). After locating a peak in the $4N$-point DFT, the actual position of the peak is further refined by fitting a quadratic polynomial to three samples about the DFT peak (two samples on either side of the peak and the DFT peak sample itself). This gives a peak location closer to the actual discrete-time Fourier transform (DTFT) peak than the DFT peak. This was done to minimize the contribution of the error in peak-picking to the error in the IF estimate. For the AZC algorithm, to locate the ZCs, the signal is oversampled four times (equivalently, four stages of the bandlimited interpolation plus bisection scheme) and linear interpolation between adjacent samples of opposite sign was performed to estimate the ZC. The cumulative MSE is shown in Figure 10. We infer that AZC technique for triangular IF estimation offers as much as 10 dB less error at low SNRs (and hence more robust) than ASPEC and AWVD-based estimates. For high SNRs, however, the performance of the AZC technique is about 5–6 dB poorer than ASPEC and AWVD. For the "sum of sinusoids" IF, the AZC technique performance is similar to the other two techniques at all SNRs.

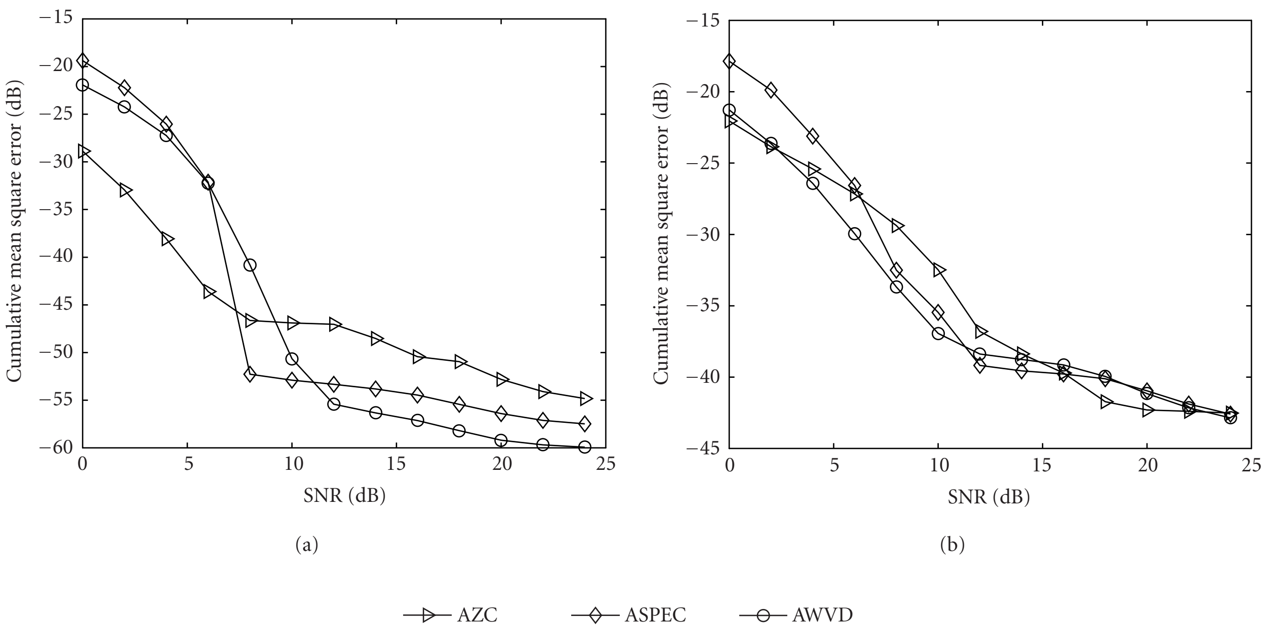

Figure 10: Comparison of AZC, ASPEC, and AWVD techniques for IF estimation (coarse search). (a) Triangular IF. (b) “Sum of sinusoids” IF.

