Statistical AM–FM Models, Extended Kalman Filter Demodulation, Cramér–Rao Bounds, and Speech Analysis

Wan-Chieh Pai and Peter C. Doerschuk, *Member, IEEE*

***Abstract*—A stochastic dynamical system model for describing time signals that are jointly amplitude (AM) and frequency (FM) modulated is presented. The signal is assumed to be bandpass, perhaps originating from a filter bank applied to a broad band signal, and includes the constraint that the magnitude of the complex base band signal is positive. Motivated by speech processing and the desire for narrowband modulating signals, time is divided into frames, and the modulating signals are smoothly interpolated across each frame. The model allows a detailed characterization of the bandwidth of the modulating signals and the statistical character of the measurement noise. An adaptive estimation algorithm based on extended Kalman filtering ideas for extracting the modulating signals from the measured signal is described and demonstrated on both voiced and unvoiced speech signals. The Cramér–Rao bound on the performance of any estimator is computed.**

***Index Terms*—AM–FM representations, Cramér–Rao bound, extended Kalman filter, nonlinear speech models, sinusoidal speech representation, speech processing.**

I. Introduction

Signal models in which slowly varying AM and FM signals modulate a sinusoidal carrier have a long history in signal processing. The simplest such model in discrete time for the measured signal $y(l)$ is

$$y(l) = a(l)\cos\left(2\pi f_c lT + 2\pi T \sum_{m=-\infty}^{l} \nu(m)\right)$$

where $a(\cdot)$ and $\nu(\cdot)$ are the AM and FM signals, respectively, f_c is the carrier, and T is the sampling period. Three important issues are models for the AM and FM signals, algorithms for extracting the AM and FM signals from the measured signal, and bounds on the smallest error that can be achieved in extracting the AM and FM signals from the measured signal.

A variety of models have been used for the AM and FM signals, including deterministic but unknown [12], [13], [17], [20], [24], [29], [30], [38], [42], [47], polynomial [35], [40], [41], linear combination of general basis functions [12], and stochastic [12], [25]–[27], [34], [47]. In this paper, we describe a new stochastic model that combines aspects of polynomial models (there is an interpolation between stochastic endpoints) and stochastic dynamical system models (the endpoints are the output of a stochastic dynamical system). This model is motivated by sinusoidal representation of speech [16], [31], [33] in which speech is broken up into frames , and the speech in each frame is represented as a sum of sinusoids in which the amplitudes, frequencies, and phases of the sinusoids vary in a very simple way, or not at all, during the frame. In our model, one frame corresponds to one interpolation interval, and the simple variation of parameters over the interpolation interval corresponds to the interpolating function.

In some work, the model represents the measured signal itself (e.g., [12], [27], [30], [34], [47]), whereas in other work, the model represents a subband of the measured signal (e.g., [8], [17]). In this paper, the model represents a subband, and the algorithms operate on the corresponding complex baseband signal [37].

A variety of algorithms for extracting the AM and FM signals from the measured signal have been proposed based on different types of models for the AM and FM signals. Examples of algorithms include algorithms based on the Teager energy operator [7], [17], [18], [36], algorithms based on time-frequency transforms [3], [24], [41], and algorithms based on recursive statistical nonlinear filtering [25]–[27], [34]. In this paper, we describe and use an extended Kalman filter [2, Sec. 8.2] algorithm in parallel with a simple adaptive algorithm for estimating certain noise variances. Related algorithms have been described for signals with constant amplitude but time-varying frequency, e.g., [4], [5], [9].

The primary method for bounding the error that can be achieved in extracting the AM and FM signals from the measured signal is the Cramér–Rao bound [12], [13], [20], [35], [40], [47]. Versions of this bound exist for both deterministic but unknown parameters and for statistical parameters [44]. In this paper, we compute this bound for a very general class of models that includes the models we apply to speech. Our computations include the transient behavior of the estimator when the system is initialized or, perhaps, reinitialized at, for example, a voiced-to-unvoiced transition. Similar reinitialization would be useful in images at, for example, an edge in the image.

Whereas AM–FM models have been used in two-dimensional (2-D) problems (e.g., [1], [19], [28]) as well as one-dimensional (1-D) problems, in this paper, we concentrate on 1-D problems. However, the models and algorithms we describe can be generalized to 2-D just like the models and algorithms associated

Manuscript received February 22, 1999; revised January 31, 2000. The associate editor coordinating the review of this paper and approving it for publication was Prof. Dr. Ir. Bart L. R. De Moor.

The authors are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA.

Publisher Item Identifier S 1053-587X(00)05996-1.

with 1-D Kalman filtering can be generalized to 2-D [21], [45], [46] or 1-D communication models and algorithms can be generalized to 2-D [22], [23].

The remainder of the paper is organized in the following fashion. In Section II, we describe the model for the speech signal, in Section III, we present the estimation algorithm, and in Section IV, we apply the model and algorithm to the speech signals. A more general model of the same AM–FM type is described in Section V, and for this more general model, we derive an algorithm to compute the Cramér–Rao bound on the estimation error of the AM and FM modulations in Section VI. In Section VII, we use the algorithm to compute the Cramér-Rao bounds for a complicated example. Finally, Section VIII contains some general remarks and conclusions.

The following notation is used. Let

$\mathbf{I}_n$	$n \times n$ identity matrix;
$\mathbf{0}_{m,n}$	$m \times n$ zero matrix;
$\mathbf{0}_n$	n-dimensional zero column vector;
$\mathrm{E}\{\cdot\}$	expectation;
$\lfloor \cdot \rfloor$	floor function;
$\delta_{m,n}$	Kronecker delta function;
mod	modular operation ($r = k$ mod N satisfies $k = nN + r$ with n an integer and $r \in \{0, 1, \cdots, N-1\}$);
$'$	transpose.

For any sequence $\mathbf{x}(n)$ let $\mathbf{x}_l^m = (\mathbf{x}'(l), \cdots, \mathbf{x}'(m))'$.

II. Model

Using a bank of bandpass filters with center frequencies f_i, we divide the input signal y into subsignals y_i. We then convert each subsignal y_i to its complex baseband representation [37] (using center frequency f_i) with in-phase and quadrature-phase components y_i^I and y_i^Q, respectively. Our statistical model is a model for the y_i^I and y_i^Q signals. It is based on three signals:

1) amplitude a_i;
2) instantaneous frequency ν_i;
3) phase ϕ_i.

In the following description, several downsampling steps and the subsignal index i are omitted in order to simplify the notation.

We model the signals y^I and y^Q as the output of a dynamical system. The dynamical system can be described by sample number or indexed by frame number and within-frame sample number. Both descriptions have advantages and disadvantages, and in the following, both descriptions are given, starting with the description indexed by sample number. The measurement equation is straightforward.

$$\begin{bmatrix} y^I(l) \\ y^Q(l) \end{bmatrix} = \begin{bmatrix} a(l)\cos(\phi(l)) \\ a(l)\sin(\phi(l)) \end{bmatrix} + \begin{bmatrix} rv^I(l) \\ rv^Q(l) \end{bmatrix} \tag{1}$$

where $v(\cdot) = [v^I(\cdot), v^Q(\cdot)]'$ is a zero-mean white Gaussian sequence with identity covariance matrix, and r is a known positive constant. The state equation for the phase ϕ follows directly from its definition:

$$\phi(l+1) = \phi(l) + 2\pi T \nu(l) \tag{2}$$

where ν is the instantaneous frequency, and T is the sampling interval. In order to reduce the bandwidth of the estimated a and ν signals, we use an unusual dynamical system to describe a and ν. Both a and ν are modeled as the result of linear interpolation between stochastic endpoints separated by a fixed interval of N samples under the constraint that the function is continuous at the endpoints. In the context of speech signals, the N sample window is exactly the traditional idea of a frame. As in speech signal processing, the user must decide on an appropriate value of N. Picking N too small tends to increase the bandwidth of the a and ν signals, whereas picking N too large tends to lead to less accurate decompositions of the original signal. We could easily replace the linear interpolation by a higher order spline but have not found that to be necessary for the speech applications in which we are interested. In addition, instead of defining a to be the linear interpolation between the endpoints $a^{\pm}$, we define a to be the linear interpolation between the endpoints $[a^{\pm}]^2$. When using a linear interpolation (although not necessarily when using a higher order interpolation), this guarantees that the amplitude is always non-negative. The computational and theoretical simplicity and the ability to enforce positivity are the primary reasons for choosing a linear interpolation method. Assume that the endpoints of both the a and ν interpolations form a random walk. With these definitions and assumptions, the four state equations that define the endpoints for the linear interpolations of a and ν are

$$\begin{aligned} a^+(l+1) &= a^+(l) + q_a w_a(l) \delta_{l \bmod N, 0} \\ a^-(l+1) &= a^-(l) + [a^+(l) - a^-(l)] \delta_{l \bmod N, 0} \\ \nu^+(l+1) &= \nu^+(l) + q_\nu w_\nu(l) \delta_{l \bmod N, 0} \\ \nu^-(l+1) &= \nu^-(l) + [\nu^+(l) - \nu^-(l)] \delta_{l \bmod N, 0} \end{aligned} \tag{3}$$

where $w_a(\cdot)$ and $w_\nu(\cdot)$ are independent zero-mean white Gaussian sequences with variances 1, and q_a and q_ν are positive constants. The linear interpolations actually occur in the ϕ state equation (2) and the measurement equation (1) by use of the definitions

$$\begin{aligned} \nu(l) &= \nu^-(l) + [\nu^+(l) - \nu^-(l)][((l+N-1) \bmod N) + 1]/N \\ a(l) &= [a^-(l)]^2 + \{[a^+(l)]^2 - [a^-(l)]^2\} \\ &\quad \cdot [((l+N-1) \bmod N) + 1]/N. \end{aligned} \tag{4}$$

Because of the mod operations, the system is time varying.

We now describe the same dynamical system again but indexed by frame number and within-frame sample number rather than by sample number. This simplifies the interpolation equations substantially but makes the measurement equation and state equation for ϕ more complex. Let $k \in \{0, 1, \cdots\}$ index the frames and $n \in \{1, \cdots, N\}$ index the within-frame samples. Therefore, $l = kN + n$. Replace the notation $a(l)$ by $a_{k,n}$ and likewise for the other variables. The advantage of the k, n notation is that $a^{\pm}$ and $\nu^{\pm}$ depend only on k and that the system is time invariant. The state equation that replaces (3) is

$$\begin{bmatrix} a^+(k+1) \\ a^-(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a^+(k) \\ a^-(k) \end{bmatrix} + \begin{bmatrix} q_a \\ 0 \end{bmatrix} w_a(k)$$

and likewise for $\nu^{\pm}$, which is easier to understand than (3) and, unlike (3), is time invariant. However, the measurement equation and state equation for ϕ are more complicated than (1) and (2). Define the interpolating function $g_n(\cdot,\cdot)$ by

$$g_n(u_1, u_0) = u_0 + (u_1 - u_0)n/N \tag{5}$$

where $n \in \{1, \cdots, N\}$. Then, $a_{k,n} = g_n([a^+(k)]^2, [a^-(k)]^2)$ and $\nu_{k,n} = g_n(\nu^+(k), \nu^-(k))$. The phase at the nth sample in the kth frame, which is denoted by $\phi_{k,n}$, is the phase at the end of the previous frame (i.e., $\phi_{k-1,N}$) plus $2\pi T$ times the sum of the interpolated frequencies (i.e., $\nu_{k,n'}$) at the first through nth points in the frame

$$\begin{aligned}\phi_{k,n} &= \phi_{k-1,N} + 2\pi T \sum_{n'=1}^{n} \nu_{k,n'}\\ &= \phi_{k-1,N} + \gamma_n \nu^+(k) + \delta_n \nu^-(k)\end{aligned}$$

where

$$\gamma_n = 2\pi T n(n+1)/(2N) \tag{6}$$

$$\delta_n = 2\pi T n[1 - (n+1)/(2N)] \tag{7}$$

and we have used the identity $\sum_{k=1}^{K} k = K(K+1)/2$. Therefore, it is possible to propagate $\phi_{k,N}$, which is independent of n and is denoted by ϕ_k, by

$$\phi_k = \phi_{k-1} + \gamma_n \nu^+(k) + \delta_n \nu^-(k)$$

which replaces (2). Finally, the aggregated measurement, which includes all N in-phase and all N quadrature-phase samples in the frame, is

$$\begin{aligned}\mathbf{y}(k) = &[a_{k,1}\cos(\phi_{k,1}), a_{k,1}\sin(\phi_{k,1}), \cdots, a_{k,N}\cos(\phi_{k,N})\\ &a_{k,N}\sin(\phi_{k,N})]' + r\mathbf{v}(k)\end{aligned}$$

where $\mathbf{v}(\cdot)$ is a $2N$-dimensional white zero-mean Gaussian noise process with an identity covariance matrix which replaces (1).

