Auditory Motivated Level-Crossing Approach to Instantaneous Frequency Estimation

S. Chandra Sekhar, *Student Member, IEEE,* and Thippur V. Sreenivas, *Senior Member, IEEE*

Abstract—We address the problem of estimating the instantaneous frequency (IF) of a phase signal using its level-crossing (LC) information based on front-end auditory processing motivation. We show that the problem of IF estimation using LC information can be cast in the framework of estimation from irregularly sampled data. The formulation has the generality of estimating different types of IF without the need for a quasistationary assumption. We consider two types of IF—polynomial and bandlimited; we use polynomial interpolating functions for the former, and for the latter, we propose a novel "*line plus sum of sines*" model. The model parameters are estimated by linear regression. Considering the noisy case, LC data for different levels is analyzed, and methods for combining different estimators from LCs are discussed. Theoretical and extensive simulation results show that the performance of the zero-crossing (ZC) based IF estimator and the level-crossing based IF estimator with smaller level values is better than those obtained with higher level values or their combinations. The new technique reaches the Cramér–Rao bound (CRB) roughly above 4 dB signal-to-noise ratio (SNR), and its performance does not deteriorate rapidly with mismatch in the IF order compared with the other techniques in the literature.

Index Terms—Auditory processing, instantaneous frequency, irregular sampling, level-crossing, zero-crossing.

I. Introduction

ALMOST all real-world systems are time-varying in nature. As a result, signals produced by them, such as speech, music, biomedical, and other natural signals show a time-varying spectral characteristic. The spectral variability with time can be decomposed as amplitude variability and frequency variability. Most sources convey information through modulations of amplitude (AM) and modulations of frequency (FM) of a steady carrier, which serves as a vehicle to transport the information contained in the modulations. The basic and simplest signal processing model of such a source is therefore an AM-FM [1]–[3] combination or a sum of such combinations. Analysis of such signals and characterization of their spectral evolution has been the focus of research for quite sometime. Most popular tools developed for this purpose are the time–frequency distributions (TFDs) [4]–[6], which represent energy variations in a joint time–frequency domain. These are designed to satisfy physically appealing notions such as time and frequency marginals, instantaneous frequency, and group delay properties, etc. Most design approaches yield TFDs that do not satisfy all the desired properties. However, this did not restrict their utility. There are several applications based on these distributions, for example, IF estimation and applications to real-world problems [6]–[9]. The alternative set of approaches are more intuitive in nature and motivated by biological signal processing of composite signals such as speech and music [10]–[12]. For example, a joint time–frequency representation based on peripheral auditory signal processing models, known as the Ensemble Interval Histogram (EIH) [10], [11], is well known among the speech processing community. This is not designed to satisfy the general properties of a TFD, but aimed at mimicking the biological functions while processing complex stimuli using signal processing techniques. A modification of this approach uses only zero-crossing information [12], [13]. These techniques assume that the signals are quasistationary (spectrum slowly changing with time) and, hence, can be analyzed, similar to short-time Fourier transform (STFT) on a frame-by-frame basis.

Deviating from the above approaches, we address the question of accurately estimating the time-varying frequency of a signal "*without making a quasistationary assumption.*" Considering single component phase signals,[1] we address the question of generalized IF estimation accurately. Keeping the auditory motivation, we use the level-crossing "*timing information*" for estimating time-varying frequency. The new approach is shown pictorially in Fig. 1. The spike train contains the timing information of the signal crossing a threshold and is a simple model of the neuron that fires every time its excitation crosses a threshold. The timing information is more suitable to handle nonstationary signals. Estimation of IF from the timing information is a problem of reconstruction from irregularly sampled data for which we have shown an effective solution. Further, the estimates from various levels is input to a combining network (denoted by **C**), which may be linear or nonlinear. The output of the combining network is a final estimate of the time-varying frequency of the input mono-component signal.

The organization of the paper is as follows. In Section II, we present in detail the new technique and show how it can be generalized to handle a variety of IF laws. In Section III, we consider the effects of additive white Gaussian noise on level-crossing information and discuss ways of combining the various level-crossing-based IF estimators and study their performance as a function of signal-to-noise ratio (SNR). In Sec-

[1]Phase signals are extensively used in many communication systems such as RADAR and their IF estimation under noisy conditions is of considerable practical importance.
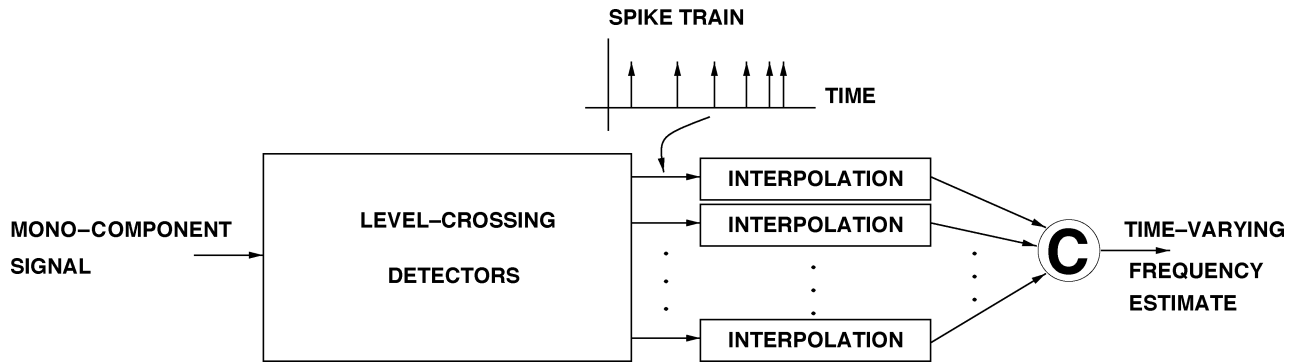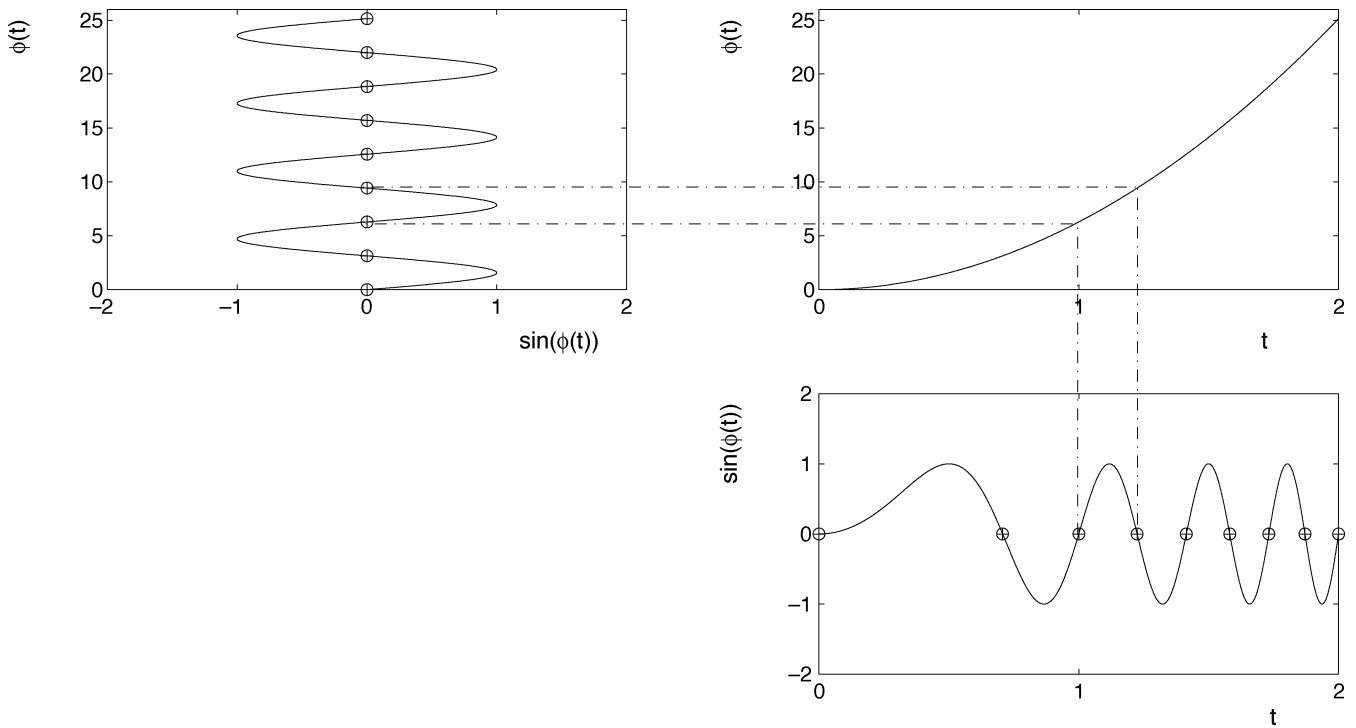
Fig. 1. Auditory motivated technique for estimating time-varying frequency.

Fig. 2. Time-warping ($\oplus$ indicates a zero-crossing).

tions IV and V, we present simulation results comparing the proposed approach with several other popular techniques existing in the literature. Section VI concludes the paper.

II. NOTION OF TIME-WARPING

Consider an arbitrary, bounded IF law $f(t)$, $0 < f(t) < f_{\max}$. We need not consider negative frequencies as they are only conceptual. The corresponding phase function $\phi(t) = 2\pi \int_{-\infty}^{t} f(\tau)d\tau$ is monotonically increasing. Now, consider the frequency modulated signal $s(t) = e^{i\phi(t)}$, $i = \sqrt{-1}$ with IF, $f(t)$, and no amplitude modulation. The inphase and quadrature components of this complex signal carry identical information about the IF, and hence, we confine our discussion to one of them, say, $s_Q(t) = \sin(\phi(t))$. A plot of t versus $s_Q(t)$ shows a sinusoidal signal that appears *warped* in the time domain, as shown in Fig. 2. For increasing/decreasing IF, the successive cycles of the sinusoid get closer/farther, respectively. However, when plotted against $\phi(t)$, $s_Q(t)$ appears as

a constant frequency sine wave. This is true for any nonlinear $\phi(t)$, which causes warping.

If we define $\mathcal{T}^0$ as the set of zero-crossing (ZC) instants of $s_Q(t)$ over an interval $[0, T]$ denoted by $\mathcal{T}^0 = \{t_{j,0}, 0 \le j \le N_0\}$, we get $\phi(t_{j,0}) = j\pi$, $0 \le j \le N_0$, where $N_0 + 1$ is the total number of ZCs within the interval.[2] Instead, if we choose them as level-ℓ crossing (LC) instants (ℓ denotes level value), $\mathcal{T}^\ell = \{t_{j,\ell}, 0 \le j \le N_\ell\}$, where $|\ell| < 1$, we get $\phi(t_{j,\ell}) = j\pi + \sin^{-1}(\ell)$, $0 \le j \le N_\ell$, where $N_\ell + 1$ is the total number of level-ℓ-crossing instants in $[0, T]$. Thus, by detecting the level-crossing instants and putting them in one-to-one correspondence with the instantaneous phase values, we automatically achieve phase unwrapping.[3] Thus, by

---

[2]The only ambiguity in the assignment of $\mathcal{T}^0$ to $\{\phi(t_{j,0}), 0 \le j \le N_0\}$ is in the choice of the initial phase, i.e., $\phi(t_{0,0})$ is arbitrarily set to zero, and the successive phase values are assigned relative to it. For IF estimation, even this ambiguity does not exist because of the derivative operation.