The experiment is repeated with the arithmetic window set. For ASPEC and AWVD, the window set consists of consecutive multiples of 2, whereas for AZC algorithm it consists of consecutive multiples of the initial window length (which is chosen as the window length encompassing $(p+1)$ farthest ZCs). Therefore, relatively, the search is more coarse for AZC algorithm than ASPEC and AWVD algorithms. The results are shown in Figure 11. Here, AZC consistently outperforms ASPEC and AWVD algorithms. This is because, the squared bias of the ZC-IF estimator varies as L^{2p} which is quite faster than the corresponding variation for spectrogram and WVD [10, 11, 24]. This implies that a coarse search in the window space can severely affect the performance of AZC more than ASPEC and AWVD. Simulations (Figures 10 and 11) strongly support this argument and emphasize the need for appropriate discretization of the window search space for robust IF estimation using the AZC algorithm.

8. CONCLUSION

We have achieved robust estimation of arbitrary IF using real ZCs of the frequency modulated signal in an adaptive window framework. This approach combines in an interesting and useful manner nonlinear measurement (ZCs), nonuniform sampling, and adaptive window techniques, resulting in superior IF estimation. Comparative simulations show that the adaptive window ZC technique can provide as much as 5–10 dB performance advantage over adaptive spectrogram and Wigner-Ville distribution-based IF estimation techniques. Extension of the new technique for estimating arbitrary instantaneous frequencies of multicomponent signals (similar to [28]) is being explored.

APPENDIX

In this appendix, we give the detailed derivation for bias and variance of the ZC-IF estimator reported in Section 4. With a change of variable from t_n to $t - t_n$, each of the quantities in (9) can be evaluated as follows ($0 \le \ell \le p$):

$$\begin{aligned}
\left.\frac{\partial \mathcal{C}}{\partial a_\ell}\right|_0 &= -2 \sum_{n=i,\, t_n \in \mathcal{I}_{0,L}}^{j} \left[\phi(t-t_n) - \sum_{k=0}^{p} a_k (t-t_n)^k\right] \\
&\qquad \times (t-t_n)^\ell h(t_n) = 0, \\
\left.\frac{\partial^2 \mathcal{C}}{\partial a_\ell^2}\right|_0 \Delta a_\ell &= 2 \sum_{n=i,\, t_n \in \mathcal{I}_{0,L}}^{j} (t-t_n)^{2\ell} h(t_n) \Delta a_\ell, \\
\left.\frac{\partial \mathcal{C}}{\partial a_\ell}\right|_0 \delta_{\Delta\phi} &= -2 \sum_{n=i,\, t_n \in \mathcal{I}_{0,L}}^{j} \Delta\phi(t,t_n)(t-t_n)^\ell h(t_n), \\
\left.\frac{\partial \mathcal{C}}{\partial a_\ell}\right|_0 \delta_{\tilde{w}} &= -2 \sum_{n=i,\, t_n \in \mathcal{I}_{0,L}}^{j} \tilde{w}(t-t_n)(t-t_n)^\ell h(t_n),
\end{aligned} \tag{A.1}$$

where $\Delta\phi(t,t_n) = \sum_{r=p+1}^{\infty}((-t_n)^r/r!)\phi^r(t)$, $r!$ denoting the factorial of r. Retaining only the significant term in the infinite summation, we can write $\Delta\phi(t,t_n) \approx ((-t_n)^{p+1}/(p+1)!)\phi^{(p+1)}(t)$.