By shifting to complex baseband processing of bandpass subsignals, algorithms for the model described here operate at a lower sampling rate than those of [27], and furthermore, the format frequency variables of [27] are no longer required. In addition, the introduction of the interpolation interval and function in the model described here, neither of which are present in [27], makes contact with the ideas of frames and sinusoidal coders in speech and achieves compression of the information in the AM and FM modulating signals. Finally, the fact that the amplitude in the measurement equation is constrained to be always positive reduces the ambiguity in the separation of AM and FM and was not used in [27].

III. Statistical Estimation

In terms of the model, the goal of extracting AM and FM modulations from the observed signal corresponds to estimating $z(l) = (a^+(l), a^-(l), \nu^+(l), \nu^-(l), \phi(l))'$ from $y^I(0)$, $y^Q(0), \cdots, y^I(l)$, $y^Q(l)$. Define $\mathbf{y}(l) = [y^I(l), y^Q(l)]'$. Let $\hat{z}(l|m)$, which is a function of $\mathbf{y}(0), \cdots, \mathbf{y}(m)$, be the estimate of $z(l)$ based on data through time m. Let $\epsilon(l|m) \doteq \mathrm{E}[(z(l) - \hat{z}(l|m))'(z(l) - \hat{z}(l|m))]$ be the mean squared error (MSE). We define as the optimal estimator, which is denoted by $\hat{z}^*(l|m)$, that estimator that minimizes $\epsilon(l|m)$ with the result that $\hat{z}^*(l|m) = \mathrm{E}[z(l)|\mathbf{y}(0), \cdots, \mathbf{y}(m)]$, and the achieved MSE is $\epsilon^*(l|m) = \mathrm{E}[(z(l) - \hat{z}^*(l|m))'(z(l) - \hat{z}^*(l|m))]$.

It is difficult to compute $\hat{z}^*(l|l)$ exactly. However, if $a(\cdot)$ was constant, then (1) is the measurement equation of a frequency modulated communication system, the extended Kalman filter (EKF) [2, Sec. 8.2][1] is essentially a phase-locked loop (PLL), and the PLL is an excellent estimator. Therefore, we compute the estimates $\hat{a}^{\pm}(l|l)$, $\hat{\nu}^{\pm}(l|l)$, and $\hat{\phi}(l|l)$ by using the EKF for this more complicated model. The computational requirements are minimal; the state equation is already linear, the one-step state transition matrix (denoted by F) is sparse so that multiplication by F is inexpensive, and the observation is a 2-D vector so the one matrix inversion is only of a 2×2 matrix.

In dynamical systems such as (3), the power in the signals is determined by the variance of the driving noise. Therefore, in order to process signals whose power varies over time, it is necessary to allow time variation in the process and measurement noise variances. Introducing dynamical system models for these variances is not simple because the variances should be varying much more slowly than the other dynamical variables (such as a^+), and furthermore, such additional dynamical models greatly complicate the estimator because they add a substantial degree of nonlinearity to the state equations that are otherwise linear [2]. Therefore, we have instead chosen to estimate the variances on-line and use the estimated values of the variances in the estimator for $a^{\pm}$, $\nu^{\pm}$, and ϕ as if the estimated values were the true values. For the speech applications in which we are interested, we have achieved adequate performance by allowing q_a^2 and r^2 to vary while keeping q_ν^2 fixed. Because q_a and r strongly influence the power in the measurements $\mathbf{y}(\cdot)$, we estimate them on-line by lowpass filtering the instantaneous measurement power signal with a one-pole filter

$$\begin{aligned}q_a^4(l+1) &= \alpha_{q_a} q_a^4(l) + \beta_{q_a}([y^I(l)]^2 + [y^Q(l)]^2)\\ r^2(l+1) &= \alpha_r r^2(l) + \beta_r([y^I(l)]^2 + [y^Q(l)]^2)\end{aligned} \tag{8}$$

where $q_a(\cdot)$ is raised to the fourth power because $a(\cdot)$ is raised to the second power in the signal and, therefore, to the fourth power in the instantaneous power. The time variation of the basic model (1)–(4) and the additional time variation introduced by the adaptation (8) does not add to the computation burden because the EKF already must linearize around the current state estimates at each time l.

The result of the EKF are the estimates $\hat{a}^{\pm}(l|l)$, $\hat{\nu}^{\pm}(l|l)$, and $\hat{\phi}(l|l)$. From these estimates, we can compute reconstructed in-phase and quadrature-phase signals, which are denoted by $\hat{y}^I(l)$ and $\hat{y}^Q(l)$, by $\hat{y}^I(l) = \hat{a}(l)\cos(\hat{\psi}(l|l))$ and $\hat{y}^Q(l) = \hat{a}(l)\sin(\hat{\psi}(l|l))$, where $\hat{a}(l)$ and $\hat{\psi}(l)$ are reconstructions of $a(l)$ and $\phi(l)$, respectively, and we also desire a reconstruction of $\nu(l)$ that is denoted by $\hat{\nu}(l)$. First, consider

[1]The EKF is a recursive estimator for the state of a finite dimensional nonlinear system. The approach is to linearize the nonlinear system and then apply the Kalman filter to the resulting linear system. The key point is how to determine the nominal state around which the linearization is performed. The choice made in the EKF is to linearize on-line around the current estimate of the state (for the state equation) and the current one-step-ahead prediction of the state (for the observation equation).

$\hat{\nu}(l)$. Let $n \in \{1, \cdots, N\}$. From (3) and (4), it follows that

$$\nu(kN+n) = \nu^-(kN+n) + [\nu^+(kN+n) - \nu^-(kN+n)]n/N \quad (9)$$

and that $\nu^-(kN+n) = \nu^-((k+1)N)$ and $\nu^+(kN+n) = \nu^+((k+1)N) = \nu^-((k+2)N)$. Using more data to construct the estimate of a random variable provides a more accurate estimate. However, the only estimates that we can extract from the EKF (in the absence of state augmentation) are $\hat{\nu}^{\pm}(l|l)$. Therefore, rather than using $\hat{\nu}^-(kN+n|kN+n)$, we use

$$\begin{aligned} E[\nu^-(kN+n)|\mathbf{y}(0), \cdots, \mathbf{y}((k+1)N)] &= E[\nu^-((k+1)N)|\mathbf{y}(0), \cdots, \mathbf{y}((k+1)N)] \\ &= \hat{\nu}^-((k+1)N|(k+1)N) \end{aligned}$$

and similarly for the other two terms of (9). Define $\xi(k) = \hat{\nu}^-((k+1)N|(k+1)N)$. The final estimate is then $\hat{\nu}(kN+n) = \xi(k) + [\xi(k+1) - \xi(k)]n/N$.

The situation for $\hat{a}(l)$ is very similar. We need to estimate $[a^{\pm}(l)]^2$. In theory, we could use estimates of the type $\Sigma_{a^\pm}(l|l) + [\hat{a}^{\pm}(l|l)]^2$, where $\Sigma_{a^\pm}(l|l)$ is the covariance from the EKF. However, we have achieved better performance by dropping the first term and using estimates of the form $[\hat{a}^{\pm}(l|l)]^2$. Define $b(k) = [\hat{a}^-((k+1)N|(k+1)N)]^2$. Then, exactly as for ν, the final estimate is $\hat{a}(kN+n) = b(k) + [b(k+1) - b(k)]n/N$.

Finally, consider $\hat{\psi}(l|l)$. We could define $\hat{\psi}(l|l) = \hat{\phi}(l|l)$ since $\hat{\phi}(l|l)$ is an EKF output. However, the EKF does not constrain the estimates it computes to satisfy the condition that phase (ϕ) is the integral of instantaneous frequency (ν). In the speech applications of interest to us, it has been possible to achieve better performance by at least locally (i.e., within one frame) maintaining the fact that ϕ is the integral of ν. Therefore, at the beginning of each frame, we set $\hat{\psi}$ equal to $\hat{\phi}$, and to compute the remaining samples in the frame, we integrate $\hat{\nu}$. Define $\chi(k) = \hat{\phi}(kN+1|kN+1)$. The estimate is

$$\hat{\psi}(kN+n) = \chi(k) + \sum_{n'=1}^{n-1} \hat{\nu}(kN+n').$$

The adaptation of the variances and the need to relate $\hat{a}^{\pm}(l|l)$, $\hat{\nu}^{\pm}(l|l)$, and $\hat{\phi}(l|l)$ to estimates of the interpolated amplitudes, frequencies, and phases are the major differences between the estimator described here and the estimator of [27]. For ease of reference, we refer to the adaptive extended Kalman filter for the AM–FM model of the output of the filter bank (see Section II) as the AEKF algorithm.

IV. Representation of Speech Signals

In this section, we demonstrate the result of applying the AEKF algorithm to speech data and compare the results with three alternative algorithms. The alternatives are the DESA-1 algorithm [29, eqs. (107)-(108)], which is a single-formant version of the MBDA algorithm [27], and the classical bank-of-filters algorithm [39, Sec. 3.2]. In the bank-of-filters algorithm the signal is passed through a nonlinearity and then a lowpass filter to provide an estimate of the instantaneous amplitude. We use the absolute value nonlinearity and a low pass filter with bandwidth equal to the signal bandwidth. Note that this algorithm does not provide an estimate of the instantaneous frequency.

The data is the sentence "Tim takes Sheila to see movies twice a week." from the TIMIT database [11, /timit/train/dr1/fcjf0/sx397.wav,]. Note that this sentence includes both voiced and unvoiced intervals, unlike the speech used in [27], which is purely voiced. The bandpass filter used to generate the subsignal used here has a passband of 2080–2200 Hz.

There are several differences between the results of the AEKF (Fig. 1) and MBDA (Fig. 2) algorithms. The amplitude estimates from AEKF are always positive, whereas those of MBDA can be negative. This is important since the complex baseband approach implies that the amplitude is positive. While taking the absolute value of the MBDA amplitude (see Fig. 3) gives a signal similar to the AEKF amplitude, it is not clear how to make the corresponding adjustment in the MBDA instantaneous frequency estimate so that the reconstruction is still accurate and the bandwidth of the instantaneous frequency estimate is small. Define the normalized mean squared error for the in-phase signal by $e^I = \left[\left\{\sum_k [y^I(k) - \hat{y}^I(k)]^2\right\} / \left\{\sum_k [y^I(k)]^2\right\}\right]^{1/2}$ and likewise for e^Q. The AEKF algorithm provides estimates that result in reconstructions with better mean squared error than does the MBDA algorithm: $e^I = 0.1481$ and $e^Q = 0.0699$ versus $e^I = 0.5780$ and $e^Q = 0.1070$.

The major difference between the results of the AEKF (Fig. 1) and DESA-1 (Fig. 3) algorithms is that the DESA-1 estimate of instantaneous phase, even after post processing with an 11 point median filter [29, Figs. 5–7], is noisier. One reason for this may be that the 2080–2200 Hz frequency band in this sentence is primarily fricative in character rather than being dominated by a formant.

Finally, comparing the results of AEKF, MBDA, and DESA-1 algorithms with the result of the classical bank-of-filters algorithm (Fig. 3), we see that the classical algorithm gives a similar result, which is an encouraging check. Note, however, that the classical algorithm does not produce an instantaneous frequency estimate. In summary, the AEKF algorithm produces estimates that, at least in some circumstances, seem superior to some other algorithms without dramatically increased computation.

V. General Model

In the speech application (see Section II), we use a simple dynamic stochastic model for the interpolation endpoints $a^{\pm}(k)$ and $\nu^{\pm}(k)$, specifically, a first-order AR process. In this section, we describe a general model where the endpoints are determined by ARMA processes. In Section VI, we compute the Cramér–Rao bound for this general model.

Unlike the simple model in Section II, the general model is described only in terms of the frame number k and the within-frame sample number $n \in \{1, \cdots, N\}$ so that the model is time invariant which is convenient for the Cramér–Rao bound calculations (see Section VI).