[3]This may be contrasted with the discrete-time approaches to phase computation where phase unwrapping problems occur because of the use of the ARC TANGENT function.

considering the ZC/LC instants of $s_Q(t)$, we obtain specific points on the IP function. A similar result also holds for $s_I(t)$. Thus, ZC/LC instants can be viewed as irregularly sampled information about $\phi(t)$ obtained from either $s_Q(t)$ or $s_I(t)$. Next, we need to perform interpolation under an assumed model to obtain a complete description of the IF for all time within the observation window.

It may be noted that the ZC/LC instants are not uniformly spaced. Interpolating these points to find the IP and IF at any other point within the window is a problem within the domain of *irregular sampling* [14], [15]. The algorithms for reconstruction from irregular samples can be broadly classified as *constrained* or *unconstrained* techniques. Standard interpolation techniques fall into the class of constrained techniques. In the unconstrained type, we seek a solution that describes the IP in an overall least squares sense, because, usually, the number of model parameters to be estimated is lesser than the data size available. The unconstrained approach is also useful when additive noise perturbs the ZC/LC information. Finding $\mathcal{T}^\ell = \{t_{j,\ell}, 0 \leq j \leq N_\ell\}$ and assigning $\phi(t_{j,\ell}) = j\pi + (-1)^j \sin^{-1}(\ell), 0 \leq j \leq N_\ell$, we can perform least-squares curve fitting using a set of interpolating functions. *A priori* information about the nature of IF can help us choose the type of interpolating functions. For polynomial IF, a polynomial curve-fitting approach is suitable. For bandlimited IF, a novel model called the "*line plus sum of sines*" (LPS) model is proposed. In the examples presented in this paper, the nature of IF variation is assumed to be known so that the correct model can be chosen. In the absence of errors due to noise in the signal, ZC/LC instant estimation errors, or model mismatch errors, accurate IF estimation can be done using the proposed approach. However, in the presence of above errors, the unconstrained method is found to be effective and useful. IF estimation based on ZCs and LCs is henceforth referred to as ZC-IF and LC-IF estimation, respectively.

In the discussion so far, we have considered only analog signals. The complex signal $s(t)$, in general, has infinite bandwidth. For the convenience of digital processing, we assume that $x(t)$ is essentially bandlimited to the frequency range $[0, B\pi]$ rad/s. When sampled at a rate of B samples/s (sampling period $= T_s$), we get $\Re\{s[nT_s]\} = cos(\phi[nT_s])$ and $\Im\{s[nT_s]\} = sin(\phi[nT_s])$ ($\Re\{.\}$ denotes the real part, and $\Im\{.\}$ denotes the imaginary part). Let the sampling period be normalized to unity, without loss of generality; thus, $\Re\{s[n]\} = cos(\phi[n])$, and $\Im\{s[n]\} = sin(\phi[n])$. We only observe discrete-time signals from which we estimate the ZC/LC information corresponding to its continuous-time counterpart and then go on to estimate $\phi(t)$. The models for IP and IF are written in continuous-time for the convenience of defining the IF as the derivative of the IP.

In general, the ZC/LC instants do not coincide with the sampling instants. Therefore, we first oversample the received signal (to get $s[rn]$, where $1/r$ is the oversampling factor) and check to see if $(s[rm]-\ell)(s[r(m+1)]-\ell) < 0$ is satisfied. If this is true, then the continuous-time counterpart $s(t)$ has a level-ℓ-crossing in the interval $[rm, r(m+1)]$. For detecting ZCs, oversampling is not required assuming alias-free sampling. The ZC/LC instant may be estimated iteratively to a desired degree of accuracy by using bandlimited interpolation [16] and a bisection approach,

similar to that used in root-finding problems [17]. It has been found experimentally that oversampling by a factor of two (i.e., $r = 0.5$) followed by bisection approach (about three to five iterations) is accurate enough for IF estimation.

A. Polynomial IF Model

For polynomial IF, we can model the IP function as $\phi(t) = \sum_{k=0}^{p} c_k t^k$. The coefficients $\{c_k, k = 0, 1, 2, \ldots, p\}$ are estimated using a least squares fit to $\mathcal{T}^\ell$ and $\{\phi(t_{j,\ell}), 0 \leq j \leq N_\ell\}$. We define a cost function as a measure of total approximation error at the LC instants as

$$\mathcal{C}_p(\mathbf{c}) = \sum_{j=0}^{N_\ell} |\phi(t_{j,\ell}) - \mathbf{c}'\mathbf{e}_{j,\ell}|^2 \tag{1}$$

where $\mathbf{c} = [c_0 \ c_1 \ \ldots \ c_p]'$, and $\mathbf{e}_{j,\ell} = [1 \ t_{j,\ell} \ \ldots \ t_{j,\ell}^p]'$, where $'$ denotes the transpose operator. Minimizing the cost function with respect to $\mathbf{c}$ yields the optimum coefficient vector, using the level-ℓ-crossing information, which is denoted by $\hat{\mathbf{c}}$ and given by

$$\hat{\mathbf{c}} = (\mathbf{H}_\ell' \mathbf{H}_\ell)^{-1} \mathbf{H}_\ell' \mathbf{\Phi}_\ell \tag{2}$$

where $\mathbf{\Phi}_\ell$ is a column vector whose j^{th} element is $\phi(t_{j,\ell}) = j\pi + (-1)^j \sin^{-1}(\ell)$, and $\mathbf{H}_\ell$ is a matrix whose j^{th} row is $\mathbf{e}_{j,\ell}'$. The IF can be easily obtained using $\hat{\mathbf{c}}$.

B. Harmonic IF Model

Polynomial IF law forms a large class of IF variations that occur in many practical situations. However, there are other kinds of IF that are not compactly modeled by polynomial variations; these are functions that can be modeled by a finite set of harmonically related trigonometric sine/cosine functions. In physical systems, since the IF corresponds to the dynamics of a physical parameter, it can be safely assumed to be a bandlimited function and, hence, the harmonic IF model. Since the IP is monotonically increasing, we cannot go in for a straightforward trigonometric polynomial, i.e., finite sum of sines/cosines model as adopted in [15]. We introduce a linear trend in addition to the sum of sines model, called the "*line plus sum of sines*" (LPS) model for IP [18]. This is equivalent to a sum of harmonically related cosines model for IF.

We model the IP $\phi(t)$ and the IF $f(t)$ as follows:

$$\phi(t) = \alpha + \beta t + \sum_{k=1}^{p} \gamma_k \sin(k\omega_o t) \tag{3}$$

$$f(t) = \frac{1}{2\pi}\left[\beta + \omega_o \sum_{k=1}^{p} k\gamma_k cos(k\omega_o t)\right] \tag{4}$$

where $\omega_0 = \pi/T$, $[0, T]$ being the observation interval of $s(t)$. Outside the interval, the specific model used above implies odd-symmetric, periodic extension of IP and, therefore, even-symmetric, periodic extension of IF. This minimizes Gibb's oscillations due to model order truncation, which is a practical necessity. When the actual IF of a physical process is only known to be essentially bandlimited, the LPS model provides a viable alternative to the polynomial model.
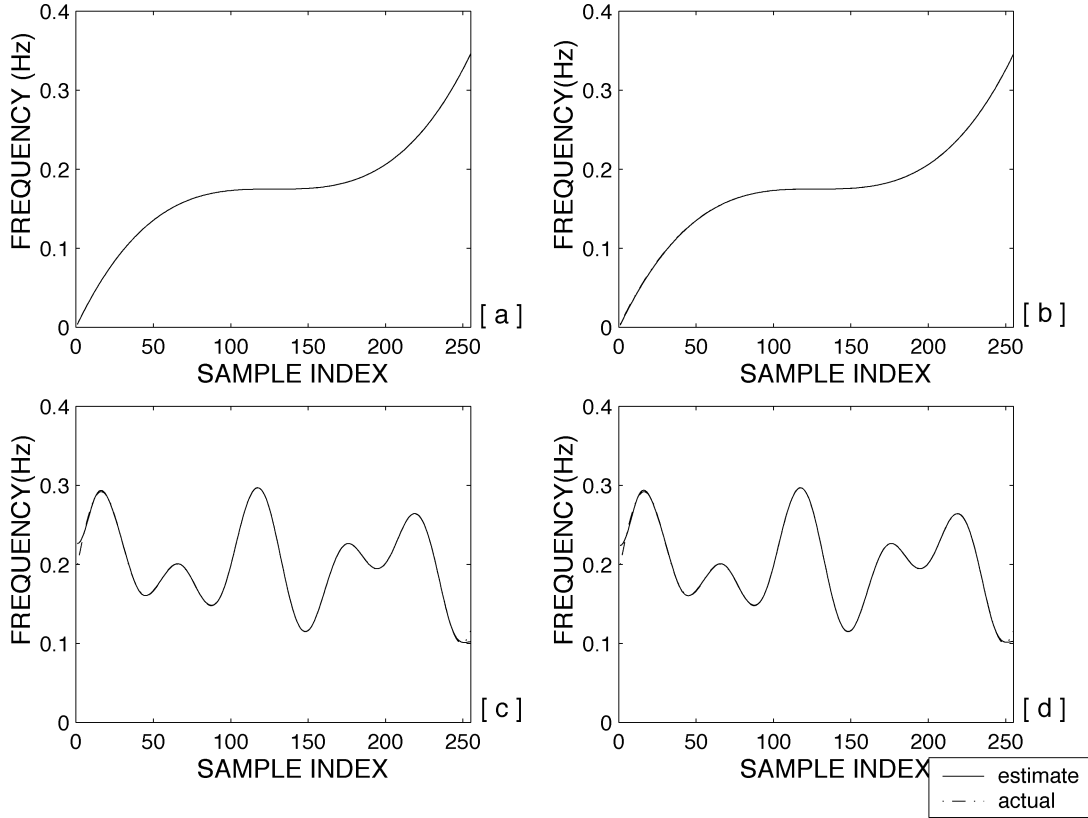
Fig. 3. Cubic polynomial IF estimation. (a) ZC-IF. (b) LC-IF ($\ell = 0.25$); "sum of sines" IF estimation. (c) ZC-IF. (d) LC-IF ($\ell = 0.25$).