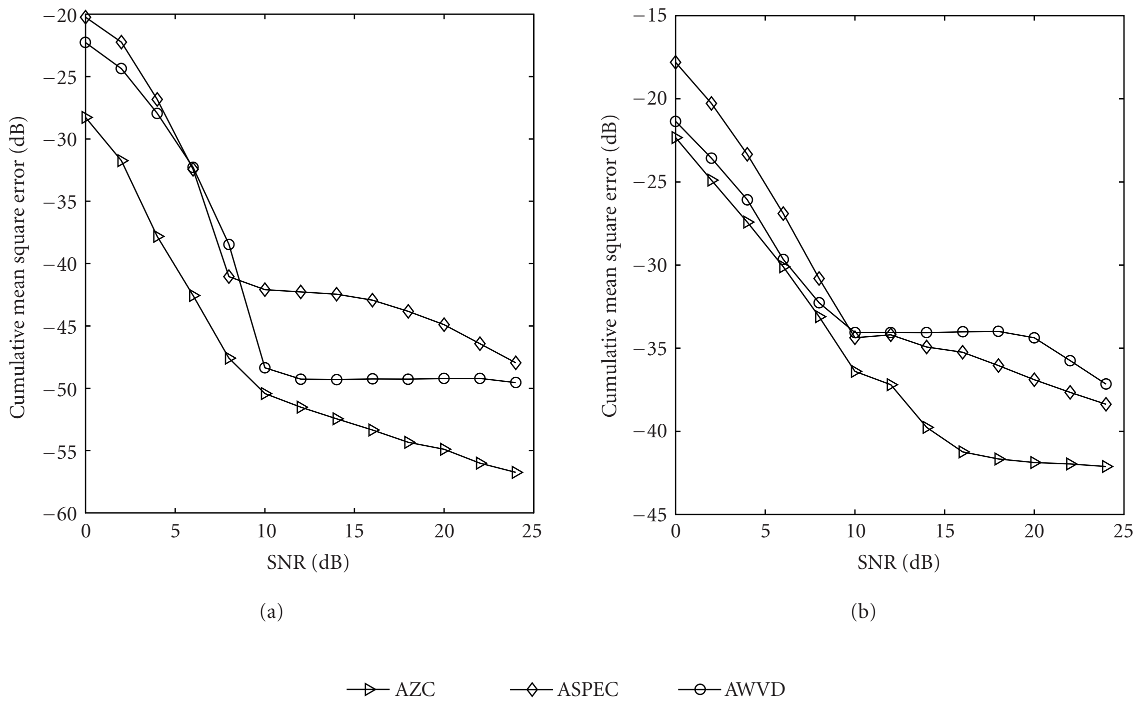

Figure 11: Comparison of AZC, ASPEC, and AWVD techniques for IF estimation (fine search). (a) Triangular IF. (b) "Sum of sinusoids" IF.

Equating $\partial\mathcal{C}/\partial a_\ell$ to zero and solving for Δa_ℓ, $0 \le \ell \le p$, we get

$$\Delta a_\ell = \frac{(-1)^{p+1}\phi^{(p+1)}(t)}{(p+1)!} \times \left[\frac{\sum_{n=i,\,t_n\in\mathcal{I}_{0,L}}^{j} t_n^{p+1}(t-t_n)^\ell h(t_n)}{\sum_{n=i,\,t_n\in\mathcal{I}_{0,L}}^{j}(t-t_n)^{2\ell}h(t_n)} + \frac{\sum_{n=i,\,t_n\in\mathcal{I}_{0,L}}^{j}\tilde{w}(t-t_n)(t-t_n)^\ell h(t_n)}{\sum_{n=i,\,t_n\in\mathcal{I}_{0,L}}^{j}(t-t_n)^{2\ell}h(t_n)}\right]. \tag{A.2}$$

The bias of the estimate, $a_\ell - \mathcal{E}\{\hat{a}_\ell\}$, or equivalently, $\mathcal{E}\{\Delta a_\ell\}$, $0 \le \ell \le p$, is given by

$$\mathcal{E}\{\Delta a_\ell\} = \frac{(-1)^{p+1}\phi^{(p+1)}(t)}{(p+1)!}\left[\frac{\sum_{n=i,\,t_n\in\mathcal{I}_{0,L}}^{j} t_n^{p+1}(t-t_n)^\ell h(t_n)}{\sum_{n=i,\,t_n\in\mathcal{I}_{0,L}}^{j}(t-t_n)^{2\ell}h(t_n)}\right]. \tag{A.3}$$

The covariance of the errors Δa_ℓ and Δa_k, denoted by $\mathrm{Cov}(\Delta a_\ell, \Delta a_k)$, and defined as $\mathrm{Cov}(\Delta a_\ell, \Delta a_k) = \mathcal{E}\{(\Delta a_\ell - \mathcal{E}\{\Delta a_\ell\})(\Delta a_k - \mathcal{E}\{\Delta a_k\})\}$, $0 \le \ell, k \le p$ is given by