The interpolated values of the amplitude and frequency at the nth sample ($n \in \{1, \cdots, N\}$) in the kth frame, which is denoted by $a_{k,n}$ and $\nu_{k,n}$, respectively, are given by $a_{k,n} =$

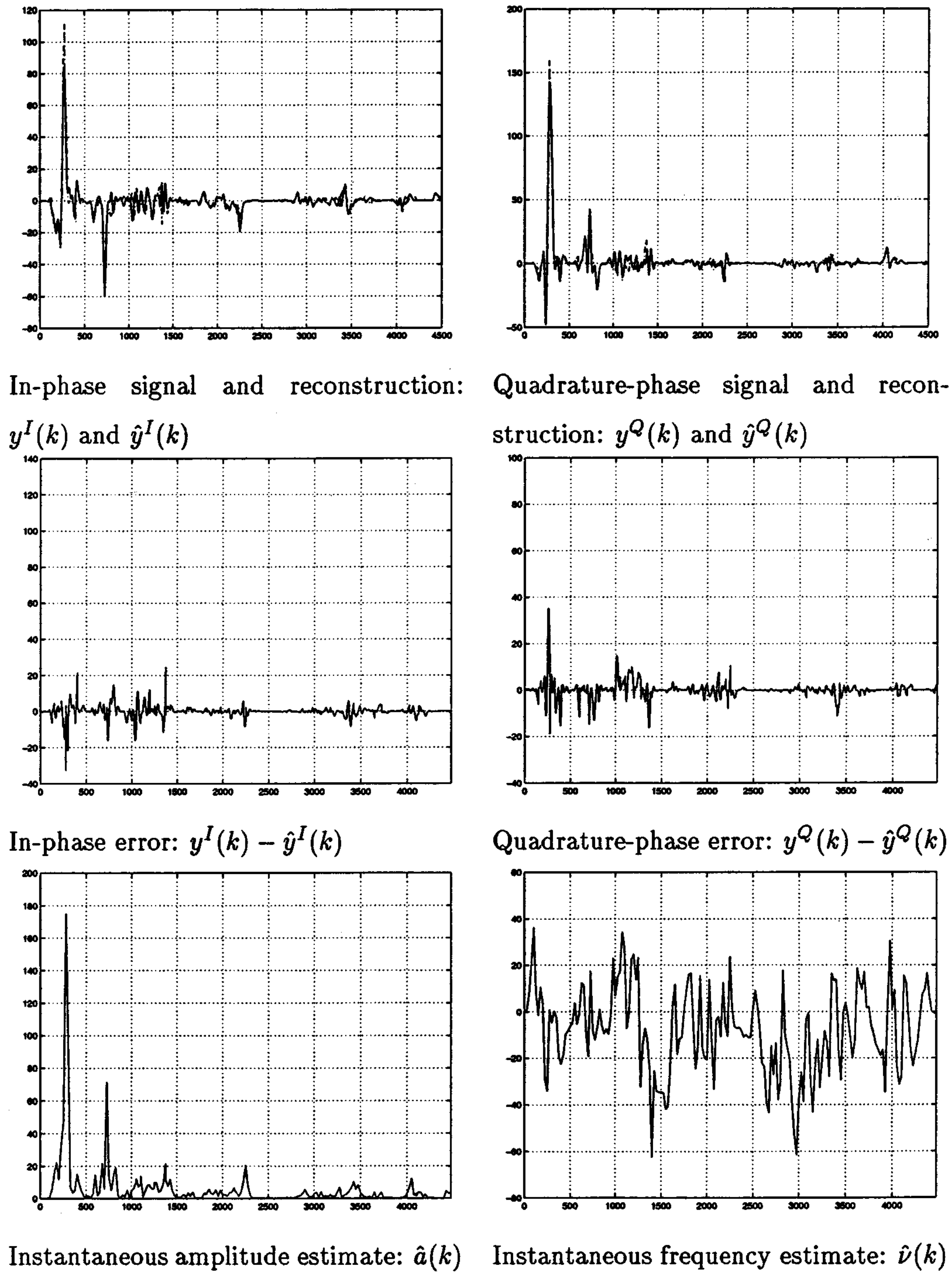

Fig. 1. Results from the AEKF algorithm. In the first pair of figures, the original and reconstructed signals are shown with solid and dashed curves, respectively.

$g_n(a^2(k), a^2(k-1))$ and $\nu_{k,n} = g_n(\nu(k), \nu(k-1))$, where g was defined in (5). The phase at the nth sample in the kth frame, which is denoted by $\phi_{k,n}$, is the phase at the end of the previous frame (i.e., $\phi_{k-1,N}$) plus $2\pi T$ times the sum of the interpolated frequencies (i.e., $\nu_{k,n'}$) at the first through nth points in the frame

$$\begin{aligned}\phi_{k,n} &= \phi_{k-1,N} + 2\pi T \sum_{n'=1}^{n} \nu_{k,n'} \\ &= \phi_{k-1,N} + \gamma_n \nu(k) + \delta_n \nu(k-1)\end{aligned}$$

where γ and δ were defined in (6) and (7). Finally, the aggregated measurement before the addition of noise is

$$\begin{aligned}&[a_{k,1}\cos(\phi_{k,1}), a_{k,1}\sin(\phi_{k,1}), \cdots, a_{k,N}\cos(\phi_{k,N}) \\ &a_{k,N}\sin(\phi_{k,N})]'.\end{aligned}$$

In order for the pairs (i.e., in-phase and quadrature-phase) of the $2N$ components of the measurement vector to have the same $a_{k,n}$ and $\phi_{k,n}$ values in the following equations, the floor function is used in the index of g, γ, and δ. The complete model

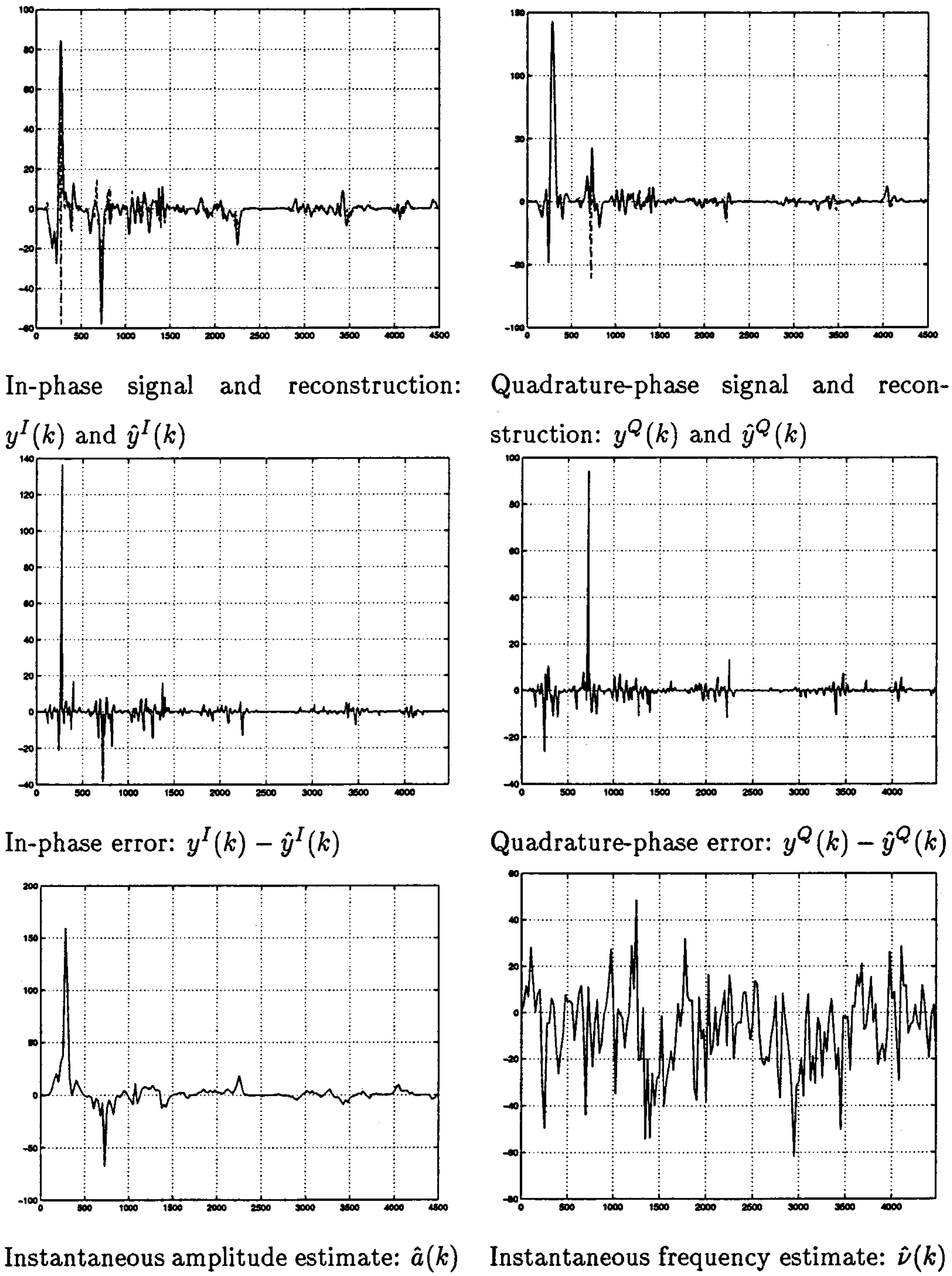

Fig. 2. Results from the MBDA algorithm. In the first pair of figures, the original and reconstructed signals are shown with solid and dashed curves, respectively.

is ($k \geq 1$)

$$\sum_{i=0}^{p_a} c_a(i)a(k-i) = \sum_{i=0}^{q_a} d_a(i)w_a(k-i) \tag{10}$$

$$\sum_{i=0}^{p_\nu} c_\nu(i)\nu(k-i) = \sum_{i=0}^{q_\nu} d_\nu(i)w_\nu(k-i) \tag{11}$$

$$\phi(k) = \phi(k-1) + \gamma_N \nu(k) + \delta_N \nu(k-1) \tag{12}$$

$$\begin{aligned}\mathbf{y}(k) &= \mathbf{s}_k(a(k), a(k-1), \phi(k), \nu(k), \nu(k-1)) \\ &\quad + \mathbf{r}\mathbf{v}(k)\end{aligned} \tag{13}$$

$$\mathbf{s}_k(\cdot) = (s_{k,1}(\cdot), \cdots, s_{k,2N}(\cdot))' \tag{14}$$

$$\begin{aligned}&s_{k,n}(a_1, a_0, \phi, \nu_1, \nu_0) \\ &\quad = g_{\lfloor (n+1)/2 \rfloor}(a_1^2, a_0^2) \\ &\quad\quad \cdot \begin{matrix}\cos \\ \sin\end{matrix} \left(\phi + \gamma_{\lfloor (n+1)/2 \rfloor}\nu_1 + \delta_{\lfloor (n+1)/2 \rfloor}\nu_0\right), \quad \begin{matrix} n \text{ odd} \\ n \text{ even}\end{matrix}\end{aligned} \tag{15}$$

where $w_a(\cdot) \in \mathcal{R}$, $w_\nu(\cdot) \in \mathcal{R}$, and $\mathbf{v}(\cdot) \in \mathcal{R}^{2N}$ are independent zero-mean white Gaussian stochastic processes with covariances 1, 1, and $\mathbf{I}_{2N}$, respectively; $\mathbf{r} \in \mathcal{R}^{2N \times 2N}$, $\mathbf{r} = r\mathbf{I}_{2N}$, and $c_a(0) = 1$, $c_\nu(0) = 1$. The initial conditions on