We specify a cost-function as a measure of total approximation error of the IP at the level-ℓ-crossing instants by the model chosen above as follows:

$$\mathcal{C}_p(\mathbf{c}) = \sum_{j=0}^{N_\ell} |\phi(t_{j,\ell}) - \mathbf{c}'\mathbf{e}_{j,\ell}|^2 \qquad (5)$$

where $\mathbf{c} = [\alpha \ \ \beta \ \ \gamma_1 \ \ \gamma_2 \ \ \dots \ \ \gamma_p]'$, and $\mathbf{e}_{j,\ell} = [1 \ \ t_{j,\ell} \ \ \sin(\omega_o t_{j,\ell}) \ \ \sin(2\omega_o t_{j,\ell}) \ \ \dots \ \ \sin(p\omega_o t_{j,\ell})]'$ $\phi(t_{j,\ell}) = j\pi + (-1)^j \sin^{-1}(\ell)$, $0 \le j \le N_\ell$. Taking the gradient of $\mathcal{C}_p(\mathbf{c})$ with respect to $\mathbf{c}$ and setting it to zero will yield the optimum (in the least squares sense[4]) coefficient vector for the pth-order IP denoted by $\widehat{\mathbf{c}}$. Thus, we obtain

$$\left[\sum_{j=0}^{N_\ell} \mathbf{e}_{j,\ell}\mathbf{e}'_{j,\ell}\right]\widehat{\mathbf{c}} = \sum_{j=0}^{N_\ell} \phi(t_{j,\ell})\mathbf{e}_{j,\ell}. \qquad (6)$$

In matrix form, this is written as $\mathbf{R}_\ell\widehat{\mathbf{c}} = \tilde{\mathbf{e}}_\ell$, where $\mathbf{R}_\ell = [\sum_{j=0}^{N_\ell} \mathbf{e}_{j,\ell}\mathbf{e}'_{j,\ell}]$, $\tilde{\mathbf{e}}_\ell = \sum_{j=0}^{N_\ell} \phi(t_{j,\ell})\mathbf{e}_{j,\ell}$; $\widehat{\mathbf{c}}$ can be found by matrix inversion. Using $\widehat{\mathbf{c}}$ in (4), the IF can be estimated.

C. Illustrations

We illustrate the performance of ZC and LC approaches to IF estimation using the models described above considering two examples: a phase signal with a polynomial IF given by $f[n] =$

[4]A nice study of linear least squares estimation with emphasis on order recursion, order selection and computational efficiency can be found in [19] and the references therein.

$0.175 + 0.175((n - (M/2) + 1)^3/(M/2)^3)$ and another with a "sum of sines" IF given by $f[n] = 0.2 + 0.05\sin(0.12n) + 0.05\sin(0.07n)$, $0 \le n \le M - 1$, $M = 256$. We assume that the nature of the underlying IF variation is known *a priori* so that the appropriate model is employed. Estimating the IF when the underlying model is not known *a priori* is reported separately [20]. The ZC-IF and LC-IF estimates (corresponding to an arbitrarily chosen level value of 0.25) are shown in Fig. 3. It is found that the proposed approaches perform very accurate IF estimation. The error in IF estimation due to various mismatches and effect of noise are discussed in the next section.

III. EFFECT OF ADDITIVE NOISE

Consider $s(t)$ corrupted with complex Gaussian noise $w(t)$, whose power spectral density $S_{ww}(\omega) = 2\sigma_w^2/B$ for $0 \le \omega \le B\pi$ and zero otherwise; σ_w^2 is thus the variance of noise. The received signal is given by $x(t) = e^{i\phi(t)} + w(t)$. Following the analysis in [21], we can write $x(t) = e^{i\phi(t)}(1 + v(t))$ with $\mathrm{var}\{v(t)\} = \sigma_w^2$ where var denotes variance. $v(t)$ can be written as $v(t) = v_I(t) + iv_Q(t)$, where $v_I(t)$ and $v_Q(t)$ are the in-phase and quadrature components of $v(t)$, respectively. We can write

$$1 + v(t) = \sqrt{(1 + v_I(t))^2 + v_Q^2(t)}\, e^{itan^{-1}\left[\frac{v_Q(t)}{1 + v_I(t)}\right]} \approx e^{iv_Q(t)} \qquad (7)$$

under high-SNR assumption ($\sigma_w^2 \ll 1$). In addition, $\mathrm{var}\{v_I(t)\} = \mathrm{var}\{v_Q(t)\} = \sigma_w^2/2$. With the above assumptions and approximations, we can write

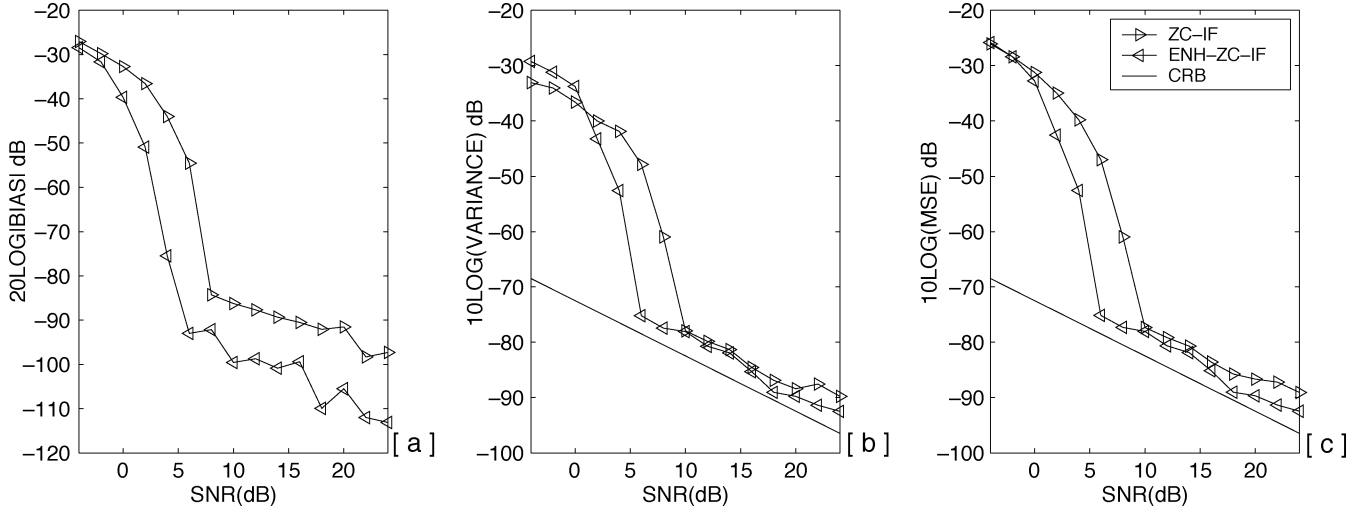$$x(t) \approx e^{i[\phi(t) + v_Q(t)]}. \qquad (8)$$

Fig. 4.   Quadratic polynomial IF estimation and performance improvement of ZC-IF estimator with TV filtering. (a) Squared bias. (b) Variance. (c) Mean square error (MSE), all in decibels, as a function of SNR (in decibels), at the center of the window (ENH refers to signal enhancement due to TV filtering).

$v_Q(t)$ can thus be treated as *additive phase noise*. The phase noise is real, zero-mean, and Gaussian distributed[5] with a spectral density $S_{v_Q v_Q}(\omega) = \sigma_w^2/2B$ for $|\omega| \leq B\pi$ and zero otherwise. $w(t)$ when sampled at a rate of B samples/s yields uncorrelated Gaussian distributed phase noise samples. The implication of this result is that, under a high-SNR assumption, the sampled instantaneous phase function gets additively corrupted with white noise. The additive phase noise property justifies least squares approach to IF estimation. The phase perturbation of (8) gets reflected as ZC/LC perturbation in the signal domain, which is discussed further in the next section.

A. Perturbation Analysis

Consider the quadrature component of the noisy phase signal. Let ℓ, $|\ell| < 1$ denote an arbitrary amplitude level. Let $t_{n,\ell}$ denote the nth-level-ℓ-crossing instant in the presence of noise. Let the corresponding level-crossing instant for the clean signal be $(t_{n,\ell} - \delta t_{n,\ell})$ with $\delta t_{n,\ell}$ denoting the perturbation in the LC instant. Thus, we have

$$\sin\left(\phi(t_{n,\ell})\right) + w_Q(t_{n,\ell}) = \ell \qquad (9)$$
$$\sin\left(\phi(t_{n,\ell} - \delta t_{n,\ell})\right) = \ell. \qquad (10)$$

We model the perturbation $\delta t_{n,\ell}$ as a zero-mean, uncorrelated random variable [13]. Consider

$$\mathcal{E}\left\{\cos\left(\phi(t_{n,\ell}) - \phi(t_{n,\ell} - \delta t_{n,\ell})\right)\right\}$$
$$= \ell^2 + \sqrt{1-\ell^2}\,\mathcal{E}\left\{\sqrt{1 - (\ell - w_Q(t_{n,\ell}))^2}\right\}. \quad (11)$$

Using a second-order Taylor series approximation to evaluate the expectations involved, we obtain, after simplification (assuming stationary additive noise, high SNR, and $\delta t_{n,\ell} \ll 1$)

$$\sigma_{\delta t_{n,\ell}}^2 = \frac{\sigma_w^2}{2\left(\phi'(t_{n,\ell})\right)^2 (1-\ell^2)}. \qquad (12)$$

[5]Juxtapose this result with that in [22], where the phase noise is shown to have a Tikhonov probability density function. In [22], the wrapped phase is considered, whereas in the present analysis, we have the unwrapped phase and, hence, the difference.

Clearly, $\delta t_{n,\ell}$ is nonstationary. The dependence of the variance of the perturbation of the LC instant on the level value is explicit. The expression also indicates that the higher level-crossing instants at lower frequencies have larger variance compared to lower level-crossing instants at higher frequencies. The above analysis is a generalization of the results in [12] and [13] to time-varying signals. For the simpler case of a constant-frequency sinusoid, $\sigma_{\delta t_{n,\ell}}^2$ is least in the case of ZC instants ($\ell = 0$) and it increases with increasing level value. Equation (12) also suggests that since the noise variance of different LC instant perturbations is different for different levels, it would be appropriate to perform LC-IF estimation differently for each level and then combine them appropriately. This is discussed in Section III-E.

B. ZC-IF Estimation

Consider IF estimation using ZC information. To study the performance of the proposed approach, a phase signal with a quadratic IF given by $f[n] = 0.689[0.25 + 0.25((n - (M/2) + 1)^2/(M/2)^2)]$, $0 \leq n \leq M - 1$ is generated with $M = 256$. A normalized sampling period of unity is assumed. Complex white Gaussian noise of a suitable variance is added to generate a signal of a desired SNR. Since polynomial IF is considered, the IF is estimated as explained in Section II-A, using the quadrature component. For each SNR chosen, 100 Monte-Carlo trials are run to compute the statistics of the estimator. SNR is varied from -4 to 24 dB in steps of 2 dB. The squared bias ($B^2[n]$), variance ($V[n]$), and mean square error ($\text{MSE}[n] = B^2[n] + V[n]$) of the IF estimator at the center of the observation window ($n = 128$) are plotted as a function of the SNR in Fig. 4 (see legend "ZC−IF"). For the sake of uniformity across all SNRs and trials, $p = 3$ is used in the simulations. It can be seen that the ZC-IF estimator is nearly efficient beyond SNR = 10 dB in the sense that it lies quite close to the CRB. The derivation of the CRB is outlined in Appendix I.