$$\begin{aligned}\mathrm{Cov}(\Delta a_\ell, \Delta a_k) &= \frac{\mathcal{E}\left\{\sum_{n=i,\,t_n\in\mathcal{I}_{0,L}}^{j}\tilde{w}(t-t_n)(t-t_n)^\ell h(t_n)\sum_{m=i,\,t_m\in\mathcal{I}_{0,L}}^{j}\tilde{w}(t-t_m)(t-t_m)^k h(t_m)\right\}}{\sum_{n=i,\,t_n\in\mathcal{I}_{0,L}}^{j}(t-t_n)^{2\ell}h(t_n)\sum_{m=i,\,t_m\in\mathcal{I}_{0,L}}^{j}(t-t_m)^{2k}h(t_m)} \\ &= \frac{\sum_{n=i,\,t_n\in\mathcal{I}_{0,L}}^{j}\sum_{m=i,\,t_m\in\mathcal{I}_{0,L}}^{j}\mathcal{E}\{\tilde{w}(t-t_n)\tilde{w}(t-t_m)\}(t-t_n)^\ell(t-t_m)^k h(t_n)h(t_m)}{\sum_{n=i,\,t_n\in\mathcal{I}_{0,L}}^{j}\sum_{m=i,\,t_m\in\mathcal{I}_{0,L}}^{j}(t-t_n)^{2\ell}(t-t_m)^{2k}h(t_n)h(t_m)}.\end{aligned} \tag{A.4}$$

Since the power spectral density of noise, $w(t)$, is known, it is easy to compute $\mathcal{E}\{\tilde{w}(t-t_n)\tilde{w}(t-t_m)\}$, because $r_{\tilde{w}\tilde{w}}(\tau) = (\sigma_w^2/A^2)(\sin(B\pi\tau)/B\pi\tau)$, and $\mathcal{E}\{\tilde{w}(t-t_n)\tilde{w}(t-t_m)\} = r_{\tilde{w}\tilde{w}}(t_n - t_m)$. However, this leads to very complicated expressions for the variance and covariance of the coefficient estimation errors. For the purposes of analysis and ease of computation, one can arrive at simplifed expressions by assuming that $\mathcal{E}\{\tilde{w}(t-t_n)\tilde{w}(t-t_m)\} \approx (\sigma_w^2/A^2)\delta_{t_n,t_m}$. This is equivalent to assuming that the autocorrelation function of noise, $w(t)$, is highly localized at zero lag and is negligible elsewhere (nearly white noise). Under this assumption, one can write the following approximate expressions (for $0 \le \ell$, $k \le p$):

$$\begin{aligned}
&\operatorname{Cov}(\Delta a_\ell, \Delta a_k) \\
&= \frac{\sigma_w^2}{A^2}\left\{\frac{\sum_{n=i,\, t_n\in\mathcal{I}_{0,L}}^{j}(t-t_n)^{\ell+k}h^2(t_n)}{\sum_{n=i,\, t_n\in\mathcal{I}_{0,L}}^{j}(t-t_n)^{2\ell}h(t_n)\sum_{m=i,\, t_m\in\mathcal{I}_{0,L}}^{j}(t-t_m)^{2k}h(t_m)}\right\}, \\
&\operatorname{Var}(\Delta a_\ell) \\
&= \frac{\sigma_w^2}{A^2}\left\{\frac{\sum_{n=i,\, t_n\in\mathcal{I}_{0,L}}^{j}(t-t_n)^{2\ell}h^2(t_n)}{\left[\sum_{n=i,\, t_n\in\mathcal{I}_{0,L}}^{j}(t-t_n)^{2\ell}h(t_n)\right]^2}\right\}.
\end{aligned} \tag{A.5}$$

Asymptotically, as data size within the window tends to infinity, we obtain

$$\begin{aligned}
\operatorname{Bias}(\Delta a_\ell) &\longrightarrow \frac{(-1)^{p+1}\phi^{(p+1)}(t)}{(p+1)!}\left[\frac{\int_{-L/2}^{+L/2}s^{p+1}(t-s)^\ell ds}{\int_{-L/2}^{+L/2}(t-s)^{2\ell}ds}\right], \\
\operatorname{Cov}(\Delta a_\ell, \Delta a_k) &\longrightarrow \frac{\sigma_w^2}{A^2}\left[\frac{\int_{-L/2}^{+L/2}(t-s)^{\ell+k}ds}{\int_{-L/2}^{+L/2}(t-s)^{2\ell}ds\int_{-L/2}^{+L/2}(t-s)^{2k}ds}\right], \\
\operatorname{Var}(\Delta a_\ell) &\longrightarrow \frac{\sigma_w^2}{A^2}\left[\frac{1}{\int_{-L/2}^{+L/2}(t-s)^{2\ell}ds}\right].
\end{aligned} \tag{A.6}$$