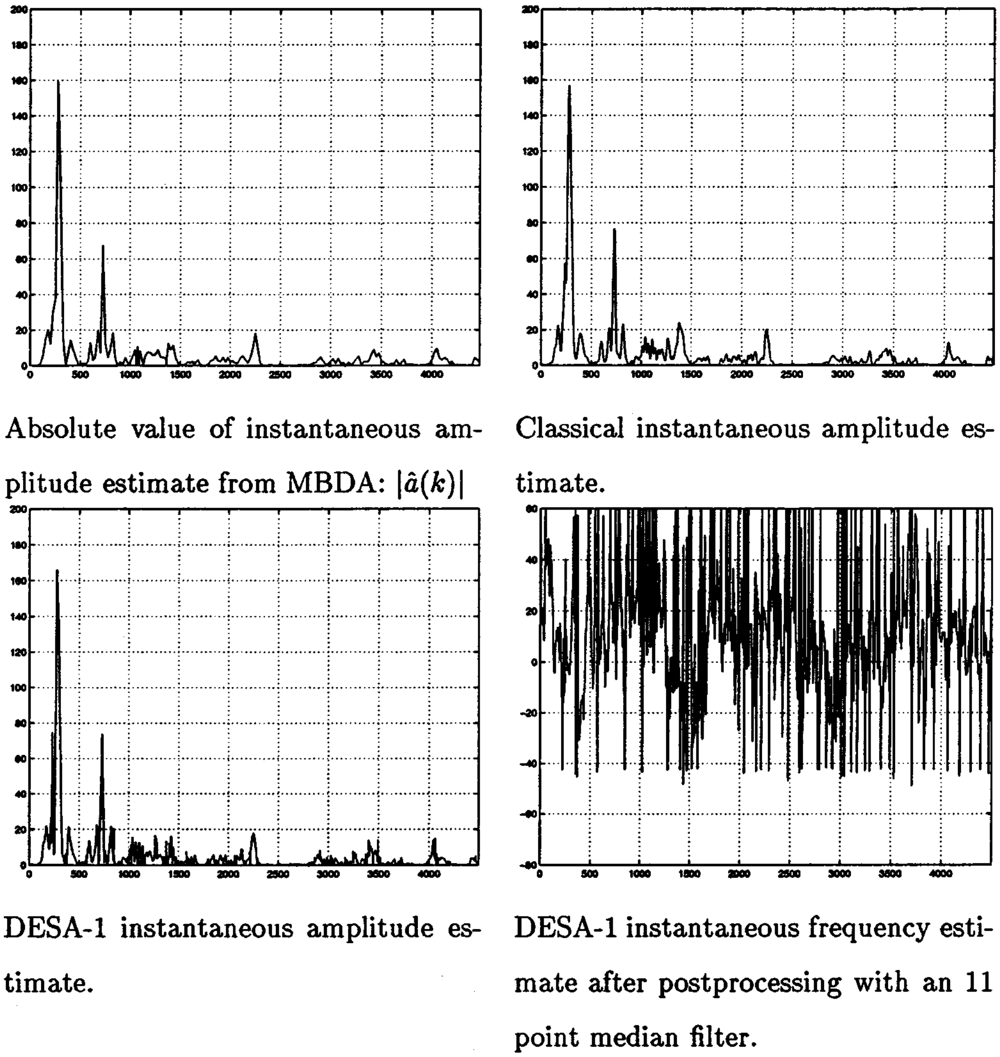

Fig. 3. Results from the MBDA, DESA-1, and Classical algorithms.

the $a(\cdot)$ process, the initial conditions on the joint $[\nu(\cdot), \phi(\cdot)]'$ process, $w_a(\cdot)$, and $w_\nu(\cdot)$ are independent of each other and jointly Gaussian. Precise specification of the initial conditions on the $a(\cdot)$ process (mean $\mathbf{m}_0^{(a)}$ and covariance $\mathbf{P}_0^{(a)}$) and jointly on the $\nu(\cdot)$ and $\phi(\cdot)$ processes (mean $\mathbf{m}_0^{(f)}$ and covariance $\mathbf{P}_0^{(f)}$) is done via state variable representations in Appendix A.

VI. Cramér–Rao Bound

Let $\mathbf{x}(\cdot)$ be a vector nonlinear AR process and $\mathbf{y}(\cdot)$ be a sequence of vector measurements. Application of the standard Cramér-Rao bound [6], [10], [14], [15], [32], [43], [44] to the entire trajectory $\mathbf{x}_{1-p}^K$ of the nonlinear AR process gives the basic bound used in [6], [10], and [14]. Define the trajectory error covariance $\Lambda \in \mathcal{R}^{(K+p)n\times(K+p)n}$ with blocks $\Lambda_{l,k} \in \mathcal{R}^{n\times n}$ given by $\Lambda_{l,k} = E\{(\mathbf{x}_l - \hat{\mathbf{x}}_l^*)(\mathbf{x}_k - \hat{\mathbf{x}}_k^*)'\}$. Define the Fisher information matrix $\mathbf{J}$ by, under appropriate regularity assumptions, $\mathbf{J} = E\{\nabla_{\mathbf{x}_{1-p}^K} \nabla_{\mathbf{x}_{1-p}^K} \ln p(\mathbf{y}_{1-p}^K, \mathbf{x}_{1-p}^K)\}$, where p is the joint pdf on the $\mathbf{x}$ and $\mathbf{y}$ trajectories. Then, since $\hat{\mathbf{x}}_k^*$ is unbiased, the standard multivariate CRB is $\Lambda - \mathbf{J}^{-1} \geq 0$, where $\geq$ means positive semidefinite. The estimation error at time K is $\Lambda_{K,K}$; therefore, the desired CRB is

$$\Lambda_{K,K} - [0 \quad \mathbf{I}_n]\, \mathbf{J}^{-1} \begin{bmatrix} 0 \\ \mathbf{I}_n \end{bmatrix} \geq 0. \tag{16}$$

There are two difficulties in applying (16). The first difficulty is that for a long trajectory, the matrix $\mathbf{J}$ is large, and it is difficult to compute $\mathbf{J}^{-1}$. This problem is circumvented in [6], [10], and [14] by finding a linear Gaussian system that has the same Fisher information matrix as the nonlinear system of interest. (An alternative method is described in [43]). The linear Gaussian system used here and in [10] is a pth-order AR process and an observation equation where the observation at time k depends on the AR process at times $k, k-1, \cdots, k-p+1$. This system can be transformed by state augmentation into a linear Gaussian state-variable system. In the state-variable system, $\Lambda_{K,K}^{\text{state}}$ can be computed exactly and without excessive computation by the Kalman filter, and furthermore, the CRB is satisfied with equality so that the known value of $\Lambda_{K,K}^{\text{state}}$ is the desired bound on the performance of any estimator for the nonlinear system. The second difficulty is that the computation of $\mathbf{J}$ requires computation of expectations. The approach of [6], [10], and [14] still requires such computations, but the approach organizes the computations so that they involve only the nonlinear AR process (i.e., not also the observation equation). In the problem considered here, the nonlinear AR process is actually linear so that these computations can be done analytically.

In order to use the methods of [10] to compute the Cramér–Rao bound, it is necessary to replace (10)–(12) by a vector autoregression process with a full-rank process noise covariance and make related changes in (13). Specifically, it is necessary to write the system in the form

$$\mathbf{x}(k+1) = \mathbf{f}_k(\mathbf{x}(k), \cdots, \mathbf{x}(k-p+1)) + \mathbf{q}\mathbf{w}(k) \tag{17}$$

$$\mathbf{y}(k) = \mathbf{h}_k(\mathbf{x}(k), \cdots, \mathbf{x}(\max(k-p+1, 1-p))) + \mathbf{r}\mathbf{v}(k) \tag{18}$$

where (17) holds for $k \in \{0, \cdots, K-1\}$, (18) holds for $k \in \{1-p, \cdots, K\}$, $\mathbf{q}\mathbf{q}'$, and $\mathbf{r}\mathbf{r}'$ are full rank; $\mathbf{w}(\cdot)$, $\mathbf{v}(\cdot)$, and the initial condition $\mathbf{x}_0 = (\mathbf{x}'(0), \cdots, \mathbf{x}'(1-p))'$ are independent and Gaussian with means 0, 0, and $\mathbf{m}_0^{\mathbf{x}}$, respectively, and covariances $\mathbf{I}$, $\mathbf{I}$, and $\mathbf{P}_0^{\mathbf{x}}$, respectively. Then, the linear Gaussian system that is used for comparison is a special case of the nonlinear system with

$$\mathbf{f}_k(\mathbf{x}(k), \cdots, \mathbf{x}(k-p+1)) = \sum_{i=0}^{p-1} \mathbf{A}_{k,i}\mathbf{x}_{k-i} \tag{19}$$

$$\mathbf{h}_k(\mathbf{x}(k), \cdots, \mathbf{x}(\max(k-p+1, 1-p))) = \sum_{i=0}^{\min(p-1, k+p-1)} \mathbf{C}_{k,i}\mathbf{x}_{k-i} \tag{20}$$

and $\mathbf{m}_0^{\mathbf{x}} = 0$, where $\mathbf{A}_{k,i}$, $\mathbf{C}_{k,i}$, $\mathbf{Q} = \mathbf{q}\mathbf{q}'$, $\mathbf{R} = \mathbf{r}\mathbf{r}'$, and $\mathbf{P}_0^{\mathbf{x}}$ must be determined. In the problem described in this paper, we will find that the dimensions of $\mathbf{x}$ and $\mathbf{y}$ in (17) and (18) are 2 and $2N$, respectively. The theory in [10] requires these dimensions to be equal; therefore, we augment $\mathbf{x}$ with unobserved components. The comparison system (19) and (20) that results from the theory of [10] then has $\mathbf{x}$ and $\mathbf{y}$, both of dimension $2N$. However, only two of the $\mathbf{x}$ components are observable, and only two of the $\mathbf{y}$ components depend on $\mathbf{x}$. Therefore, we can reduce the dimensions of both $\mathbf{x}$ and $\mathbf{y}$ to 2, and it is this smaller system that we describe in this paper.

A. Transformation to Standard Form

We transform (10)–(15) to the form of (17) and (18) in four steps. The first step is to rewrite the ARMA $a(\cdot)$ process as an AR process, which is denoted by $\alpha(\cdot)$ and defined by

$$\sum_{i=0}^{p_a} c_a(i)\alpha(k-i) = w_a(k) \tag{21}$$

which drives an MA process that is defined by

$$a(k) = \sum_{i=0}^{q_a} d_a(i)\alpha(k-i) \tag{22}$$

which can be incorporated into the measurement equation (18). This transformation was already used in Appendix A, and the initial condition on $a(\cdot)$ is given as an initial condition on $\alpha(\cdot)$.

The second step is to define a scalar AR process $\beta(\cdot)$ from which both $\nu(\cdot)$ and $\phi(\cdot)$ can be reconstructed by using $\beta(\cdot)$ to drive two MA processes. This can be done by expressing the transfer function from $w_\nu(\cdot)$ to $\nu(\cdot)$ and $\phi(\cdot)$ in the form

$$\breve{\beta}(z)/\breve{w}_\nu(z) = 1\Big/\left[(1-z^{-1})\sum_{i=0}^{p_\nu} c_\nu(i)z^{-i}\right]$$

$$\breve{\nu}(z)/\breve{\beta}(z) = (1-z^{-1})\sum_{i=0}^{q_\nu} d_\nu(i)z^{-i}$$

$$\breve{\phi}(z)/\breve{\beta}(z) = (\gamma_N + \delta_N z^{-1})\sum_{i=0}^{q_\nu} d_\nu(i)z^{-i}$$

where $\breve{\cdot}$ indicates z transform. Multiplying out the polynomials leads to the definitions

$$c_\beta(i) = \begin{cases} c_\nu(0), & i=0 \\ c_\nu(i) - c_\nu(i-1), & 1 \le i \le p_\nu \\ -c_\nu(p_\nu), & i = p_\nu + 1 \\ 0, & \text{otherwise} \end{cases} \tag{23}$$

$$d_{\beta,\nu}(i) = \begin{cases} d_\nu(0), & i=0 \\ d_\nu(i) - d_\nu(i-1), & 1 \le i \le q_\nu \\ -d_\nu(q_\nu), & i = q_\nu + 1 \\ 0, & \text{otherwise} \end{cases} \tag{24}$$

$$d_{\beta,\phi}(i) = \begin{cases} \gamma_N d_\nu(0), & i=0 \\ \gamma_N d_\nu(i) + \delta_N d_\nu(i-1), & 1 \le i \le q_\nu \\ \delta_N d_\nu(q_\nu), & i = q_\nu + 1 \\ 0, & \text{otherwise} \end{cases} \tag{25}$$

and the time domain equations

$$\sum_{i=0}^{p_\nu+1} c_\beta(i)\beta(k-i) = w_\nu(k) \tag{26}$$

$$\nu(k) = \sum_{i=0}^{q_\nu+1} d_{\beta,\nu}(i)\beta(k-i) \tag{27}$$

$$\phi(k) = \sum_{i=0}^{q_\nu+1} d_{\beta,\phi}(i)\beta(k-i) \tag{28}$$

where (27) and (28) can be incorporated in the measurement equation (18). The initial condition on $\beta(\cdot)$ is determined in Appendix B.