C. Performance Improvement

A typical FM receiver has a front end that consists of a bandpass filter and a limiter [3]. The bandpass filter serves to cut off
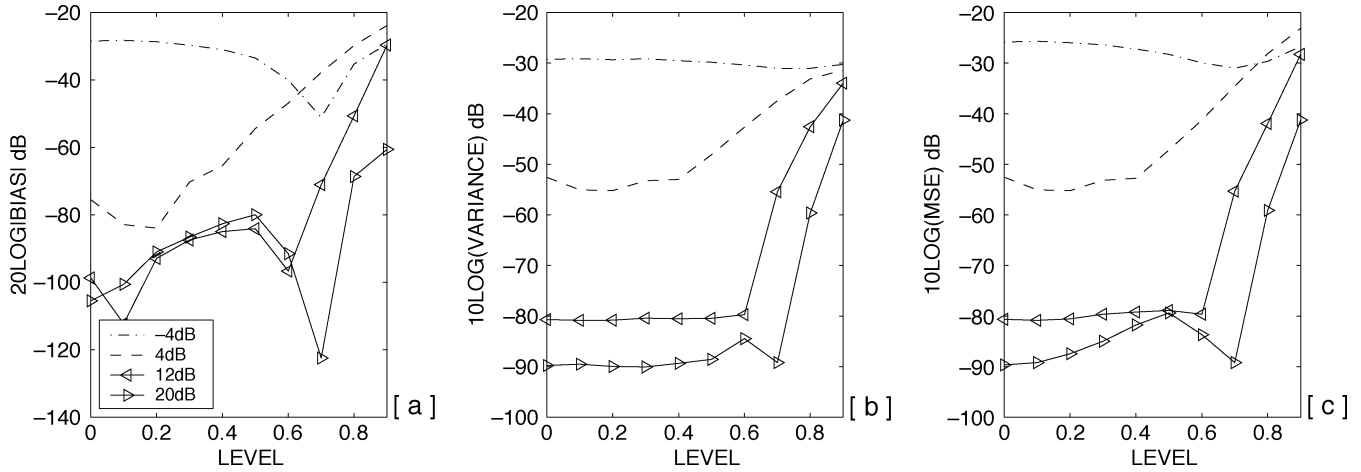
Fig. 5. Quadratic polynomial IF estimation; performance comparison of different ENH-LC-IF estimators. (a) Squared bias. (b) Variance. (c) MSE, all in decibels, as a function of level, with SNR (in decibels) as the parameter.

out-of-band noise that can otherwise cause distortion in the demodulated message. The limiter retains only ZCs for demodulation. We adopt a similar front-end to improve the performance of the proposed IF estimator. Since we are dealing with nonstationary signals and the spectrum of *out-of-band* noise changes with time, we use an STFT-based time-varying (TV) bandpass filter prior to IF estimation. A limiter used in conventional FM receivers is not used here because we also explore the use of LC information for IF estimation (Section III-D).

Consider the following specification for the TV bandpass filter (the notation means that n is discrete and ω is continuous):

$$\mathcal{L}[n,\omega] = \begin{cases} 1, & \text{for } |\omega - \omega[n]| < \sigma[n], \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

where $\omega[n]$ is the center frequency of the filter at time n, and $\sigma[n]$ is the spread/bandwidth about $\omega[n]$. The output of such a system to an input, $x[n]$ is given by

$$\widehat{s}[n] = \frac{1}{2\pi h[0]} \int_{-\pi}^{+\pi} \mathcal{X}[n,\omega]\mathcal{L}[n,\omega]e^{i\omega n}d\omega \tag{14}$$

where $h[n]$ is the window function used in obtaining the STFT of $x[n]$ denoted by $\mathcal{X}[n,\omega]$ and defined as $\mathcal{X}[n,\omega] = \sum_m h[n-m]x[m]e^{-j\omega m}$. The TV bandpass filter center frequency is estimated as $\omega[n] = arg\max_\omega |\mathcal{X}[n,\omega]|^2$. The bandwidth factor $\sigma[n]$ can be chosen as the width of the main lobe about $\omega[n]$; this is fixed, depending on the analysis window. More insights into the TV filtering process are given in [23], [24], and the references therein.

The experiment of quadratic IF estimation is repeated with the above preprocessing step. The results are shown in Fig. 4 (see legend "ENH − ZC − IF"). It can be seen that the threshold SNR of the estimator has been reduced by about 4 dB. Noting this advantage, henceforth, in all simulations, we perform ZC/LC-IF estimation on $\widehat{s}[n]$ whenever noise is present. The case of polynomial IF is considered only for the sake of illustration; the conclusions reached in this section hold for

bandlimited IF as well. In the absence of any *a priori* information, we found that a wideband spectrogram gives a reasonable improvement in SNR. Therefore, a 17-point Hamming window is used in all simulations.

D. LC-IF Estimation

We repeat the simulations of Section III-B for LC-IF estimation. Based on the unity amplitude phase signal assumption, ten level values were chosen equispaced between 0 and 0.9, both inclusive. To make a comparative observation, the squared bias, variance, and MSE at the center of the window are plotted in Fig. 5 with SNR as the parameter. We also show similar results in Fig. 6 for a "sum of sines" IF $f[n] = 0.2 + 0.05\sin(0.12n) + 0.05\sin(0.07n), 0 \le n \le M-1$ (parameters first chosen arbitrarily and later modified to ensure that $0 < f[n] < 0.5$). $M = 256$ was used in the simulations. For the sake of uniformity in comparison across all SNRs and realizations, a fixed-order, $p = 24$, is used in the Monte-Carlo analysis. From these plots, we observe that for moderate to high SNRs, ENH-ZC-IF and ENH-LC-IF estimators with low-level value exhibit similar performance characteristics. In general, at any SNR, the performance of ENH-LC-IF estimators is the same (for low level value) or worse (for high level value) compared with ENH-ZC-IF estimation. As SNR increases, the ENH-LC-IF estimators corresponding to the higher level values become more reliable.

E. Combined Estimators

Keeping the foregoing discussion in mind, it would be natural to ask the following question: **Can different LC-IF estimators be combined to get an estimator that possibly offers improved performance than the individual estimators?** We attempt to answer this question by considering some combinations.

1) The simplest combination would be a weighted average, i.e., define a new IP estimator $\widehat{\phi}(t)$ as $\widehat{\phi}(t) = \sum_{l=0}^{L-1} w_l \widehat{\phi}^l(t)$, where L is the total number of levels, and $\widehat{\phi}^l(t)$ is the IP estimate from the lth
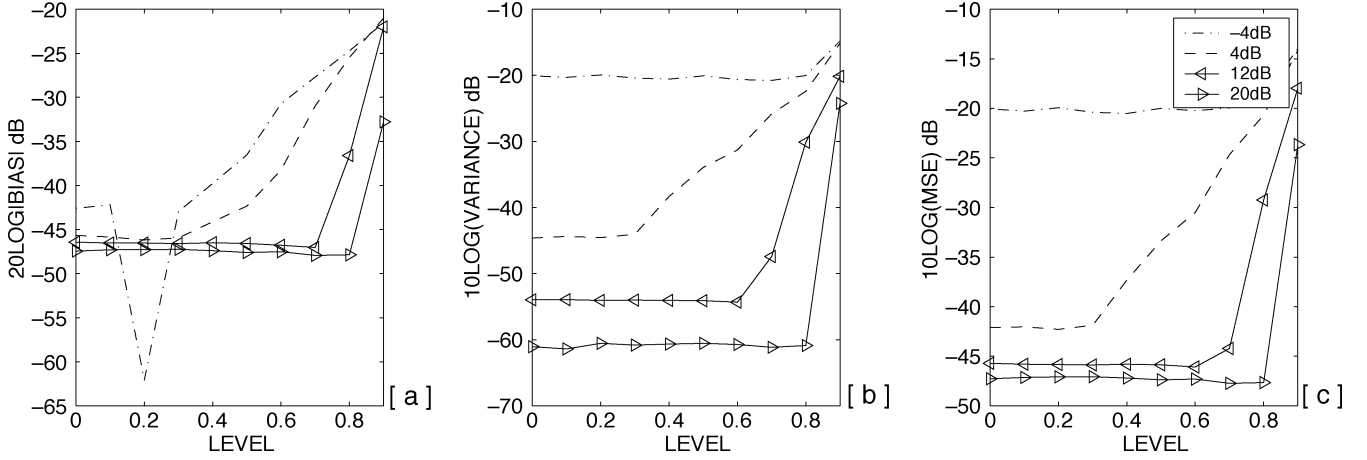
Fig. 6. "Sum of sines" IF estimation; performance comparison of different ENH-LC-IF estimators. (a) Squared bias. (b) Variance. (c) MSE, all in decibels, as a function of level, with SNR (in decibels) as the parameter.

level-crossing[6] information. The IF can be estimated as $\hat{f}(t) = (1/2\pi)(d\hat{\phi}(t)/dt)$. With the individual estimators assumed approximately unbiased, for $\hat{f}(t)$ to be unbiased as well, we must have $\sum_{l=0}^{L-1} w_l = 1$ with each $w_l \geq 0$ to make the resulting estimator positive. Equal weightage, i.e., mean of all estimators, implies that, irrespective of the variance in each LC estimate, the confidence in all estimators is the same. Intuitively, this is not expected to offer any improvement in performance and might result in an estimator with poorer performance than any of those being combined.

2) The result of Section III-A can be used to find optimal weights in a *minimum variance sense*, i.e., we can find weights, $\{w_l, l = 0, 1, \ldots, L-1\}$ such that var$\{\hat{\phi}(t)\}$, where $\hat{\phi}(t) = \sum_{l=0}^{L-1} w_l \hat{\phi}^l(t)$ is the least, subject to the constraints $\sum_{l=0}^{L-1} w_l = 1$ and $w_l \geq 0, 0 \leq l \leq L-1$.