The IF at time t is estimated as

$$\hat{f}(t) = \frac{1}{2\pi}\sum_{\ell=1}^{p}\ell\hat{a}_\ell t^{\ell-1}. \tag{A.7}$$

The expressions for bias and variance of the IF can be obtained as

$$\begin{aligned}
\operatorname{Bias}(\hat{f}(t)) &= \frac{1}{2\pi}\sum_{\ell=1}^{p}\operatorname{Bias}(\Delta a_\ell)\ell t^{\ell-1}, \\
\operatorname{Var}(\hat{f}(t)) &= \frac{1}{4\pi^2}\sum_{\ell=1}^{p}\sum_{m=1}^{p}\ell m t^{\ell+m-2}\operatorname{Cov}(\Delta a_\ell, \Delta a_m).
\end{aligned} \tag{A.8}$$

ACKNOWLEDGMENTS

We thank the Speech and Audio Group at Department of Electrical Communication Engineering, Indian Institute of Science, for fruitful discussions in the course of this work. Special thanks to Kiran, Krishna, and Srini for help and support during this work. Thanks are due to the reviewers for constructive criticism which greatly helped improve the quality of the manuscript.

REFERENCES

[1] L. Cohen, *Time-Frequency Analysis*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1995.

[2] T. W. Parsons, *Voice and Speech Processing*, McGraw Hill, New York, NY, USA, 1986.

[3] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1978.

[4] J. G. Proakis, *Digital Communication*, McGraw Hill, New York, NY, USA, 2000.

[5] B. Picinbono, "Some remarks on instantaneous amplitude and frequency of signals," in *Proc. IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, pp. 293–300, Pittsburgh, Pa, USA, October 1998.

[6] T. F. Quatieri, T. E. Hanna, and G. C. O'Leary, "AM-FM separation using auditory-motivated filters," *IEEE Trans. Speech, and Audio Processing*, vol. 5, no. 5, pp. 465–480, 1997.

[7] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "Energy separation in signal modulations with application to speech analysis," *IEEE Trans. Signal Processing*, vol. 41, no. 10, pp. 3024–3051, 1993.

[8] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal. Part 2: Algorithms and applications," *Proceedings of the IEEE*, vol. 80, no. 4, pp. 540–568, 1992.

[9] S. Chandra Sekhar and T. V. Sreenivas, "Auditory motivated level-crossing approach to instantaneous frequency estimation," to appear in *IEEE Trans. Signal Processing*.

[10] L. Stankovic and V. Katkovnik, "Algorithm for the instantaneous frequency estimation using time-frequency distributions with adaptive window width," *IEEE Signal Processing Letters*, vol. 5, no. 9, pp. 224–227, 1998.

[11] V. Katkovnik and L. Stankovic, "Instantaneous frequency estimation using the Wigner distribution with varying and data-driven window length," *IEEE Trans. Signal Processing*, vol. 46, no. 9, pp. 2315–2325, 1998.

[12] L. Stankovic and V. Katkovnik, "Instantaneous frequency estimation using higher order L-Wigner distributions with data-driven order and window length," *IEEE Transactions on Information Theory*, vol. 46, no. 1, pp. 302–311, 2000.

[13] H. B. Voelcker, "Toward a unified theory of modulation, Part I, phase envelope relationships," *Proceedings of the IEEE*, vol. 63, no. 3, pp. 340–353, 1966.

[14] R. Kumaresan and A. Rao, "Model-based approach to envelope and positive instantaneous frequency estimation of signals with speech applications," *J. Acoust. Soc. Amer.*, vol. 105, no. 3, pp. 1912–1924, 1999.

[15] M. A. Poletti, "The homomorphic analytic signal," *IEEE Trans. Signal Processing*, vol. 45, no. 8, pp. 1943–1953, 1997.

[16] B. Kedem, *Time Series Analysis by Higher Order Crossings*, IEEE Press, Piscataway, NJ, USA, 1994.