After the first two steps, (13) is replaced by

$$\mathbf{y}((k) = \mathbf{s}_k\left(\sum_{i=0}^{q_a} d_a(i)\alpha(k-i), \sum_{i=0}^{q_a} d_a(i)\alpha(k-1-i) \cdot \sum_{i=0}^{q_\nu+1} d_{\beta,\phi}(i)\beta(k-i), \sum_{i=0}^{q_\nu+1} d_{\beta,\nu}(i)\beta(k-i) \cdot \sum_{i=0}^{q_\nu+1} d_{\beta,\nu}(i)\beta(k-1-i)\right) + \mathbf{r}\mathbf{v}(k). \tag{29}$$

Equations (14), (15), (21), (26), and (29) define an intermediate system. The third step is to express the right-hand side of (29) as a function of $\alpha(\cdot)$ and $\beta(\cdot)$ rather than as a function of functions of $\alpha(\cdot)$ and $\beta(\cdot)$. For $n \in \{1, \cdots, N\}$, define

$$\breve{\epsilon}_n = n/N$$

$$\breve{e}_{\beta,n}(i) = \begin{cases} d_{\beta,\phi}(0) + \gamma_n d_{\beta,\nu}(0) \\ \quad i = 0 \\ d_{\beta,\phi}(i) + \gamma_n d_{\beta,\nu}(i) + \delta_n d_{\beta,\nu}(i-1) \\ \quad 1 \le i \le q_\nu + 1 \\ \delta_n d_{\beta,\nu}(q_\nu + 1) \\ \quad i = q_\nu + 2 \\ 0, \quad \text{otherwise.} \end{cases} \tag{30}$$

In order to have notation that easily incorporates the presence of both an in-phase and a quadrature-phase measurement at each sampling time, let $\epsilon_n = \breve{\epsilon}_{\lfloor (n+1)/2 \rfloor}$ and $e_{\beta,n}(i) = \breve{e}_{\beta,\lfloor (n+1)/2 \rfloor}(i)$ for $n \in \{1, \cdots, 2N\}$. Then, elementary calculations show that (14), (15), and (29) can be replaced by ($k \ge 1$)

$$\mathbf{y}(k) = \mathbf{h}_k((\alpha(k), \cdots, \alpha(k - q_a - 1), \beta(k) \cdots, \beta(k - q_\nu - 2)) + \mathbf{r}\mathbf{v}(k) \tag{31}$$

$$\mathbf{h}_k(\cdot) = (h_{k,1}(\cdot), \cdots, h_{k,2N}(\cdot))' \tag{32}$$

$$\begin{aligned} &h_{k,n}(\alpha(k), \cdots, \alpha(k - q_a - 1), \beta(k), \cdots, \beta(k - q_\nu - 2)) \\ &= \left\{ \left[\sum_{i=0}^{q_a} d_a(i)\alpha(k-1-i) \right]^2 (1 - \epsilon_n) \right. \\ &\quad \left. + \left[\sum_{i=0}^{q_a} d_a(i)\alpha(k-i) \right]^2 \epsilon_n \right\} \\ &\quad \times \begin{matrix} \cos \\ \sin \end{matrix} \left(\sum_{i=0}^{q_\nu+2} e_{\beta,n}(i)\beta(k-i) \right), \quad \begin{matrix} n \text{ odd} \\ n \text{ even.} \end{matrix} \end{aligned} \tag{33}$$

The fourth step is to define $\mathbf{x}(k) = (\alpha(k), \beta(k))' \in \mathcal{R}^2$, $\mathbf{w}(k) = (w_a(k-1), w_\nu(k-1))' \in \mathcal{R}^2$, $\mathbf{q} = \mathbf{I}_2$, $p = \max(p_a, p_\nu+1, q_a+2, q_\nu+3)$, $\mathbf{c}(i) = \operatorname{diag}(c_a(i), c_\beta(i)) \in \mathcal{R}^{2\times 2}$, and $\mathbf{f}_k(\mathbf{x}(k), \cdots, \mathbf{x}(k-p+1)) = \sum_{i=0}^{p-1} \mathbf{c}(i)\mathbf{x}(k-i)$. Finally, extend the definition of $h_{k,n}$ in (33) to negative times $k \in \{1-p, \cdots, 0\}$ ($n \in \{1, \cdots, 2N\}$) by setting

$$h_{k,n}(\cdot) = 0. \tag{34}$$

With these definitions of $\mathbf{x}(\cdot)$, $\mathbf{w}(\cdot)$, $\mathbf{q}$, p, and $\mathbf{f}_k(\cdot)$, (17) and (21), and (26) are equivalent. Furthermore, after replacing $\alpha(k)$ by $x_1(k)$ and $\beta(k)$ by $x_2(k)$ in (31)–(34), the resulting equations and (18) are equivalent.

B. *Calculation of the Expectation of Moments of* $\mathbf{f}$, $\mathbf{h}$, *and Their Gradients*

Now that the system is in the standard form of (17) and (18), we compute the Cramér–Rao bound by the method of [10]. The calculations are simplified by two facts: Because the state and measurement noises are uncorrelated (in the notation of [10], $U_k = 0$), it follows that a formula for $\mathrm{E}\{\nabla_{\mathbf{x}(k)}\mathbf{h}_i\}$ is not required. Because $\nabla_{\mathbf{x}(k)}\mathbf{f}_i$ is deterministic, there is no need to compute its moments.

If $\boldsymbol{\psi}$: $\mathcal{R}^n \to \mathcal{R}^m$, then define $\nabla\boldsymbol{\psi}$ to be the matrix with components (i,j) defined by $\partial\psi_i/\partial x_j$. For later use, define

$$\theta(k,n) = \sum_{i=0}^{q_\nu+2} e_{\beta,n}(i)\beta(k-i) \tag{35}$$

$$A_{l,n}(i) = 2d_a(i-1-l)(1-\epsilon_n)a(i-1) + 2d_a(i-l)\epsilon_n a(i) \tag{36}$$

$$B_{l,n}(i) = e_{\beta,n}(i-l)\left[a^2(i-1)(1-\epsilon_n) + a^2(i)\epsilon_n\right]. \tag{37}$$

The facts that $\theta(k, 2m-1) = \theta(k, 2m)$, $A_{l,2m-1}(i) = A_{l,2m}(i)$, and $B_{l,2m-1}(i) = B_{l,2m}(i)$ ($m \in \{1, \cdots, N\}$) will be important in the derivation of (42)–(44). The first step [10] to compute the Cramér–Rao bound is to evaluate $\nabla_{\mathbf{x}(k)}\mathbf{f}_i$ and $\nabla_{\mathbf{x}(k)}\mathbf{h}_i$ with the results that $\nabla_{\mathbf{x}(k)}\mathbf{f}_i = \mathbf{c}(i-k)$ and

$$\nabla_{\mathbf{x}(k)}\mathbf{h}_i = \begin{bmatrix} \dfrac{\partial h_{i,1}}{\partial x_1(k)} & \dfrac{\partial h_{i,1}}{\partial x_2(k)} \\ \dfrac{\partial h_{i,2}}{\partial x_1(k)} & \dfrac{\partial h_{i,2}}{\partial x_2(k)} \\ \vdots & \vdots \\ \dfrac{\partial h_{i,2N}}{\partial x_1(k)} & \dfrac{\partial h_{i,2N}}{\partial x_2(k)} \end{bmatrix} \tag{38}$$

where

$$\frac{\partial h_{i,n}}{\partial x_1(l)} = \frac{\partial h_{i,n}}{\partial \alpha(l)} = A_{l,n}(i) \begin{matrix} \cos \\ \sin \end{matrix} (\theta(i,n)), \quad \begin{matrix} n \text{ odd} \\ n \text{ even} \end{matrix} \tag{39}$$

$$\frac{\partial h_{i,n}}{\partial x_2(l)} = \frac{\partial h_{i,n}}{\partial \beta(l)} = \mp B_{l,n}(i) \begin{matrix} \sin \\ \cos \end{matrix} (\theta(i,n)), \quad \begin{matrix} n \text{ odd} \\ n \text{ even.} \end{matrix} \tag{40}$$

The second step [10] is to compute the expected value of

$$\nabla_{\mathbf{x}(l)}\mathbf{h}_i'\nabla_{\mathbf{x}(k)}\mathbf{h}_i = \begin{bmatrix} \displaystyle\sum_{n=1}^{2N} \frac{\partial h_{i,n}}{\partial x_1(l)}\frac{\partial h_{i,n}}{\partial x_1(k)} & \displaystyle\sum_{n=1}^{2N} \frac{\partial h_{i,n}}{\partial x_1(l)}\frac{\partial h_{i,n}}{\partial x_2(k)} \\ \displaystyle\sum_{n=1}^{2N} \frac{\partial h_{i,n}}{\partial x_2(l)}\frac{\partial h_{i,n}}{\partial x_1(k)} & \displaystyle\sum_{n=1}^{2N} \frac{\partial h_{i,n}}{\partial x_2(l)}\frac{\partial h_{i,n}}{\partial x_2(k)} \end{bmatrix}. \tag{41}$$

The summands in (41) are

$$\frac{\partial h_{i,n}}{\partial x_1(l)}\frac{\partial h_{i,n}}{\partial x_1(k)} = A_{l,n}(i)A_{k,n}(i)\left[\begin{matrix} \cos \\ \sin \end{matrix} (\theta(i,n))\right]^2 \quad \begin{matrix} n \text{ odd} \\ n \text{ even} \end{matrix}$$

$$\frac{\partial h_{i,n}}{\partial x_1(l)}\frac{\partial h_{i,n}}{\partial x_2(k)} = \mp A_{l,n}(i)B_{k,n}(i)\sin(\theta(i,n)) \cdot \cos(\theta(i,n)), \quad \begin{matrix} n \text{ odd} \\ n \text{ even.} \end{matrix}$$

$$\frac{\partial h_{i,n}}{\partial x_2(l)}\frac{\partial h_{i,n}}{\partial x_2(k)} = B_{l,n}(i)B_{k,n}(i)\left[\begin{matrix} \sin \\ \cos \end{matrix} (\theta(i,n))\right]^2 . \quad \begin{matrix} n \text{ odd} \\ n \text{ even.} \end{matrix}$$

Therefore, the components of (41) are

$$\sum_{n=1}^{2N} \frac{\partial h_{i,n}}{\partial x_1(l)} \frac{\partial h_{i,n}}{\partial x_1(k)} = \sum_{m=1}^{N} A_{l,2m}(i) A_{k,2m}(i) \quad (42)$$

$$\sum_{n=1}^{2N} \frac{\partial h_{i,n}}{\partial x_1(l)} \frac{\partial h_{i,n}}{\partial x_2(k)} = 0 \quad (43)$$

$$\sum_{n=1}^{2N} \frac{\partial h_{i,n}}{\partial x_2(l)} \frac{\partial h_{i,n}}{\partial x_2(k)} = \sum_{m=1}^{N} B_{l,2m}(i) B_{k,2m}(i). \quad (44)$$

In order to compute the expectation of (42)–(44) [and, thereby, (41)], we define

$$m_a(k) = \mathrm{E}\{a(k)\} \quad (45)$$

$$P_a(k_1,k_2) = \mathrm{E}\{[a(k_1) - m_a(k_1)][a(k_2) - m_a(k_2)]\} \quad (46)$$

which can easily be computed from (10). The stochastic processes $\theta(\cdot,\cdot)$, $A_{\cdot,\cdot}(\cdot)$, and $B_{\cdot,\cdot}(\cdot)$ are jointly Gaussian. Furthermore, $\theta(\cdot,\cdot)$ is independent of $\{A_{\cdot,\cdot}(\cdot), B_{\cdot,\cdot}(\cdot)\}$, but $A_{\cdot,\cdot}(\cdot)$ and $B_{\cdot,\cdot}(\cdot)$ are dependent. Using (45), (46), (36), and (37), we have

$$\begin{aligned}
&\mathrm{E}\{A_{l,2m}(i)A_{k,2m}(i)\} \\
&\quad = 4\left\{d_a(i-1-l)d_a(i-1-k)(1-\epsilon_{2m})^2\right. \\
&\quad\quad \cdot \left[P_a(i-1,i-1) + m_a^2(i-1)\right] \\
&\quad\quad + [d_a(i-1-l)d_a(i-k) + d_a(i-l)d_a(i-1-k)] \\
&\quad\quad \cdot (1-\epsilon_{2m})\epsilon_{2m}\left[P_a(i-1,i) + m_a(i-1)m_a(i)\right] \\
&\quad\quad \left. + d_a(i-l)d_a(i-k)\epsilon_{2m}^2\left[P_a(i,i) + m_a^2(i)\right]\right\} \quad (47)
\end{aligned}$$

$$\begin{aligned}
&\mathrm{E}\{B_{l,2m}(i)B_{k,2m}(i)\} \\
&\quad = e_{\beta,2m}(i-l)e_{\beta,2m}(i-k)\left\{\left[3P_a^2(i-1,i-1)\right.\right. \\
&\quad\quad \left. + 6P_a(i-1,i-1)m_a^2(i-1) + m_a^4(i-1)\right](1-\epsilon_{2m})^2 \\
&\quad\quad + 2\left[P_a(i,i)P_a(i-1,i-1) + 2P_a^2(i,i-1)\right. \\
&\quad\quad + m_a^2(i)P_a(i-1,i-1) + m_a^2(i-1)P_a(i,i) \\
&\quad\quad \left. + 4m_a(i)m_a(i-1)P_a(i,i-1) + m_a^2(i)m_a^2(i-1)\right] \\
&\quad\quad \cdot (1-\epsilon_{2m})\epsilon_{2m} \\
&\quad\quad \left. + \left[3P_a^2(i,i) + 6P_a(i,i)m_a^2(i) + m_a^4(i)\right]\epsilon_{2m}^2\right\} \quad (48)
\end{aligned}$$

where Gaussian moment factoring [44, p. 229] was used.