The problem of LC-IF estimation is essentially one of fitting a curve to the data sets $\{t_{j,\ell} + \delta t_{j,\ell}, 0 \leq j \leq N_\ell\}$ and $\{\phi(t_{j,\ell}) = j\pi + (-1)^j \sin^{-1}(\ell), 0 \leq j \leq N_\ell\}$. The same can be recast, using a first-order Taylor series approximation, to a problem of curve fitting to the data sets $\{t_{j,\ell}, 0 \leq j \leq N_\ell\}$ and $\{\phi(t_{j,\ell} + \delta t_{j,\ell}) = j\pi + (-1)^j \sin^{-1}(\ell) = \phi(t_{j,\ell}) + \phi'(t_{j,\ell})\delta t_{j,\ell}\}$, where $\phi'(t_{j,\ell})\delta t_{j,\ell}$ is a zero-mean uncorrelated random variable and has variance $\sigma_w^2/(2(1 - \ell^2))$. Although it is strictly not true, we assume that $\hat{\phi}^l(t)$, for different l, are independent of each other. We have var$\{\hat{\phi}^l(t)\} \propto (\sigma_w^2/(1 - (l/L)^2))$ [by ignoring differences in the observation matrix, $\mathbf{H}_\ell$ (see Section II-A), which is different for different levels]. The problem of finding optimal weights can be solved easily using the method of *Lagrange multipliers*. The optimal weights, which are denoted by w_l^{opt}, can be found to be $w_l^{\text{opt}} = (1 - (l/L)^2)/(\sum_{k=0}^{L-1} 1 - (k/L)^2), 0 \leq l \leq L-1$. This choice of weights assures the minimum possible vari-

[6]We explain the notation followed: L is the total number of equispaced levels between 0 and 1 (including 0, excluding 1), and ℓ is a level value and takes values $0, 1/L, 2/L, \ldots, (L-1)/L$; the corresponding level index is denoted by l, which takes values $0, 1, 2, \ldots, L-1$. Thus, $\ell = l/L$, i.e., the lth level value is ℓ.

ance for $\hat{\phi}(t)$ under the assumptions made. However, it may be noted that the variance cannot go below the CRB.

3) Define $\hat{f}(t) = \text{median}\{\hat{f}^0(t), \hat{f}^1(t), \hat{f}^2(t), \ldots, \hat{f}^{L-1}(t)\}$, where $\hat{f}^l(t)$ is the IF estimate from the lth level-crossing information.

The performance of the proposed combinations was studied and the results are summarized in Fig. 7(a) and (b) for the examples considered earlier. It is clear that ENH-ZC-IF estimation is the best. This is followed by the median. The performance improvement with the median combination is due to the fact that outlier estimates that usually correspond to higher level values, and of higher variance, are rejected by the median. Any robust combination with an outlier rejection property can also be used instead of the median.

Apart from the above combinations, *ad hoc* combinations were also tried by choosing the weights as shown in Fig. 7(c). The choice of the weights is motivated by the observation that the ENH-LC-IF estimators with level value close to zero performed similar to the ENH-ZC-IF estimator. In such a case, assuming independent ENH-LC-IF estimators, using such weights should give us an estimator with a variance smaller than any of the individual estimators. However, from Fig. 7(c), where the variance is plotted for the quadratic IF example, it turns out that, at best, the combined estimator achieves the performance of the ENH-ZC-IF estimator, indicating that our assumption of independence of various ENH-LC-IF estimators was not strictly true. The same is true for the "sum of sines" IF example as well.

To summarize the results, the ENH-ZC-IF estimator and the ENH-LC-IF estimators with small level-values exhibit similar performance. This implies robustness of ZC and low level-crossing information compared with higher level-crossing information (which is also substantiated by the analysis in Section III-A). Linear and nonlinear combinations, such as median, of different ENH-LC-IF estimators, at best, offer the same performance as ENH-ZC-IF estimator/low-level ENH-LC-IF estimator. The lack of improvement in performance of the combined estimator is due to the fact that the different ENH-LC-IF estimators are probably not mutually independent.
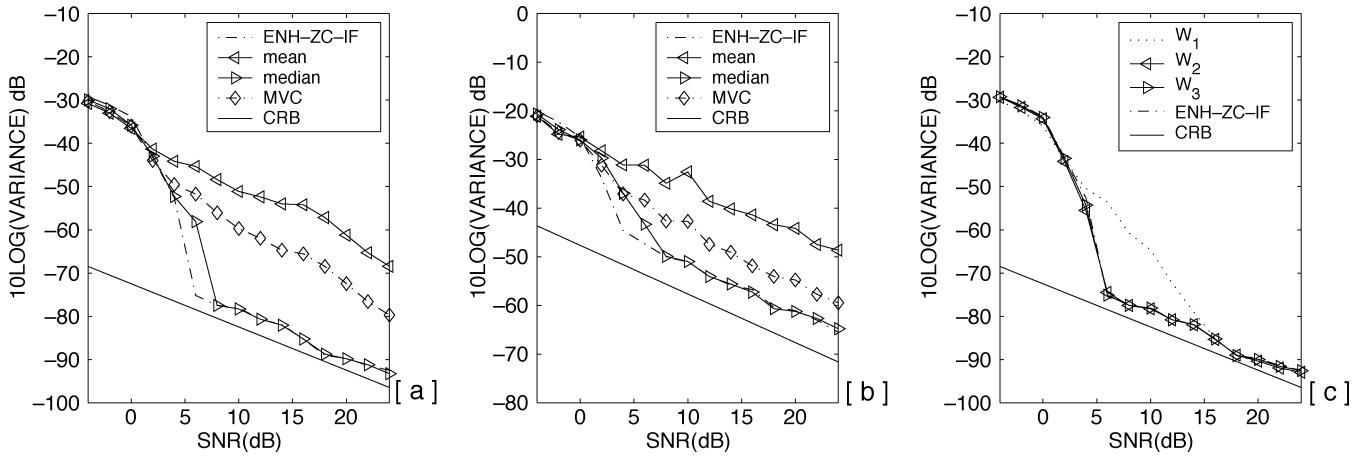
Fig. 7. Variance of different combinations of IF estimators as a function of SNR (in decibels). (a) Quadratic polynomial IF. (b) "Sum of sines" IF. (c) Quadratic polynomial IF. $W_1 = \{1/8, 1/8, 1/8, 1/8, 1/8, 1/8, 1/8, 1/8, 0, 0\}$, $W_2 = \{1/4, 1/4, 1/4, 1/4, 0, 0, 0, 0, 0, 0\}$, $W_3 = \{1/2, 1/2, 0, 0, 0, 0, 0, 0, 0, 0\}$. In (a) and (b), MVC stands for minimum variance combination.
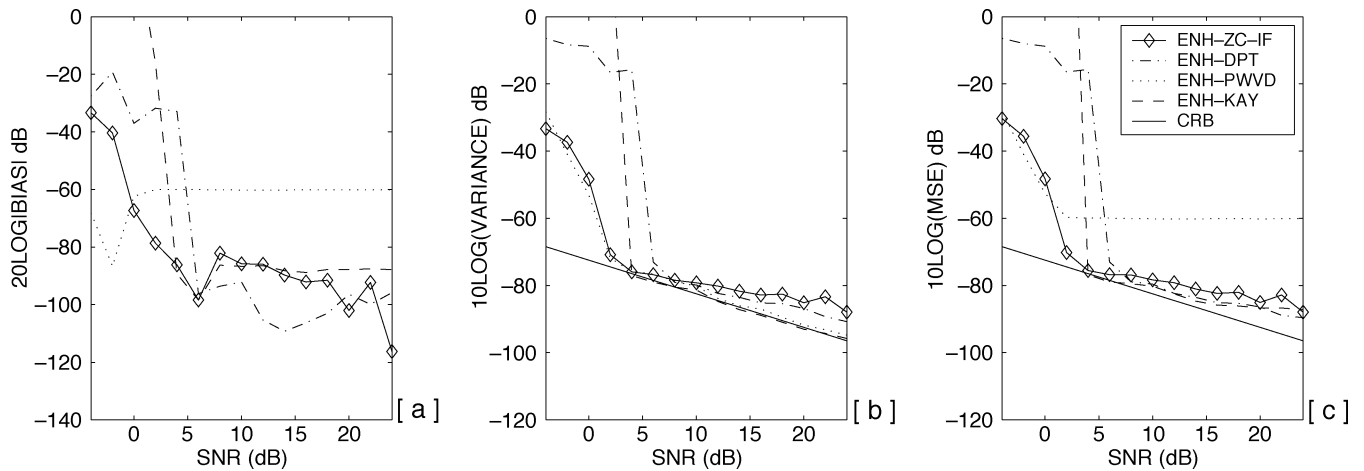

Fig. 8. Cubic polynomial IF estimation. Performance comparison of different techniques. (a) Squared bias (in decibels), (b) variance (in decibels), and (c) MSE (in decibels) versus SNR (in decibels) at the center of the window.

IV. PERFORMANCE ANALYSIS AND COMPARISON

In this section, we compare the performance of the new approach with some popular techniques in the literature. For the purpose of comparison, we choose the ENH-ZC-IF estimator. Among the TFD-based approaches, we choose polynomial Wigner–Ville distribution (PWVD) [25]; among the non-TFD-based approaches, we choose Kay's regression method [21] and the discrete polynomial-phase transform (DPT) approach [26], [27] for comparison. By far, these are the most popular techniques in the literature on IF estimation. For the sake of uniformity across techniques, range of SNRs, and realizations, we assume that the order of the IF is known *a priori*. The performance of the techniques when the order is unknown will be investigated later.

We consider cubic polynomial IF given by $f[n] = 0.150 + 0.150((n - (M/2) + 1)^3/(M/2)^3)$, $0 \le n \le M - 1$, where $M = 256$. Complex white Gaussian noise of a required variance is added to the phase signal to achieve a desired SNR. For ZC-IF estimation, only the quadrature component of the resulting noisy phase signal is used, whereas all other techniques use the complex noisy signal itself. In Section III-C, it was shown that