[17] O. Ghitza, "Auditory nerve representations as a front-end for speech recognition in a noisy environment," *Computer Speech and Language*, vol. 1, no. 2, pp. 109–131, 1986.

[18] T. V. Sreenivas and R. J. Niederjohn, "Zero-crossing based spectral analysis and SVD spectral analysis for formant frequency estimation in noise," *IEEE Trans. Signal Processing*, vol. 40, no. 2, pp. 282–293, 1992.

[19] S. Chandra Sekhar and T. V. Sreenivas, "Polynomial instantaneous frequency estimation using level-crossing information," in *Proc. IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing*, Grado, Italy, June 2003.

[20] A. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, USA, 3rd edition, 2000.

[21] E. V. Krishnamurthy and S. K. Sen, *Numerical Algorithms: Computations in Science and Engineering*, Affiliated East-West Press, New Delhi, India, 2001.

[22] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.

[23] P. M. Djuric and S. M. Kay, "Parameter estimation of chirp signals," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 38, no. 12, pp. 2118–2126, 1990.

[24] S. Chandra Sekhar and T. V. Sreenivas, "Adaptive spectrogram vs. adaptive pseudo-Wigner-Ville distribution for instantaneous frequency estimation," *Signal Processing*, vol. 83, no. 7, pp. 1529–1543, 2003.

[25] B. Ristic and B. Boashash, "Comments on "The Cramer-Rao lower bounds for signals with constant amplitude and polynomial phase"," *IEEE Trans. Signal Processing*, vol. 46, no. 6, pp. 1708–1709, 1998.

[26] S. M. Kay, *Modern Spectral Estimation-Theory and Application*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1988.

[27] B. Boashash, Ed., *Time Frequency Signal Processing—A Comprehensive Reference*, Elsevier, Oxford, UK, 2003.

[28] Z. M. Hussain and B. Boashash, "Adaptive instantaneous frequency estimation of multicomponent FM signals using quadratic time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 50, no. 8, pp. 1866–1876, 2002.

S. Chandra Sekhar was born in Gollapalem, Andhra Pradesh, India on August 5, 1976. He obtained B.E. degree in electronics and communication engineering with first rank from Osmania University in 1999. Apart from several distinctions at the Andhra Pradesh State level, he also won the University's prestigious Prof. K. K. Nair commemoration gold medal in 1999 for the outstanding undergraduate student. In 1999, he joined Indian Institute of Science, Bangalore, India, where he is presently pursuing the Ph.D. degree. He has been a recipient of IBM (India Research Laboratory, New Delhi) Research Fellowship from 2001 to 2004. His main research interests are nonstationary signal processing methods with applications to speech/audio, time-frequency signal processing, auditory modeling, estimation and detection theory, sampling theory, and zero-crossing techniques. He is a student member of IEEE.

T. V. Sreenivas graduated from Bangalore University in 1973, obtained M.E. from Indian Institute of Science(IISc), Bangalore, in 1975 and Ph. D. degree from Indian Institute of Technology, Bombay, India, in 1981, working as Research Scholar at Tata Institute of Fundamental Research, Bombay. During 1982–1985, he worked with LRDE, Bangalore, in the area of low bitrate speech coding. During 1986–1987, he worked with Norwegian Institute of Technology, Trondheim, Norway, developing new techniques for speech coding and speech recognition. During 1988–1989, he was Visiting Assistant Professor at Marquette University, Milwaukee, USA, teaching and researching in speech enhancement and spectral estimation. Since 1990, he has joined the faculty of IISc, Bangalore, where he is currently Associate Professor. At IISc, he leads the activity of Speech and Audio Group. His research is focussed on auditory spectral estimation, speech/audio modeling and novel algorithms for speech/audio compression, recognition and enhancement. He is also a Faculty Entrepreneur and has jointly founded "Esqube Communication Solutions Pvt. Ltd.," a startup company in Bangalore. He has been a Visiting Faculty at Fraunhofer Institute for Integrated Circuits, Erlangen, Germany and Griffith University, Australia. He is a Senior Member of IEEE and currently Chairman of IEEE Signal Processing Society, Bangalore Chapter.