Define $\boldsymbol{\mathcal{H}}(i,k,l) = \mathrm{E}\{\nabla_{\mathbf{x}(l)}\mathbf{h}_i' \nabla_{\mathbf{x}(k)}\mathbf{h}_i\}$ with components

$$\mathcal{H}_{j,j'}(i,k,l) = \mathrm{E}\left\{\sum_{n=1}^{2N} [\partial h_{i,n}/\partial x_j(l)]\,[\partial h_{i,n}/\partial x_{j'}(k)]\right\} \quad (49)$$

where $n \in \{1,\cdots,2N\}$, and $j,j' \in \{1,2\}$. By (34), it follows that $\mathcal{H}_{j,j'}(i,k,l) = 0$ for $i \in \{1-p,\cdots,0\}$. From (43), we immediately have that $\mathcal{H}_{j,j'}(i,k,l) = 0$ for all $j \neq j'$. Finally, using (47) in (42) determines $\mathcal{H}_{1,1}(i,k,l)$ and using (48) in (44) determines $\mathcal{H}_{2,2}(i,k,l)$. In both of these calculations, there is a sum over the m index that can be done analytically with the aid of analytical formulas for $\sum_{n=1}^{N} n^l$ ($l = 1$ or $l = 2$), but the resulting formulas are too complicated to be included here.

```
for(k = K; k >= 1 - p; k - -){
    C_{k,0} = ( D_{k,k} - Σ_{i=k+1}^{min(k+p-1,K)} C'_{i,i-k} C_{i,i-k} )^{T/2}
    for(l = k - 1; l >= max(k - p + 1, 1 - p); l - -){
        C_{k,k-l} = [ ( D_{k,l} - Σ_{i=k+1}^{min(l+p-1,K)} C'_{i,i-l} C_{i,i-k} ) C_{k,0}^{-1} ]'
    }
}
```

$$\mathbf{C}_{k,0} = \left(\mathbf{D}_{k,k} - \sum_{i=k+1}^{\min(k+p-1,K)} \mathbf{C}'_{i,i-k}\mathbf{C}_{i,i-k}\right)^{T/2}$$

$$\mathbf{C}_{k,k-l} = \left[\left(\mathbf{D}_{k,l} - \sum_{i=k+1}^{\min(l+p-1,K)} \mathbf{C}'_{i,i-l}\mathbf{C}_{i,i-k}\right)\mathbf{C}_{k,0}^{-1}\right]'$$

Fig. 4. Algorithm for the computation of $\mathbf{C}_{k,i}$. The control structures are written in the C programming language. The notation $^{1/2}$ indicates matrix square roots [$\mathbf{R} = \mathbf{R}^{1/2}\left(\mathbf{R}^{1/2}\right)'$, $\mathbf{R}^{T/2} \doteq \left(\mathbf{R}^{1/2}\right)'$].

C. Determination of the Linear System of (19) and (20)

After extensive computations to evaluate (5)–(8) and (10)–(16) and [10] followed by dimensional reduction [as described following (20)], we find the following results. It is sufficient to choose a comparison system [(19) and (20)] with parameters $\mathbf{A}_{k,k-l} = \mathbf{c}(k-l)$, $\mathbf{Q} = \mathbf{I}_2$, $\mathbf{R} = \mathbf{I}_2$, and $\mathbf{P}_0^{\mathbf{x}} = f\left(\mathbf{P}_0^{(\alpha)}, \mathbf{P}_0^{(\beta)}\right)$, where the function $f(\cdot,\cdot)$ is described in Appendix C. $\mathbf{C}_{k,i} \in \mathcal{R}^{2\times 2}$ cannot be computed explicitly but, rather, is the solution of (17) and (18) in [10], which can be solved by the following algorithm. Define

$$\mathbf{D}_{k,l} = \frac{1}{r^2} \sum_{i=k}^{\min(l+p-1,K)} \boldsymbol{\mathcal{H}}(i,k,l). \quad (50)$$

Then, $\mathbf{C}_{k,i}$ is determined by the algorithm shown in Fig. 4. Once $\mathbf{A}_{k,k-l}$, $\mathbf{C}_{k,i}$, $\mathbf{P}_0^{\mathbf{x}}$, $\mathbf{Q}$, and $\mathbf{R}$ have been determined, then the state augmentation and Kalman filter computations needed to determine the bound are standard [2].

$\boldsymbol{\mathcal{H}}$ is diagonal, and therefore, $\mathcal{D}_{k,l}$ is diagonal, and therefore, $\mathbf{C}_{k,i}$ is also diagonal. For this reason, the algorithm of Fig. 4 can be run independently on the scalar diagonal elements of $\mathcal{D}_{k,l}$ and $\mathbf{C}_{k,l}$. The fact that the amplitude and phase variables become uncoupled in the Cramér–Rao bound is not unexpected because of the following.

1) Both in-phase and quadrature-phase measurements are made.
2) The noise variances on the in-phase and quadrature-phase measurements are equal, and the correlation coefficient is zero.
3) The state equations for amplitude and phase variables are uncoupled.

Exactly the same behavior occurs in the following simple static problem: Let $(a,\phi,r_I,r_Q)'$ be Gaussian with mean $(m_a,m_\phi,0,0)'$ and covariance $\mathrm{diag}(\sigma_a^2,\sigma_\phi^2,\sigma_r^2,\sigma_r^2)$. The goal is the estimate $(a,\phi)'$ based on measurements $y_I = a^2\ \cos(\phi) + r_I$ and $y_Q = a^2\ \sin(\phi) + r_Q$. The resulting Fisher information matrix [44, p. 84] is

$$J = \mathrm{diag}\left(4\,\frac{\sigma_a^2}{\sigma_r^2} + \frac{1}{\sigma_a^2}, \frac{3\sigma_a^4 + 6\sigma_a^2 m_a^2 + m_a^4}{\sigma_r^2} + \frac{1}{\sigma_\phi^2}\right)$$

which has the same diagonal structure that is present in the current problem.

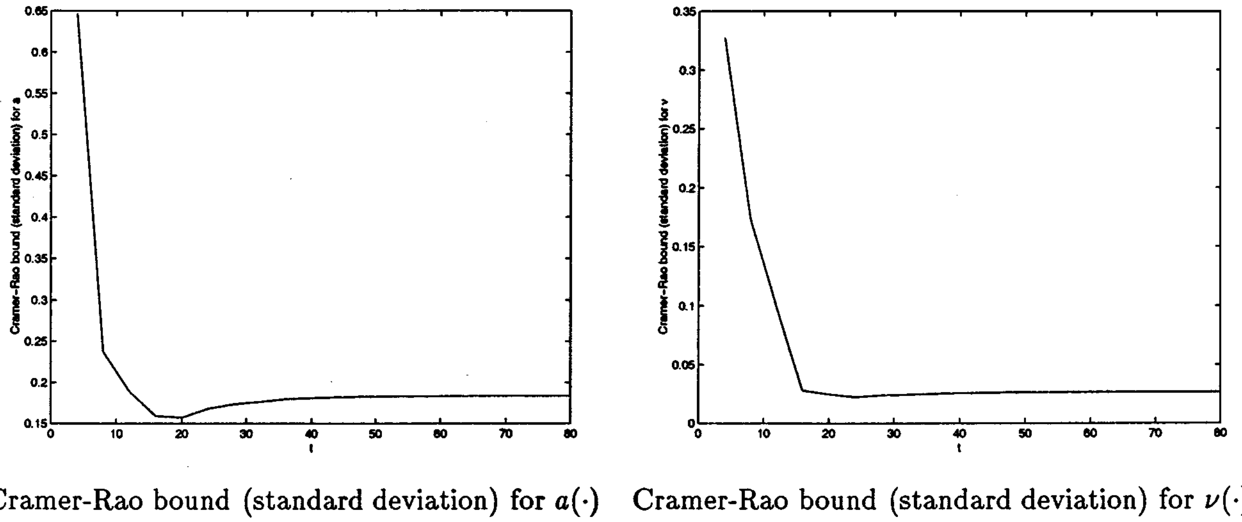

Fig. 5. Cramér-Rao bounds as a function of time for $a(\cdot)$ and $\nu(\cdot)$.

Since $a(\cdot)$, $\nu(\cdot)$, and $\phi(\cdot)$ are related by FIR filters to $\alpha(\cdot)$ and $\beta(\cdot)$, any second-order statistic (i.e., the Cramér–Rao bound) that applies to $\alpha(\cdot)$ and $\beta(\cdot)$ can be transformed into the corresponding second-order statistic that applies to $a(\cdot)$, $\nu(\cdot)$, and $\phi(\cdot)$.

VII. Application of the Cramér-Rao Bound

The Cramér-Rao bound (see Section IV) applies to the model of Section V that is much more general than the speech model of Section II. In order to demonstrate the generality of the Cramér–Rao bound algorithm and our software [2] for computing the bound, we apply it to a model in which $a(\cdot)$ is an ARMA model with $p_a = 6$ and $q_a = 3$ whose frequency response is lowpass with digital frequency cut-off 0.4π, and $\nu(\cdot)$ is an ARMA model with $p_a = 8$ and $q_a = 5$ whose frequency response is lowpass with digital frequency cut-off 0.3π. Both ARMA frequency responses are generalized Butterworth filters that were designed with the algorithm `maxflat` in Matlab. The resulting Cramér–Rao bounds, which are functions of time, are shown in Fig. 5. Notice the prominent initial transient, which is important since we envision reinitializing these methods frequently at, for example, an unvoiced to voiced transition in speech or an edge in a texture segmentation of an image.

VIII. Discussion

In this paper, we

1) proposed a stochastic dynamical system model for the complex baseband representation of AM–FM signals, which describes the signal in terms of frames that are modeled on the frames used in speech processing;
2) proposed an adaptive nonlinear estimator for extracting the AM and FM modulating signals, which requires very modest amounts of computation;
3) demonstrated the approach of 1) and 2) on a speech signal;
4) computed the Cramér–Rao bound on the estimation error variance performance of any estimator for extracting the AM and FM signals in the general class of model proposed in 1).

The model [Item 1)] pervades all aspects of the work.

There are several aspects of the amplitude part of the model that deserve further comment. Ideally, we would like to have a dynamical system for a positively-constrained amplitude. For instance, a positively-constrained amplitude might be one component of the state. It is difficult to propose such a model, especially if the dynamical system is driven by an additive noise. For that reason, we proposed a dynamical system for the square root of the amplitude and then squared the output of the dynamical system in the measurement equation. This squaring operation interacts with the interpolation operation: square the interpolated values versus interpolate the squared values. We choose to interpolate the squared values because squaring the interpolated values can lead to cusps in the amplitude function at the frame boundaries, and such cusps increase the bandwidth. In this paper, we have only considered linear interpolants. It is possible to get smoother and, therefore, lower bandwidth signals by using higher order interpolants, but it is harder to guarantee that the interpolated amplitude will be positive. This squaring operation interacts with the Cramér–Rao bound. If the initial condition on $a(\cdot)$ is zero, then the conditional probability density on the $a(\cdot)$ trajectory given the measurements is symmetrically bimodal: For any trajectory, the value of the probability for that trajectory and for the negative of that trajectory are equal. If the initial condition on $a(\cdot)$ is nonzero, then the probability density is not symmetrically bimodal, but for a stable model, the initial condition decays to zero exponentially in time, and therefore, the probability density is asymmetrically bimodal. The bimodal character need not hurt the performance of a nonlinear filter but does distort the Cramér–Rao bound since if most of the probability is concentrated in the regions $a_0 \pm \delta$ and $-a_0 \pm \delta$, then the Cramér–Rao bound will give a standard deviation on the order of $a_0 + \delta$, whereas a filter could have a typical error in estimating $a^2(\cdot)$ on the order of δ.

[2] See http://www.ece.purdue.edu/~doerschu/4SPpaper.README.html for the software used in this section.

In this paper, we propose a statistical parametric approach to AM–FM modeling and processing. An advantage of this approach is that it allows the characteristics of the modulating signals (e.g., bandwidth) and the measurement noise (i.e., variance) to be precisely described. Furthermore, the characteristics of the signals and noise imply optimal (typically impractical) processing and bounds on the performance achievable with any processing so that practical suboptimal processing can be compared with an absolute standard.