time-varying filtering results in improved IF estimation performance; hence, for an even comparison, we use the same front-end for all techniques. The prefix ENH is used to denote the same. One hundred Monte-Carlo realizations of the noisy signal are generated, and IF estimation is performed for each realization. From this, the bias, variance, and MSE of the IF estimators are computed. These quantities corresponding to the center of the window, as a function of SNR, are displayed in Fig. 8. From the figure, it is clear that ENH-ZC-IF and ENH-PWVD follow the CRB closely above 4 dB SNR. Below 4 dB SNR, both ZC-IF and ENH-PWVD perform equally well and are better than the other techniques. The MSE plots show that above 4 dB, all techniques perform nearly identically, whereas ENH-ZC-IF and ENH-PWVD are better in the low SNR region. The performance of all the techniques as a function of time is shown for SNR = 2 dB (close to the *knee-SNR* for ENH-ZC-IF) in Fig. 9, from which we observe that ENH-ZC-IF and ENH-PWVD are much better than ENH-KAY and ENH-DPT. ENH-PWVD has a variance close to the CRB only at the center of the window; elsewhere, it is slightly away from the bound. This is true even for high SNRs. Amongst ENH-ZC-IF and ENH-PWVD, the former is marginally superior at low SNRs. At high SNRs, ENH-ZC-IF, ENH-KAY, and ENH-DPT exhibit
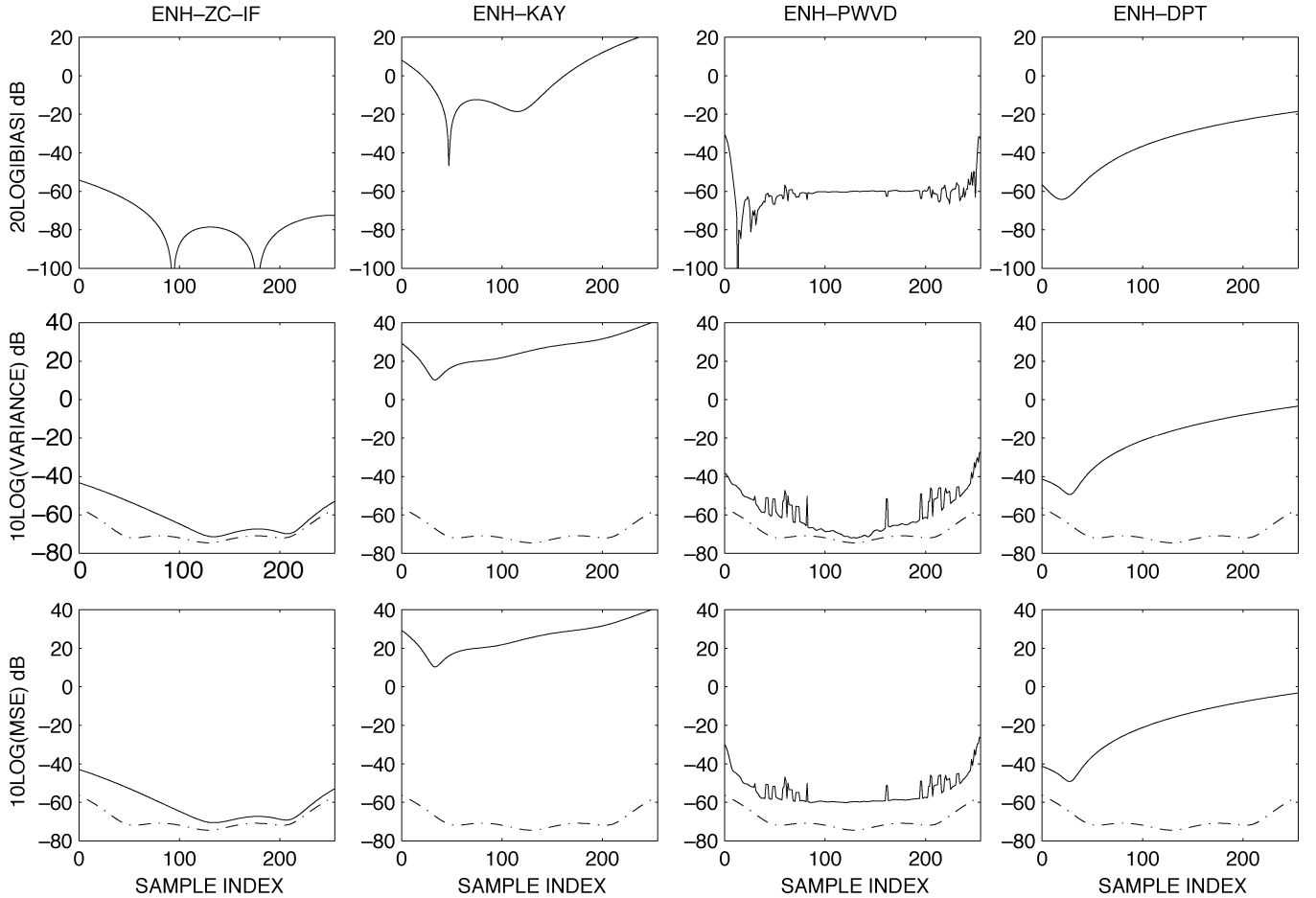
Fig. 9.   Cubic polynomial IF estimation (SNR = 2 dB). Performance comparison of different IF estimation techniques. (a) Squared bias (in decibels), (b) variance (in decibels), and (c) MSE (in decibels) as a function of the sample index.

variance close to the CRB throughout the window, unlike ENH-PWVD. From the bias characteristics, we find that the least bias is exhibited by the ENH-ZC-IF estimator. The PWVD- and DPT-based IF estimators, though theoretically unbiased, practically exhibit small nonzero bias due to finite nonzero error in peak-picking, even with zero-padded discrete Fourier transform (DFT) computation. Such errors also arise in detecting ZCs, but the subsequent least squares regression seems to reduce that error and, hence, lower bias compared with other techniques. From Fig. 9, we also observe that the number of valleys of the CRB within the observation window is equal to the order of the IF polynomial.

V. DISCUSSION

We discuss certain characteristics of the proposed IF estimator vis-a-vis other techniques in the literature. The asymptotic bias and variance analysis is given in Appendix II.

A. Threshold Effect

The LC-IF estimator, like many other nonlinear estimators, exhibits "*threshold effect*," i.e., beyond a particular SNR ("*threshold/knee-SNR*"), the variance of the estimator closely follows the CRB, whereas below the threshold SNR, the performance deteriorates rapidly. The threshold SNR is 4 dB for linear IF and 6 dB for quadratic and cubic IF. This implies that with increase in the order of the polynomial, the threshold

SNR does not increase rapidly. On the contrary, for the DPT approach, the threshold SNR increases by about 6 dB whenever the order is incremented by one [28].

B. Identifiability

In parameter estimation from sampled phase signals, the issue of identifiability arises in two different contexts.

1)   The signal $s(t) = e^{i \sum_{k=0}^{p} c_k t^k}$ is a function of the parameters $\{c_k, k = 0, 1, 2, \ldots, p\}$. Assuming a normalized sampling period of unity, we get a discrete-time signal $s[n]$, which is also a function of the parameters $\{c_k, k = 0, 1, 2, \ldots, p\}$. However, the parameter set $\{c_k + 2\pi m, k = 0, 1, 2, \ldots, p\}$ for arbitrary integer m also yields the same signal $s[n]$. This ambiguity arises because of aliasing of the polynomial phase parameters [29]. This is concerned with uniform sampling and cannot be taken care of at the parameter estimation level, however accurate the estimation technique may otherwise be.

2)   Starting with uniformly sampled signals, approaches such as the DPT, which use the discrete-time Fourier transform (DTFT), give rise to ambiguous parameter estimates. The constraints on the parameters such that these ambiguities do not arise is discussed in [27]. If any of the parameters violates the constraint, the DPT
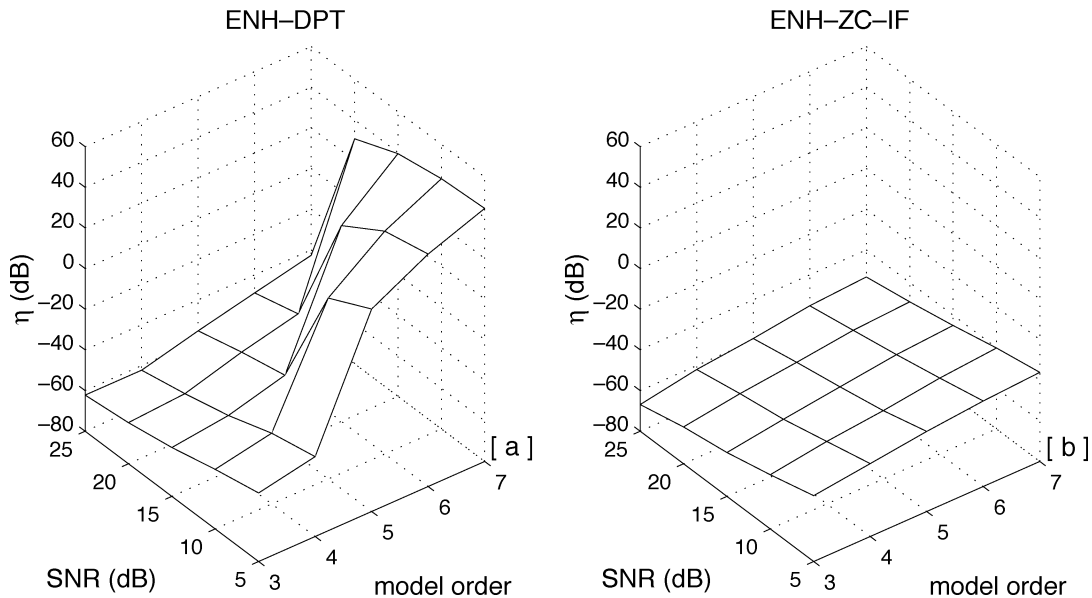
Fig. 10. Quadratic polynomial IF estimation performance comparison of (a) ENH-DPT and (b) ENH-ZC-IF techniques as a function of SNR (in decibels) and model order of the IP.

approach yields inaccurate results, even if the IF lies in the normalized frequency range [0, 0.5]. Since the present technique does not use DTFT, such ambiguities do not arise. To illustrate this point, consider a phase signal with cubic polynomial IF given by $f[n] = 0.175 + 0.175((n - (M/2) + 1)^3/(M/2)^3), 0 \le n \le M - 1$, where $M = 256$. The IF lies in the range [0, 0.5]. However, since all parameters do not obey the constraints, the DPT approach fails to estimate the IF correctly, whereas the ZC/LC technique can accurately estimate the IF.

C. Knowledge of IF Order

The simulations in Section IV were carried out assuming *a priori* knowledge of the order of the polynomial IF only to eliminate the effects of order mismatch in the analysis. In practice, we may not know the true order, and hence, the performance of the IF estimator will depend on the assumed model order. To study the effect of order mismatch on the performance of the two techniques ENH-ZC-IF and ENH-DPT, an experiment was conducted. A phase signal with a quadratic polynomial IF (same as in Section III-B) was generated. Complex white Gaussian noise was added to achieve a desired SNR. The SNR was varied from 5 to 25 dB in steps of 5 dB. For every SNR, 100 Monte-Carlo trials were run to obtain the statistics of the estimator. The IP model order was varied from 3 (actual order) to 7. An average IF estimation error measure was used to study the effect of the assumed order on the performance of the estimators. It is denoted by η and defined as

$$\eta = \frac{1}{RM} \sum_{m=1}^{R} \sum_{n=0}^{M-1} \left(f[n] - \widehat{f}[n] \right)^2 \qquad (15)$$

where R is the total number of realizations (100 in the present experiment), and M is the data size. It may be noted that this error is different from the model fitting error in (1) and (5). η is

plotted on a decibel scale in Fig. 10 for different SNRs and orders. From the plots, it is clear that at all SNRs, the ENH-ZC-IF technique shows gradual increase in error as order is increased (over-fitting) and/or SNR is decreased. On the other hand, the ENH-DPT technique shows a steep increase in error for high orders and low SNRs. We thus conclude that ZC-IF estimation is less sensitive to order mismatch than DPT.

VI. CONCLUSIONS

We have addressed the problem of estimating generalized polynomial and bandlimited IF from the corresponding phase signal, taking an auditory front-end processing motivation. Unlike other zero-crossing-based approaches, it is shown that without the need for a quasistationary assumption, the signal level-crossing information can be effectively used for IF estimation from both clean and noisy signals. It is found that ZC and LC-IF estimators with low-level value are the best among all LC-IF estimators, and their IF estimation performance is close to the CR bound. The proposed approach is general enough to handle a variety of IF laws. A detailed comparison with other techniques in the literature is also shown. In this paper, we have demonstrated the utility of ZC/LC *instant* information, which is different from the conventional ZC/LC rate [30]/interval information, for time-varying signal parameter estimation. The link between irregular sampling and level-crossing analysis for time-varying signal parameter estimation holds promise for further research.

APPENDIX I
CRAMÉR–RAO LOWER BOUND

A. Polynomial Model

In this Appendix, we briefly outline the derivation of the CRB in the level-crossing framework. For a detailed discussion on performance bounds for general nonuniform and random sampling schemes, see [31]. Consider the phase signal $s(t) = e^{i\phi(t)}$

in zero-mean, complex additive Gaussian noise $w(t)$ of variance σ_w^2 to yield the noisy signal $x(t)$. The instantaneous phase is given by $\phi(t) = \sum_{k=0}^{p} c_k t^k$. To perform IF estimation, we need to estimate $\mathbf{c} = [c_0, c_1, \ldots, c_p]$. Consider the set of level-ℓ crossing instants $\mathcal{T}^\ell = \{t_{j,\ell}, 0 \leq j \leq N_\ell\}$. Correspondingly, we have $x(t_{j,\ell}) = s(t_{j,\ell}) + w(t_{j,\ell}), 0 \leq j \leq N_\ell$. In vector form, this can be written as $\mathbf{x} = \mathbf{s} + \mathbf{w}$, each of dimension $(N_\ell + 1) \times 1$. The mean of the data vector, which is denoted by $\mu(\mathbf{c})$, is $\mathbf{s}$. The covariance matrix C has $\mathcal{E}\{w(t_{g,\ell})w^*(t_{h,\ell})\}$ as the gth row, hth column entry. The matrix C is not perfectly diagonal but is diagonally dominant, i.e., the diagonal elements are the maximum elements in a given row and column. The corresponding Fisher information matrix $\mathcal{J}^\ell$ is a square matrix of size $(p+1) \times (p+1)$. The gth row, hth column entry of $\mathcal{J}^\ell$ is given by the following (H denotes Hermitian transpose):

$$
\begin{aligned}
(\mathcal{J}^\ell)_{g,h} &= \frac{\partial \mu^H(\mathbf{c})}{\partial c_g} C^{-1} \frac{\partial \mu(\mathbf{c})}{\partial c_h} \\
&= \mathbf{s}^H T_\ell^{g-1} C^{-1} T_\ell^{h-1} \mathbf{s} \\
&\approx \sigma_w^2 \mathbf{s}^H T_\ell^{g+h-2} \mathbf{s}, \quad 1 \leq g, h \leq p+1
\end{aligned}
\tag{16}
$$

where the uth row, vth column entry of T_ℓ^{g+h} is given by $(T_\ell^{g+h})_{u,v} = t_{u,\ell}^{g+h} \delta_{u,v}$ ($\delta_{u,v}$ is the Kronecker delta), $1 \leq u, v \leq N_\ell + 1, 1 \leq g, h \leq p+1$, and the approximation in the last step is that $C \approx \sigma_w^2 I_{(N_\ell+1)\times(N_\ell+1)}$, I being the identity matrix. The variance bound on the gth parameter estimate $\widehat{c}_g$ is given by $C_{\widehat{c}_g} \geq (\mathcal{J}^{\ell^{-1}})_{g,g}$.

The IF is given by $\widehat{f}^l[n] = (1/2\pi) \sum_{k=1}^{p} k \widehat{c}_k n^{k-1}$ and the CRB is given by

$$
\mathrm{CRB}\left\{\widehat{f}^l[n]\right\} = \frac{1}{4\pi^2} \tau C_{\widehat{c}_{(2:p+1,2:p+1)}} \tau'
\tag{17}
$$

where $\tau = [1 \ 2n \ 3n^2 \ \ldots \ pn^{p-1}]$, and $(C_{\widehat{c}})_{(2:p+1,2:p+1)}$ is obtained from $C_{\widehat{c}}$ by removing its first row and column (indicated by the subsubscript $(2:p+1, 2:p+1)$).

This bound is signal-specific, i.e., it depends on the level-crossing instants that are specific to that signal and depend on the actual IF. It is not general enough to be used across different kinds of IF even if they are of the same polynomial order. To compute the bound accurately also requires exact knowledge of the actual IF, which may not be available in practice. From a practical viewpoint, we can forgo the specificity and exactness of the bound. A general bound, even if approximate, will be practically useful. This can be done by extending the analysis in [32] to yield a large sample approximation on any IF estimator $\widehat{f}[n]$ as

$$
\begin{aligned}
&\mathrm{CRB}\left\{\widehat{f}[n]\right\} \\
&= \frac{1}{4\pi^2} \sum_{i=0}^{p} \sum_{j=0}^{p} \frac{ijn^{i+j-2}\sigma_w^2}{2M^{i+j+1}} \\
&\quad \cdot \left[\frac{1}{i+j+1} + \frac{(p+1)^2}{2M(i+1)(j+1)} - \frac{1}{2M}\right] \\
&\quad \cdot \left[(-1)^{i+j}(p+i+1)(p+j+1)C_i^{p+i} C_i^p C_j^{p+j} C_j^p\right] \\
&\quad 0 \leq n \leq M-1
\end{aligned}
\tag{18}
$$

where M is the observed data size, and p is the order of the IP. $C_j^p = (p!/(j!(p-j)!))$, where $!$ denotes factorial. Simulations show that the above bound is general and can be used to assess the performance of a wide variety of polynomial IF estimators.

B. Harmonic Model

A similar analysis for the LPS model will yield signal specific bounds. An approximate bound can be derived for the LPS model as well. The entries in the Fisher information matrix $\mathcal{J}$ can be computed as follows: $\mathcal{J}_{\alpha,\alpha} = 2T/\sigma_w^2$, $\mathcal{J}_{\beta,\beta} = 2T^3/3\sigma_w^2$, $\mathcal{J}_{\alpha,\beta} = 0$, $\mathcal{J}_{\alpha,\gamma_k} = 0$, $\mathcal{J}_{\beta,\gamma_k} = -2T^2/\pi k \sigma_w^2$, $\mathcal{J}_{\gamma_k,\gamma_m} = (T/\sigma_w^2)\delta_{k,m}$. The approximate bound is given by $\mathrm{CRB}\{\widehat{f}[n]\} = (1/4\pi^2)v\mathcal{J}^{-1}v^T$, where $v = [0 \ 1 \ \omega_0\cos(\omega_0 n) \ 2\omega_0\cos(2\omega_0 n) \ \ldots \ p\omega_0\cos(p\omega_0 n)]$, $0 \leq n \leq M-1$.

APPENDIX II
BIAS-VARIANCE ANALYSIS

In this Appendix, we study the conditions under which the LC-IF estimator becomes unbiased and consistent.

Let the actual model for the IP be given by $\phi(t) = \sum_{k=0}^{p} \alpha_k t^k$. Let the model chosen be $\phi(t) = \sum_{k=0}^{q} c_k t^k$ and its estimate be given by $\widehat{\phi}(t) = \sum_{k=0}^{q} \widehat{c}_k t^k$. If $q = p$, there is no model mismatch; otherwise, there is a mismatch between the actual and the estimated model. The LCs $\{\widehat{t}_{j,\ell}; j = 0, 1, 2, \ldots, N_\ell\}$ are estimated from the noisy signal. Under a high-SNR assumption, we assume that they are perturbed due to noise, i.e., the instants $\widehat{t}_{j,\ell}$ can be written as $\widehat{t}_{j,\ell} = t_{j,\ell} + \delta t_{j,\ell}$, where $\{t_{j,\ell}, j = 0, 1, 2, \ldots, N_\ell\}$ are the level-ℓ-crossing instants of $s_Q(t)$. The coefficients are estimated in a least squares sense by a qth-order polynomial fit to the data sets, $\mathcal{T}^\ell = \{\widehat{t}_{j,\ell}; j = 0, 1, 2, \ldots, N_\ell\}$ and $\{\phi(t_{j,\ell}) = j\pi + (-1)^j \sin^{-1}(\ell); j = 0, 1, 2, \ldots, N_\ell\}$. Using a first-order Taylor series approximation, the problem can be recast to one of fitting a curve to the data sets $\{t_{j,\ell}; j = 0, 1, 2, \ldots, N_\ell\}$ and $\{\phi(t_{j,\ell}) + \phi'(t_{j,\ell})\delta t_{j,\ell}; j = 0, 1, 2, \ldots, N_\ell\}$. Define $\mathbf{\Phi}_\ell$ as a column vector whose jth element is $j\pi + (-1)^j \sin^{-1}(\ell)$ and $\mathbf{\Delta}_{\mathbf{\Phi}_\ell}$ as a column vector whose jth element is $\phi'(t_{j,\ell})\delta t_{j,\ell}$. The estimated optimum q^{th} order coefficient vector can be written as $\widehat{\mathbf{c}} = (\mathbf{H}'_\ell \mathbf{H}_\ell)^{-1}\mathbf{H}'_\ell(\mathbf{\Phi}_\ell + \mathbf{\Delta}_{\mathbf{\Phi}_\ell}) = \mathbf{c} + \mathbf{\Delta}_{\mathbf{c}}$, where $\mathbf{\Delta}_{\mathbf{c}} = (\mathbf{H}'_\ell \mathbf{H}_\ell)^{-1}\mathbf{H}'_\ell \mathbf{\Delta}_{\mathbf{\Phi}_\ell}$. In the absence of noise, $\widehat{\mathbf{c}} = \mathbf{c}$. The estimate of instantaneous phase is given by $\widehat{\phi}[n] = \widehat{\mathbf{c}}'\mathbf{n}$, where $\mathbf{n} = [1 \ n \ n^2 \ \ldots \ n^p]'$ and the estimate of the instantaneous frequency is given by $\widehat{f}[n] = (1/2\pi)\sum_{k=1}^{q} k\widehat{c}_k n^{k-1}$, where $\widehat{c}_k$ is the kth element of $\widehat{\mathbf{c}}$. The bias of the IF estimator, which is denoted by $\mathcal{B}[n]$ is given by

$$
\mathcal{B}[n] = f[n] - \frac{1}{2\pi}\mathcal{E}\left\{(\mathbf{Sc})'\tau + (\mathbf{S}\mathbf{\Delta}_{\mathbf{c}})'\tau\right\}
\tag{19}
$$

$$
= f[n] - \frac{1}{2\pi}(\mathbf{Sc})'\tau
\tag{20}
$$

where $\boldsymbol{\tau} = [1 \ 2n \ 3n^2 \ \ldots \ pn^{p-1}]'$. The matrix $\mathbf{S}$ is obtained by prefixing an all-zero row and column to a $p \times p$ identity matrix. The statistical bias component $\mathcal{E}\{(\mathbf{S}\mathbf{\Delta}_{\mathbf{c}})'\tau\}$ is essentially due to additive noise. The expected value of $\mathbf{\Delta}_{\mathbf{c}} = (\mathbf{H}'_\ell \mathbf{H}_\ell)^{-1}\mathbf{H}'_\ell \mathbf{\Delta}_{\mathbf{\Phi}}$ is zero (see Section III-A and E). The residual component in

bias is the modeling bias that arises mainly due to model order mismatch and is given by

$$\mathcal{B}[n] = \sum_{k=1}^{p} k\alpha_k n^{k-1} - \sum_{k=1}^{q} kc_k n^{k-1}. \qquad (21)$$

If the model order is chosen correctly, i.e., $q = p$, then modeling bias equals zero.