APPENDIX A
STATE VARIABLE REPRESENTATIONS OF THE ARMA SYSTEMS

First, consider the amplitude ARMA process (10). Assume that $c_a(p_a) \neq 0$ and that $d_a(q_a) \neq 0$. Assume that the polynomials in z defined by $\prod_{i=0}^{q_a} d_a(i)z^{-i}$ and $\prod_{i=0}^{p_a} c_a(i)z^{-i}$ have no common factors. Define $\mathbf{c}_a = [c_a(p_a), \cdots, c_a(1)]' \in \mathcal{R}^{p_a}$, $\mathbf{d}_a = [d_a(q_a), \cdots, d_a(1), d_a(0)]' \in \mathcal{R}^{q_a+1}$, and $n = \max(p_a, q_a+2)$. n is the dimension of the state variable realization. Because the measurement equation (13) depends on both $a(k)$ and $a(k-1)$, we require a 2-D measurement equation in the state variable realization. If $p_a \geq q_a + 2$, then $n = p_a$, and the realization is

$$\begin{bmatrix} x_a(k+1-n) \\ \vdots \\ x_a(k) \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{n-1} & \mathbf{I}_{n-1} \\ & -\mathbf{c}_a' \end{bmatrix} \begin{bmatrix} x_a(k-n) \\ \vdots \\ x_a(k-1) \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{n-1} \\ 1 \end{bmatrix} w_a(k) \quad (51)$$

$$\begin{bmatrix} a(k) \\ a(k-1) \end{bmatrix} = \begin{bmatrix} \mathbf{0}'_{p_a-q_a-1} & \mathbf{d}'_a \\ \mathbf{0}'_{p_a-q_a-2} & \mathbf{d}'_a \quad 0 \end{bmatrix} \cdot \begin{bmatrix} x_a(k+1-n) \\ \vdots \\ x_a(k) \end{bmatrix} \quad (52)$$

whereas if $p_a < q_a+2$, then $n = q_a+2$, and the realization is

$$\begin{bmatrix} x_a(k+1-n) \\ \vdots \\ x_a(k) \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{n-1} & \mathbf{I}_{n-1} \\ \mathbf{0}'_{q_a+2-p_a} & -\mathbf{c}'_a \end{bmatrix} \cdot \begin{bmatrix} x_a(k-n) \\ \vdots \\ x_a(k-1) \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{n-1} \\ 1 \end{bmatrix} w_a(k) \quad (53)$$

$$\begin{bmatrix} a(k) \\ a(k-1) \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{d}'_a \\ \mathbf{d}'_a & 0 \end{bmatrix} \begin{bmatrix} x_a(k+1-n) \\ \vdots \\ x_a(k) \end{bmatrix}. \quad (54)$$

Make the obvious definitions of $\mathbf{x}_a(k)$, $\mathbf{A}_a$, $\mathbf{b}_a$, and $\mathbf{C}_a$ to get

$$\mathbf{x}_a(k) = \mathbf{A}_a\mathbf{x}_a(k-1) + \mathbf{b}_a w_a(k) \quad (55)$$

$$\begin{bmatrix} a(k) \\ a(k-1) \end{bmatrix} = \mathbf{C}_a\mathbf{x}_a(k). \quad (56)$$

The initial condition (mean $\mathbf{m}_0^{(a)}$ and covariance $\mathbf{P}_0^{(a)}$) is on $\mathbf{x}_a(0)$.

Now, consider the instantaneous frequency ARMA process (11). Assume that $c_\nu(p_\nu) \neq 0$ and $d_\nu(q_\nu) \neq 0$. Assume that the polynomials in z defined by $\prod_{i=0}^{q_\nu} d_\nu(i)z^{-i}$ and $\prod_{i=0}^{p_\nu} c_\nu(i)z^{-i}$ have no common factors. Define $\mathbf{c}_\nu = [c_\nu(p_\nu), \cdots, c_\nu(1)]' \in \mathcal{R}^{p_\nu}$, $\mathbf{d}_\nu = [d_\nu(q_\nu), \cdots, d_\nu(1), d_\nu(0)]' \in \mathcal{R}^{q_\nu+1}$, and $n = \max(p_\nu, q_\nu+2)$. n is the dimension of the state variable realization. Because the measurement equation (13) depends on both $\nu(k)$ and $\nu(k-1)$, we require a 2-D measurement equation in the state variable realization. If $p_\nu \geq q_\nu + 2$, then $n = p_\nu$, and the realization is analogous to (51) and (52) with $\mathbf{d}_a$ replaced by $\mathbf{d}_\nu$ and $\mathbf{c}_a$ replaced by $\mathbf{c}_\nu$, whereas if $p_\nu < q_\nu + 2$, then $n = q_\nu + 2$, and the realization is analogous to (53) and (54) with the same replacements. Make the obvious definitions of $\mathbf{x}_\nu(k)$, $\mathbf{A}_\nu$, $\mathbf{b}_\nu$, and $\mathbf{C}_\nu$ to get

$$\mathbf{x}_\nu(k) = \mathbf{A}_\nu\mathbf{x}_\nu(k-1) + \mathbf{b}_\nu w_\nu(k)$$

$$\begin{bmatrix} \nu(k) \\ \nu(k-1) \end{bmatrix} = \mathbf{C}_\nu\mathbf{x}_\nu(k).$$

The phase AR process (12) has the realization

$$\begin{bmatrix} \phi(k) \\ \nu(k) \end{bmatrix} = \begin{bmatrix} 1 & \delta_N \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \phi(k-1) \\ \nu(k-1) \end{bmatrix} + \begin{bmatrix} \gamma_N \\ 1 \end{bmatrix} \nu(k)$$

$$[\phi(k)] = [1 \quad 0] \begin{bmatrix} \phi(k) \\ \nu(k) \end{bmatrix}.$$

The instantaneous frequency and phase state variable systems can be combined. Let $\mathbf{h}'$ be the first row of $\mathbf{C}_\nu$. Then, $\phi(k) = \phi(k-1) + \delta_N\nu(k-1) + \gamma_N\nu(k) = \phi(k-1) + \delta_N\mathbf{h}'\mathbf{x}_\nu(k-1) + \gamma_N\mathbf{h}'\mathbf{x}_\nu(k) = \phi(k-1) + \delta_N\mathbf{h}'\mathbf{x}_\nu(k-1) + \gamma_N\mathbf{h}'[\mathbf{A}_\nu\mathbf{x}_\nu(k-1) + \mathbf{b}_\nu w_\nu(k)] = \phi(k-1) + (\delta_N\mathbf{h}' + \gamma_N\mathbf{h}'\mathbf{A}_\nu)\mathbf{x}_\nu(k-1) + \gamma_N\mathbf{h}'\mathbf{b}_\nu w_\nu(k)$. Therefore

$$\begin{bmatrix} \mathbf{x}_\nu(k) \\ \phi(k) \end{bmatrix} = \begin{bmatrix} \mathbf{A}_\nu & \mathbf{0}_n \\ \delta_N\mathbf{h}' + \gamma_N\mathbf{h}'\mathbf{A}_\nu & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}_\nu(k-1) \\ \phi(k-1) \end{bmatrix} + \begin{bmatrix} \mathbf{b}_\nu \\ \gamma_N\mathbf{h}'\mathbf{b}_\nu \end{bmatrix} w_\nu(k)$$

$$\begin{bmatrix} \nu(k) \\ \nu(k-1) \\ \phi(k) \end{bmatrix} = \begin{bmatrix} \mathbf{C}_\nu & \mathbf{0}_2 \\ \mathbf{0}'_n & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}_\nu(k) \\ \phi(k) \end{bmatrix}.$$

Make the obvious definitions of $\mathbf{x}_f(k)$, $\mathbf{A}_f$, $\mathbf{b}_f$, and $\mathbf{C}_f$ to get

$$\mathbf{x}_f(k) = \mathbf{A}_f\mathbf{x}_f(k-1) + \mathbf{b}_f w_\nu(k) \quad (57)$$

$$\begin{bmatrix} \nu(k) \\ \nu(k-1) \\ \phi(k) \end{bmatrix} = \mathbf{C}_f\mathbf{x}_f(k). \quad (58)$$

The initial condition (mean $\mathbf{m}_0^{(f)}$ and covariance $\mathbf{P}_0^{(f)}$) is on $\mathbf{x}_f(0)$.

APPENDIX B
STATE VARIABLE REALIZATION USING $\beta(\cdot)$ AND $\mathbf{m}_0^{(\beta)}$ AND $\mathbf{P}_0^{(\beta)}$

The first step is to realize $[\nu(k), \nu(k-1), \phi(k)]'$ as the output of a state variable system where the state is composed of lagged versions of $\beta(\cdot)$. We have already assumed in Appendix A that $c_\nu(p_\nu) \neq 0$ and $d_\nu(q_\nu) \neq 0$ and that the polynomials in z defined by $\prod_{i=0}^{q_\nu} d_\nu(i)z^{-i}$ and $\prod_{i=0}^{p_\nu} c_\nu(i)z^{-i}$ have no common factors. Define $p_\beta = p_\nu + 1$, $q_\beta = q_\nu + 1$, $n = \max(p_\beta, q_\beta + 2)$, $\mathbf{c}_\beta = [c_\beta(p_\beta), \cdots, c_\beta(1)]' \in \mathcal{R}^{p_\beta}$, $\mathbf{d}_{\beta,\nu} = [d_{\beta,\nu}(q_\beta), \cdots, d_{\beta,\nu}(1), d_{\beta,\nu}(0)]' \in \mathcal{R}^{q_\beta}$, and $\mathbf{d}_{\beta,\phi} = [d_{\beta,\phi}(q_\beta), \cdots, d_{\beta,\phi}(1), d_{\beta,\phi}(0)]' \in \mathcal{R}^{q_\beta}$. n is the dimension of the state variable realization. Because the measurement equation (13) depends on $\nu(k)$, $\nu(k-1)$, and $\phi(k)$, we require a 3-D measurement equation in the state variable realization. If $p_\beta \geq q_\beta + 2$, then $n = p_\beta$ and the realization is

$$\begin{bmatrix} \beta(k+1-n) \\ \vdots \\ \beta(k) \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{n-1} & \mathbf{I}_{n-1} \\ & -\mathbf{c}_\beta' \end{bmatrix} \begin{bmatrix} \beta(k-n) \\ \vdots \\ \beta(k-1) \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{n-1} \\ 1 \end{bmatrix} w_\nu(k)$$

$$\begin{bmatrix} \nu(k) \\ \nu(k-1) \\ \phi(k) \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{p_\beta-q_\beta-1}' & \mathbf{d}_{\beta,\nu}' & \\ \mathbf{0}_{p_\beta-q_\beta-2}' & \mathbf{d}_{\beta,\nu}' & 0 \\ \mathbf{0}_{p_\beta-q_\beta-1}' & \mathbf{d}_{\beta,\phi}' & \end{bmatrix} \begin{bmatrix} \beta(k+1-n) \\ \vdots \\ \beta(k) \end{bmatrix}$$

whereas if $p_\beta < q_\beta + 2$, then $n = q_\beta + 2$, and the realization is

$$\begin{bmatrix} \beta(k+1-n) \\ \vdots \\ \beta(k) \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{n-1} & \mathbf{I}_{n-1} \\ \mathbf{0}_{q_\beta+2-p_\beta}' & -\mathbf{c}_\beta' \end{bmatrix} \begin{bmatrix} \beta(k-n) \\ \vdots \\ \beta(k-1) \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{n-1} \\ 1 \end{bmatrix} w_\nu(k)$$

$$\begin{bmatrix} \nu(k) \\ \nu(k-1) \\ \phi(k) \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{d}_{\beta,\nu}' \\ \mathbf{d}_{\beta,\nu}' & 0 \\ 0 & \mathbf{d}_{\beta,\phi}' \end{bmatrix} \begin{bmatrix} \beta(k+1-n) \\ \vdots \\ \beta(k) \end{bmatrix}.$$

Make the obvious definitions of $\mathbf{x}_\beta(k)$, $\mathbf{A}_\beta$, $\mathbf{b}_\beta$, and $\mathbf{C}_\beta$ to get

$$\mathbf{x}_\beta(k) = \mathbf{A}_\beta \mathbf{x}_\beta(k-1) + \mathbf{b}_\beta w_\nu(k) \tag{59}$$

$$\begin{bmatrix} \nu(k) \\ \nu(k-1) \\ \phi(k) \end{bmatrix} = \mathbf{C}_\beta \mathbf{x}_\beta(k). \tag{60}$$

We now compute the initial condition (mean $\mathbf{m}_0^{(\beta)}$ and covariance $\mathbf{P}_0^{(\beta)}$) on $\mathbf{x}_\beta(0)$. Let $\mathbf{A}_f$ and $\mathbf{C}_f$ be as implied by (57) and (58). Define $\mathbf{h}_f'$ and $\mathbf{t}_f'$ to be the second and third rows of $\mathbf{C}_f$, respectively. Let $\mathbf{A}_\beta$ and $\mathbf{C}_\beta$ be as implied by (59) and (60). Define $\mathbf{h}_\beta'$ and $\mathbf{t}_\beta'$ to be the second and third rows of $\mathbf{C}_\beta$, respectively. Define $\mathbf{T}_f = \left[\mathbf{h}_f, (\mathbf{h}_f'\mathbf{A}_f)', \cdots, (\mathbf{h}_f'\mathbf{A}_f^{p_\beta-2})', \mathbf{t}_f\right]'$, $\mathbf{T}_\beta = \left[\mathbf{h}_\beta, (\mathbf{h}_\beta'\mathbf{A}_\beta)', \cdots, (\mathbf{h}_\beta'\mathbf{A}_\beta^{p_\beta-2})', \mathbf{t}_\beta\right]'$, and $\mathbf{T} = \mathbf{T}_\beta^{-1}\mathbf{T}_f$. Then, computing the response to just the initial condition, we have $\mathbf{T}_\beta \mathbf{x}_\beta(0) = [\nu(-1), \cdots, \nu(n-2), \phi(0)]' = \mathbf{T}_f \mathbf{x}_f(0)$ so that $\mathbf{x}_\beta(0) = \mathbf{T}_\beta^{-1}\mathbf{T}_f \mathbf{x}_f(0) = \mathbf{T}\mathbf{x}_f(0)$. Therefore, the mean and covariance of $\mathbf{x}_\beta(0)$ are $\mathbf{m}_0^{(\beta)} = \mathbf{T}\mathbf{m}_0^{(f)}$ and $\mathbf{P}_0^{(\beta)} = \mathbf{T}\mathbf{P}_0^{(f)}\mathbf{T}'$, respectively.