The expression for variance of the LC-IF estimator, which is denoted by $\mathcal{V}[n]$, is given by

$$\mathcal{V}[n] = \mathcal{E}\left\{\left[\hat{f}[n] - \mathcal{E}\left\{\hat{f}[n]\right\}\right]^2\right\} \qquad (22)$$

$$= \frac{1}{4\pi^2}\mathcal{E}\left\{\left[(\mathbf{S}\mathbf{\Delta_c})'\tau\right]^2\right\}. \qquad (23)$$

Simplifying using the results in Section III-A, we get

$$\mathcal{V}[n] \approx \frac{\sigma_w^2}{8\pi^2(1-\ell^2)}\tau'\left\{\mathbf{S}\left(\mathbf{H}_\ell'\mathbf{H}_\ell\right)^{-1}\mathbf{S}\right\}\tau \qquad (24)$$

where σ_w^2 is the variance of noise. As $\sigma_w^2 \to 0$, $\mathcal{V}[n] \to 0$, and $0 \le n \le M - 1$. Thus, the LC-IF estimator is consistent with respect to the increase in SNR.

The variance expression can be further simplified by noting that the (g, h)th element of the matrix $\mathbf{H}_\ell'\mathbf{H}_\ell$ is given by $(\mathbf{H}_\ell'\mathbf{H}_\ell)_{g,h} = \sum_{k=0}^{N_\ell} t_{k,\ell}^{g+h-2}, 1 \le g \le q + 1, 1 \le h \le q + 1$. Since the columns of $\mathbf{H}_\ell$ are linearly independent, the matrix $\mathbf{H}_\ell'\mathbf{H}_\ell$ is always invertible [19]. The matrix entries consist of countably finite sums of unevenly spaced real numbers and closed-form solutions are not possible.

Under the asymptotic condition of large data size (i.e., $M \to \infty$), the matrix $\mathbf{H}_\ell'\mathbf{H}_\ell$ becomes diagonally dominant with entries given by $[\sum_{k=0}^{N_\ell} t_{k,\ell}^0, \sum_{k=0}^{N_\ell} t_{k,\ell}^2, \sum_{k=0}^{N_\ell} t_{k,\ell}^4, \ldots, \sum_{k=0}^{N_\ell} t_{k,\ell}^{2q}]$. Using this large-sample approximation, we can write the asymptotic variance as

$$\mathcal{V}[n] \approx \frac{\sigma_w^2}{8\pi^2(1-\ell^2)}\sum_{m=1}^{q}\frac{1}{\sum_{k=0}^{N_\ell} t_{k,\ell}^{2m}}(mn^{m-1})^2,$$
$$0 \le n \le M - 1. \qquad (25)$$

As $M \to \infty$, $N_\ell \to \infty$, $t_{N_\ell,\ell} \to \infty$, $\mathcal{V}[n] \to 0$. Hence, the LC-IF estimator is also consistent with respect to increase in the data size.

ACKNOWLEDGMENT

REFERENCES

[1] B. Picinbono, "On instantaneous amplitude and phase of signals," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 552–560, Mar. 1997.

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

[3] S. Haykin, *Communication Systems*. New York: Wiley, 1978.

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

[5] L. J. Stankovic, "A time-frequency distribution concentrated along the instantaneous frequency," *IEEE Signal Process. Lett.*, vol. 3, no. 3, pp. 89–91, Mar. 1996.

[6] B. Boashash, Ed., *Time-Frequency Signal Analysis and Processing*. Oxford, U.K.: Elsevier, 2003.

[7] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—part 2: algorithms and applications," *Proc. IEEE*, vol. 80, no. 4, pp. 539–568, Apr. 1992.

[8] D. C. Reed, A. M. Zoubir, and B. Boashash, "Aircraft flight parameter estimation based on passive acoustic techniques using the polynomial Wigner-Ville distribution," *J. Acous. Soc. Amer.*, vol. 102, no. 1, Jul. 1997.

[9] A. K. Barros and N. Ohnishi, "Heart Instantaneous Frequency (HIF): an alternative approach to extract heart rate variability," *IEEE Trans. Biomed. Engg*, vol. 48, no. 8, pp. 850–855, Aug. 2001.

[10] O. Ghitza, "Auditory nerve representations as a front-end for speech recognition in a noisy environment," *Comput. Speech Language*, vol. 1, pp. 109–130, 1986.

[11] ——, "Auditory models and human performance in tasks related to speech coding and speech recognition," *IEEE Trans. Speech Audio Process.*, vol. 2, no. 1, pp. 115–132, Jan. 1994.

[12] D. S. Kim, S. Y. Lee, and R. M. Kil, "Auditory processing of speech signals for robust speech recognition in real-world noisy environments," *IEEE Trans. Speech Audio Process.*, vol. 7, no. 1, pp. 55–69, Jan. 1999.

[13] 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 Process.*, vol. 40, no. 2, pp. 282–293, Feb. 1992.

[14] M. Unser, "Sampling-50 years after Shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, Apr. 2000.

[15] R. Vio, T. Strohmer, and W. Wamsteker, "On the reconstruction of irregularly sampled time series," *Publ. Astr. Soc. Pacific*, vol. 112, pp. 74–90, Jan. 2000.

[16] A. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*, Second ed. Englewood Cliffs, NJ: Prentice-Hall, 2000.

[17] E. V. Krishnamurthy and S. K. Sen, *Numerical Algorithms—Computations in Science and Engineering*, India: Affiliated East-West Press Pvt. Ltd., 2001.

[18] S. C. Sekhar and T. V. Sreenivas, "Instantaneous frequency estimation using level-crossing information," in *Proc. IEEE, Intl. Conf. Acoust., Speech. Signal Process.*, Hong Kong, 2003.

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

[20] S. C. Sekhar and T. V. Sreenivas, "Adaptive window zero-crossing-based instantaneous frequency estimation," *EURASIP J. Applied Signal Process., Special Issue on Non-linear Signal Image Process.*, vol. 12, pp. 1791–1806, Dec. 2004.

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

[22] H. Leib and S. Pasupathy, "The phase of a vector perturbed by Gaussian noise and differentially coherent receivers," *IEEE Trans. Inf. Theory*, vol. 34, no. 6, pp. 1491–1501, Nov. 1988.

[23] A. P. Suppappola, Ed., *Applications in Time-Frequency Signal Processing*. Boca Raton, FL: CRC, 2003.

[24] S. C. Sekhar and T. V. Sreenivas, "Pseudo-time-varying filters for improved instantaneous frequency estimation," *Signal Process.*, submitted for publication.

[25] B. Barkat and B. Boashash, "Instantaneous frequency estimation of polynomial FM signals using the peak of the PWVD: statistical performance in the presence of additive Gaussian noise," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2480–2490, Sep. 1999.

[26] S. Peleg, B. Porat, and B. Friedlander, "The Discrete Polynomial Transform (DPT), its properties and applications," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.*, 1991, pp. 116–119.

[27] S. Peleg and B. Friedlander, "The discrete polynomial-phase transform," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 1901–1914, Aug. 1995.

[28] B. Porat and B. Friedlander, "Asymptotic statistical analysis of the high-order ambiguity function for parameter estimation of polynomial-phase signals," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 995–1001, May 1996.

[29] J. Angeby, "Aliasing of polynomial-phase signal parameters," *IEEE Trans. Signal Process.*, vol. 48, no. 5, pp. 1488–1491, May 2000.

[30] B. Kedem, *Time Series Analysis by Higher Order Crossings*. New York: IEEE , 1994.

[31] J. A. Legg and D. A. Gray, "Performance bounds for polynomial phase parameter estimation with nonuniform and random sampling schemes," *IEEE Trans. Signal Process.*, vol. 48, no. 2, pp. 331–337, Feb. 2000.

[32] S. Peleg and B. Porat, "The Cramer-Rao lower bound for signals with constant amplitude and polynomial phase," *IEEE Trans. Signal Process.*, vol. 39, no. 3, pp. 749–752, Mar. 1991.

S. Chandra Sekhar (S'00) was born in Gollapalem, Andhra Pradesh, India, on August 5, 1976. He received the B.E degree in electronics and communication engineering, with first rank, from Osmania University, Hyderabad, India, in 1999. He is presently pursuing Ph.D degree at the Indian Institute of Science, Bangalore, India. He received the IBM (India Research Laboratory, New Delhi) Research Fellowship for the period 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.

Mr. Sekhar won Osmania University's prestigious Prof. K. K. Nair commemoration Gold medal in 1999 for the outstanding undergraduate student, as well as several distinctions at the Andhra Pradesh state level.

Thippur V. Sreenivas (SM'93) was born in Chintamani, Karnataka, India, in 1954. After graduating from Bangalore University in 1973, he received the M.E. degree from the Indian Institute of Science, Bangalore, India, in 1975 and the Ph.D. degree from the Indian Institute of Technology, Bombay, India, in 1981, while working as a Research Scholar at Tata Institute of Fundamental Research, Bombay, in the area of speech signal processing.

From 1982 to 1985, he was with the Electronics and Radar Development Establishment, Bangalore, as Scientist-C and contributed to low bitrate speech coding systems.. From 1986 to 1987, he was Research Scientist with ELAB, Norwegian Institute of Technology (NTH), Trondheim, Norway, developing new techniques for speech coding and speech recognition. From 1988 to 1989, he was a Visiting Assistant Professor with Marquette University, Milwaukee, WI, teaching and doing research in speech enhancement and spectral estimation. Since 1990, he has been with the faculty of Indian Institute of Science (IISc), Bangalore, where he is currently Associate Professor and leads the activities of the Speech and Audio Lab, where 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.," which is a startup company in Bangalore. He was a visiting faculty at Fraunhofer Institute for Integrated Circuits, Erlangen, Germany, from 1997 to 1998, and Griffith University, Brisbane, Australia, during the summer of 2000.

Dr. Sreenivas served as the chairman of Bangalore Chapter of the IEEE Signal Processing Society during 2003 and 2004.