Appendix C
Initial Condition on $\mathbf{P}_0^{\mathbf{x}}$

$\mathbf{P}_0^{\mathbf{x}}$ is constructed from $\mathbf{P}_0^{(\alpha)}$ and $\mathbf{P}_0^{(\beta)}$. Indicate components by $(\cdot)_{i,j}$. If i is even and j is odd or if i is odd and j is even, then $(\mathbf{P}_0^{\mathbf{x}})_{i,j} = 0$ because this is a cross covariance between the α and β stochastic processes. If i and j are both odd, then if $i \leq \max(p_a, q_a + 2)$ and $j \leq \max(p_a, q_a + 2)$, then $(\mathbf{P}_0^{\mathbf{x}})_{i,j} = (\mathbf{P}_0^{(\alpha)})_{(i+1)/2,(j+1)/2}$, and otherwise, $(\mathbf{P}_0^{\mathbf{x}})_{i,j}$ is arbitrary. If i and j are both even, then if $i \leq \max(p_\nu + 1, q_\nu + 3)$ and $j \leq \max(p_\nu + 1, q_\nu + 3)$, then $(\mathbf{P}_0^{\mathbf{x}})_{i,j} = (\mathbf{P}_0^{(\beta)})_{i/2,j/2}$, and otherwise, $(\mathbf{P}_0^{\mathbf{x}})_{i,j}$ is arbitrary.

References

[1] F. A. Cheikh, R. Hamila, M. Gabbouj, and J. Astola, "Impulse noise removal in highly corrupted color images," in *Proc. IEEE Int. Conf. Image Process.*, vol. I, Lausanne, Switzerland, Sept. 16–19, 1996, pp. 997–1000.

[2] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

[3] S. Barbarossa, "Analysis of multicomponent LFM signals by combined Wigner-Hough transform," *IEEE Trans. Signal Processing*, vol. 46, pp. 1511–1515, June 1995.

[4] R. R. Bitmead, A. C. Tsoi, and P. J. Parker, "A Kalman filtering approach to short-time fourier analysis," *IEEE Trans. Acoust., Speech Signal Processing*, vol. ASSP-34, pp. 1493–1501, June 1986.

[5] B. Boashash, P. J. O'Shea, and M. J. Arnold, "Algorithms for instantaneous frequency estimation: A comparative study," in *Advanced Signal Processing Algorithms, Architectures and Implementations*. Bellingham, WA: SPIE, Aug. 1990, vol. 1348, pp. 24–46.

[6] B. Z. Bobrovsky and M. Zakai, "A lower bound on the estimation error for Markov processes," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 785–788, Dec. 1975.

[7] A. C. Bovik, J. P. Havlicek, M. D. Desai, and D. S. Harding, "Limits on discrete modulated signals," *IEEE Trans. Signal Processing*, vol. 45, pp. 867–879, Apr. 1997.

[8] A. C. Bovik, P. Maragos, and T. F. Quatieri, "AM–FM energy detection and separation in noise using multiband energy operators," *IEEE Trans. Signal Processing*, vol. 41, pp. 3245–3265, Dec. 1993.

[9] F. S. Cohen, S. Kadambe, and G. F. Boudreaux-Bartels, "Tracking of unknown nonstationary chirp signals using unsupervised clustering in the Wigner distribution," *IEEE Trans. Signal Processing*, vol. 41, pp. 3085–3101, Nov. 1993.

[10] P. C. Doerschuk, "Cramer–Rao bounds for discrete-time nonlinear filtering problems," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 1465–1469, Aug. 1995.

[11] W. M. Fisher, V. Zue, J. Bernstein, and D. Pallett, An accoustic-phonetic database, in 113th Meet. Acoust. Soc. Amer., 1987.

[12] J. M. Francos and B. Friedlander, "Bounds for estimation of multicomponent signals with random amplitude and deterministic phase," *IEEE Trans. Signal Processing*, vol. 43, pp. 1161–1172, May 1995.

[13] B. Friedlander and J. M. Francos, "Estimation of amplitude and phase parameters of multicomponent signals," *IEEE Trans. Signal Processing*, vol. 43, pp. 917–926, Apr. 1995.

[14] J. I. Galdos, "A Cramer-Rao bound for multidimensional discrete-time dynamical systems," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 117–119, Feb. 1980.

[15] J. D. Gorman and A. O. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1285–1301, Nov. 1990.

[16] D. W. Griffin and J. S. Lim, "Multiband excitation vocoder," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1223–1235, Aug. 1988.

[17] J. H. L. Hansen, L. Gavidia-Ceballos, and J. F. Kaiser, "A nonlinear operator-based speech feature analysis method with application to vocal fold pathology assessment," *IEEE Trans. Biomed. Eng.*, vol. 45, pp. 300–312, Mar. 1998.

[18] ——, "A nonlinear operator-based speech feature analysis method with application to vocal fold pathology assessment," *IEEE Trans. Biomed. Eng.*, vol. 45, pp. 300–312, Mar. 1998.

[19] J. P. Havlicek, D. S. Harding, and A. C. Bovik, "Discrete quasieigenfunction approximation for AM–FM image analysis," in *Proc. IEEE Int. Conf. Image Process.*, Lausanne, Switzerland, Sept. 16–19, 1996, pp. 633–636.

[20] B. James, B. D. O. Anderson, and R. C. Williamson, "Conditional mean and maximum likelihood approaches to multiharmonic frequency estimation," *IEEE Trans. Signal Processing*, vol. 42, pp. 1366–1375, June 1994.

[21] F.-C. Jeng and J. W. Woods, "Compound Gauss-Markov random fields for image estimation," *IEEE Trans. Signal Processing*, vol. 39, pp. 683–697, Mar. 1991.

[22] S. R. Kadaba, S. B. Gelfand, and R. L. Kashyap, "Bayesian decision feedback for segmentation of binary images," *IEEE Trans. Image Processing*, pp. 1163–1178, July 1996.

[23] ——, "Recursive estimation of images using non-Gaussian autoregressive models," *IEEE Trans. Image Processing*, vol. 7, pp. 1439–1452, Oct. 1998.

[24] V. Katkovnik and LJ. Stankovic, "Instantaneous frequency estimation using the Wigner distribution with varying and data-driven window length," *IEEE Trans. Signal Processing*, vol. 46, pp. 2315–2325, Sept. 1998.

[25] B. F. La Scala and R. R. Bitmead, "Design of an extended Kalman filter frequency tracker," *IEEE Trans. Signal Processing*, vol. 44, pp. 739–742, Mar. 1996.

[26] B. F. La Scala, R. R. Bitmead, and R. J. Matthew, "Conditions for stability of the extended Kalman filter and their application to the frequency tracking problem," *Math. Contr., Signals, Syst.*, vol. 8, pp. 1–26, 1995.

[27] S. Lu and P. C. Doerschuk, "Nonlinear modeling and procesing of speech based on sums of AM–FM formant models," *IEEE Trans. Signal Processing*, vol. 44, pp. 773–782, Apr. 1996.

[28] P. Maragos and A. C. Bovik, "Image demodulation using multidimensional energy separation," *J. Opt. Soc. Amer. A*, vol. 12, no. 9, pp. 1867–1876, Sept. 1995.

[29] 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, pp. 3024–3051, Oct. 1993.

[30] ——, "On amplitude and frequency demodulation using energy operators," *IEEE Trans. Signal Processing*, vol. 41, pp. 1532–1550, Apr. 1993.

[31] J. S. Marques, L. B. Almeida, and J. M. Tribolet, "Harmonic coding at 4.8 kb/s," in *Proc. ICASSP*, vol. 1, 1990, pp. 17–20.

[32] T. L. Marzetta, "A simple derivation of the constrained multiple parameter Cramer-Rao bound," *IEEE Trans. Signal Processing*, vol. 41, pp. 2247–2249, June 1993.

[33] R. J. McAulay and T. F. Quatieri, "Speech analysis/synthesis based on a sinusoidal representation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 744–754, Aug. 1986.

[34] K. Nishiyama, "A nonlinear filter for estimating a sinusoidal signal and its parameters in white noise: On the case of a single sinusoid," *IEEE Trans. Signal Processing*, vol. 45, pp. 970–981, Apr. 1997.

[35] S. Peleg and B. Porat, "The Cramer-Rao lower bound for signals with constant amplitude and polynomial phase," *IEEE Trans. Signal Processing*, vol. 39, pp. 749–752, Mar. 1991.

[36] A. Potamianos and P. Maragos, "Speech formant frequency and bandwidth tracking using multiband energy demodulation," *J. Acoust. Soc. Amer.*, vol. 99, no. 6, pp. 3795–3806, June 1996.

[37] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.

[38] T. F. Quatieri, T. E. Hanna, and G. C. O'Leary, "AM–FM separation using auditory-motivated filters," *IEEE Trans. Speech Audio Processing*, vol. 5, pp. 465–480, Sept. 1997.

[39] L. Rabiner and J. Biing-Hwang, *Fundamentals of Speech Recognition*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[40] 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, pp. 1708–1709, June 1988.

[41] ——, "Instantaneous frequency estimation of quadratic and cubic FM signals using the cross polynomial Wigner–Ville distribution," *IEEE Trans. Signal Processing*, vol. 44, pp. 1549–1553, May 1996.

[42] H. M. Teager, "Some observations on oral air flow during phonation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 599–601, Oct. 1980.

[43] P. Tichavsky, C. H. Muravchik, and A. Nehorai, "Posterior Cramer–Rao bounds for discrete-time nonlinear filtering," *IEEE Trans. Signal Processing*, vol. 46, pp. 1386–1396, May 1998.

[44] H. L. van Trees, *Detection, Estimation, and Modulation Theory: Part I*. New York: Wiley, 1968.

[45] J. W. Woods and C. H. Radewan, "Kalman filtering in two dimensions," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 232–240, Mar. 1972.

[46] J. W. Woods, S. Dravida, and R. Mediavilla, "Image estimation using doubly stochastic Gaussian random field models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, pp. 245–253, Mar. 1987.

[47] G. Zhou and G. B. Giannakis, "Harmonics in gaussian multiplicative and additive noise: Cramer-Rao bounds," *IEEE Trans. Signal Processing*, vol. 43, pp. 1217–1231, May 1995.

Wan-Chieh Pai received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, R.O.C., in 1992 and the M.S. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, in 1995 and 1998, respectively.

In 1998, he joined DIVX, Herndon, VA, to work with the Encription Research Team. Since 1999, he has been with Sony Component Company, San Jose, CA, working with the Consumer Audio Engineering Group. His research interests are in digital signal processing, especially with speech and audio signals.

Peter C. Doerschuk (M'86) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, in 1977, 1979, and 1985, respectively, and the M.D. degree from Harvard University Medical School, Cambridge, in 1987.

After post-graduate training at Brigham and Womens' Hospital, Boston, MA, he held a post-doctoral appointment at the Laboratory for Information and Decision Systems, MIT, from January 1988 to August 1990. Since August 1990, he has been with the faculty of the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN.