Estimation of Modulation Based on FM-to-AM Transduction: Two-Sinusoid Case

Wade P. Torres, *Student Member, IEEE,* and Thomas F. Quatieri, *Fellow, IEEE*

***Abstract*—A method is described for estimating the amplitude modulation (AM) and the frequency modulation (FM) of the components of a signal that consists of two AM–FM sinusoids. The approach is based on the *transduction* of FM to AM that occurs whenever a signal of varying frequency passes through a filter with a nonflat frequency response. The objective is to separate the AM and FM of the sinusoids from the *amplitude envelopes* of the output of two transduction filters, where the AM and FM are nonlinearly combined in the amplitude envelopes. A current scheme is first refined for AM–FM estimation of a single AM–FM sinusoid by iteratively inverting the AM and FM estimates to reduce error introduced in transduction. The transduction filter pair is designed relying on both a time- and frequency-domain characterization of transduction error. The approach is then extended to the case of two AM–FM sinusoids by essentially reducing the problem to two single-component AM–FM estimation problems. By exploiting the beating in the amplitude envelope of each filter output due to the two-sinusoidal input, a closed-form solution is obtained. This solution is also improved upon by iterative refinement. The AM–FM estimation methods are evaluated through an error analysis and are illustrated for a wide range of AM–FM signals.**

***Index Terms*— AM–FM estimation, bandpass amplituse envelope, FM-to-AM transduction, signal separation, time–frequency distribution.**

I. Introduction

AMPLITUDE modulation (AM) and frequency modulation (FM) are found in many naturally occurring signals. For example, AM–FM is an important information carrier in speech and biological signals, characterizing both their resonant and harmonic components. Motivation for the approach of this paper in estimating AM–FM in signals is the hypothesis that the AM–FM of such signals is represented in and estimated from the *amplitude envelopes* of bandpass cochlear filter outputs in the front-end auditory system. We rely on two basic properties of a signal's amplitude envelope: first, that FM is *transduced* to AM by the shape of a bandpass filter, similar to the original Armstrong FM demodulator [1] and, second, that the amplitude envelope of a filter output reflects the *beating* of two sinusoids at the filter input. It has been hypothesized that the auditory system uses both this FM-to-AM transduction and the beating phenomenon for the neural coding of AM–FM in sounds [9], [20], [27].

Manuscript received March 27, 1998; revised April 22, 1999. This work was supported by the Naval Submarine Medical Research Laboratory. The associate editor coordinating the review of this paper and approving it for publication was Prof. Moeness Amin.

The authors are with the Lincoln Laboratory, Massachusetts Institute of Tecnology, Lexington, MA 02420 USA (e-mail: quatieri@ll.mit.edu).

Publisher Item Identifier S 1053-587X(99)08327-0.

In discrete time, the goal of AM–FM estimation is to estimate the amplitude functions $a_i[n]$ and frequency functions $\dot{\theta}_i[n]$ of the composite discrete-time signal

$$s[n] = \sum_{i=1}^{N} a_i[n] \cos(\theta_i[n])$$

where $\theta_i[n]$ denotes the discrete-time phase function, and $\dot{\theta}_i[n]$ is obtained from sampling the derivative of the counterpart continuous-time phase $\theta_i(t)$. AM–FM estimation has been approached from several perspectives of which we describe a few. Perhaps one of the earliest and simplest methods is based on the short-time Fourier transform magnitude (STFTM) squared of a sequence, i.e.,

$$|X(n,\omega)|^2 = \left| \sum_m x[m] w[n-m] e^{j\omega n} \right|^2$$

which measures the energy in $x[n]$ in a time–frequency neighborhood of (n, ω), where $w[n]$ is a sliding short-time window. One AM–FM estimation method, which is sometimes referred to as the "ridge algorithm," computes the instantaneous frequencies in the sequence $x[n]$ from local maxima in $|X(n,\omega)|^2$. The STFTM squared is one of many time–frequency distributions (TFD's) of a signal. A TFD describes a signal simultaneously in the time and frequency domains from which AM and FM components can be estimated from their ridges or other TFD properties [3], [5]. This includes the Wigner distribution (first introduced in the 1930's to calculate the quantum mechanical distributions of position and momentum) and a large related class of TFD's [5]. A method of harmonic ridge tracking has also been used in the context of speech analysis/synthesis [13]. In the 1960's, Costas proposed a different approach to AM–FM estimation referred to as residual signal analysis [6]. This approach uses a set of "trackers": one for each AM–FM component contained in the signal. At the input to each tracker, the other component estimates are subtracted so that the assigned component to a particular tracker is predominant. The estimates in Costas' original system are calculated by a linear projection of amplitude and phase. Recently, a more elaborate projection method has improved upon Costas' original system [18]. Another approach uses a nonlinear operator referred to as the Teager energy operator for AM–FM estimation [12]. Although derived for a single AM–FM sine, this method has been used in speech analysis by first separating sines. One approach for sine separation in this application iteratively removes speech resonant contributions [10]. An alternate approach to the

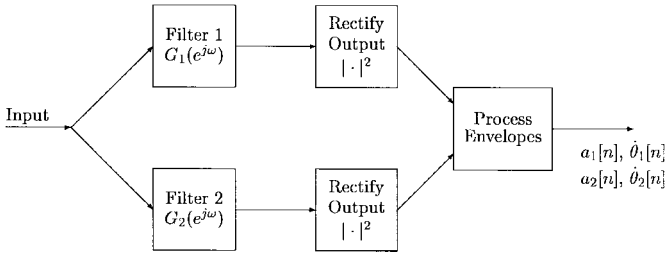


Fig. 1. Approach to two-component AM–FM estimation.

initial sine separation formulates the problem as a system of linear equations [21]. The solution to the system of equations, which have additional constraint equations relating to the periodicity of the components, is the solution to an augmented least-squares problem. The estimates of the AM–FM sine components are then processed with the Teager energy-based algorithm to obtain the AM–FM estimates of each component. Other generalizations of Teager energy-based AM–FM estimation have also been recently proposed for the multicomponent problem [11].

An approach to AM–FM estimation of a single-component AM–FM signal exploited within this paper is based on the transduction of FM to AM by linear filters. FM is transduced to AM whenever an FM signal passes through a filter with a nonflat frequency response; as the instantaneous frequency of the FM moves across the nonflat passband of the filter, a change in the output amplitude envelope occurs. Using two overlapping transduction filters, it is possible to estimate the AM and FM of a single-component AM–FM signal using only the amplitude envelopes of the filter outputs [14], [17]. In this paper, we generalize this approach to signals that consists of two AM–FM sinusoids.

A two-component AM–FM signal is given in discrete time by

$$s[n] = a_1[n]\cos\left(\theta_1[n]\right) + a_2[n]\cos(\theta_2[n])$$

where $a_1[n]$ and $a_2[n]$ are the AM functions, $\theta_1[n]$ and $\theta_2[n]$ are the phase functions, and $\dot{\theta}_1[n]$ and $\dot{\theta}_2[n]$ are the FM functions. The structure of our two-component estimation approach is shown in Fig. 1. After passing the signal through the pair of transduction filters $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$, the amplitude envelopes of the filter outputs are computed; this operation is denoted in Fig. 1 as the rectify stage. The amplitude envelopes themselves each contain a single-component AM–FM signal due to beating between the two sinusoidal components within each filter; each filter output envelope is a function of $a_1[n]$, $a_2[n]$, $\dot{\theta}_1[n]$, and $\dot{\theta}_2[n]$ and has frequency $\dot{\theta}_2[n] - \dot{\theta}_1[n]$. By applying the single-component AM–FM estimation algorithm to the envelopes, we can obtain the AM and FM of each component in the input signal. Two sources of error are introduced in the estimation process. In the single-component algorithm, the error is due to the inexactness of the assumed transduction process, which is referred to as *transduction error* [17]. Because the single-component algorithm is embedded in the two-component algorithm, the latter suffers from this error as well. In addition, the two-component algorithm has error due to the inexactness of each filter's envelope output representation as a "pure" AM–FM component. Iterative methods are introduced to reduce both forms of error.

The paper is organized as follows. In Section II, we discuss considerations of uniqueness in AM–FM representations of one- and two-component signals and review the related concept of the analytic signal. In Section III, we describe the concept of FM-to-AM transduction and describe a transduction approximation in terms of amplitude envelopes and linear, time-invariant filters when the input is a single-component AM–FM signal. We then review the AM–FM estimation algorithm, which is based on FM-to-AM transduction and applied to a single-component signal, and describe its implementation used in the remainder of the paper. A frequency-domain characterization of the transduction error is used as a basis for obtaining the filter pair for transduction. Section III ends with a technique that improves performance of this algorithm by iteratively inverting the modulation to reduce transduction error. In Section IV, the algorithm for estimating the AM and FM of a signal composed of two AM–FM components is presented, and refinements are then made based on inverse modulation techniques conceptually similar to those of Section III. Section V concludes with a summary, including a brief discussion on the paper's generalization to more than two AM–FM components.

II. Preliminaries

Consider a real signal $s[n]$ that we desire to represent in the form

$$s[n] = a[n]\cos(\theta[n])$$

where $a[n]$ is assumed to be positive. In finding this representation, we could choose any value for $a[n]$ at each n such that $a[n] \geq |s[n]|$, resulting in $\theta[n]$ equal to $\arccos(s[n]/a[n])$. Because our only constraint is that $a[n] \geq |s[n]|$, there are an infinite number of choices for $a[n]$ and $\theta[n]$ at each sample point. One approach to making the decomposition unique is through the *analytic signal* [5]. An analytic signal is, by definition, a signal that has no negative frequency components. Under certain conditions, a real signal $s[n]$ has an analytic signal counterpart $s_a[n]$ such that if

$$s[n] = a[n]\cos(\theta[n]) \tag{1}$$

then

$$s_a[n] = a[n]e^{j\theta[n]}. \tag{2}$$

The conditions under which (1) and (2) hold are as follows [3].

1) The Fourier transform of $a[n]$ lies entirely in the region $|\omega| < \omega_0$ for some ω_0.
2) The Fourier transform of $\cos(\theta[n])$ exists only outside of this region.

We assume that these conditions hold. Defined in this manner, the analytic signal provides a unique expression for $s[n]$ as an AM–FM sinusoid where the amplitude function is $a[n] = |s_a[n]|$ and $\theta[n] = -j\ln(s_a[n]/a[n])$, which is unique (modulo 2π). Although this uniqueness property can

be established for any complex signal representation of $s[n]$, Vakman [26] showed that the analytic signal is the only complex signal representation that simultaneously satisfies the following desirable properties.

1) The amplitude envelope is continuous and differentiable (in continuous time).
2) The phase and frequency are not affected by scalar multiplication of the signal.
3) If the real signal has constant amplitude and frequency, the complex signal also has constant amplitude and frequency.

The analytic signal also allows a definition of instantaneous frequency as the derivative of the phase of the analytic signal[1] [5].

In generalizing the representation problem to writing $s[n]$ as a sum of two AM–FM sinusoids of the form of (1),[2] it is again possible to argue that there are an infinite number of ways to decompose $s[n]$ in this form [25]. However, we will see that our scheme of reducing the two-component problem into two one-component problems leads naturally to uniqueness of the AM–FM components.

III. Single-Component AM–FM Estimation

We first describe FM-to-AM transduction and then review and further develop the single-component AM–FM estimation algorithm based on FM-to-AM transduction [17]. The method of *inverse modulation* is proposed for improving the estimation for sinusoids with large AM and FM.

A. FM-to-AM Transduction

FM-to-AM transduction occurs when an FM signal passes through a filter that has a nonflat spectral shape. As the frequency of the signal sweeps across the passband of the filter, the amplitude envelope of the filter output changes. The amplitude envelope of the filter output is therefore a function of both the AM and the FM of the filter input. For an input of the form $s[n] = a[n]\cos(\theta[n])$, the output of a linear time-invariant (LTI) filter can be approximated as [4], [17]

$$y[n] \approx a[n]e^{j\theta[n]}H\left(e^{j\dot{\theta}[n]}\right) = s[n]H\left(e^{j\dot{\theta}[n]}\right). \tag{3}$$

It is assumed that $H(\omega) = 0$ for $\omega < 0$, making the resulting signal analytic. If $s[n]$ has neither AM nor FM, then the approximation (3) is exact.

B. Algorithm

For a single-component AM–FM sinusoid, there are two parameters being estimated at each time point $a[n]$ and $\dot{\theta}[n]$. By passing the signal through two filters and using the transduction approximation (3), we have two equations and two unknowns at each time sample. The filters can be chosen so that the equations are readily solved. In fact, it is only the *relationship* between the two filters that establishes a unique solution. The actual filter shape is chosen to minimize the error in the approximation (3).

Suppose we choose the filters $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$ to have the relation

$$G_2(e^{j\omega}) = \omega G_1(e^{j\omega}). \tag{4}$$

We assume that $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$ are real and nonnegative in frequency, thus having the property of being symmetric and localized around the time origin, as well as leading to a convenient solution formulation. In addition, we assume that $G_1(e^{j\omega})$, and thus, $G_2(e^{j\omega})$ is zero for negative frequency. From (3), the square of the envelopes of the filter outputs can be approximated as

$$|y_1[n]|^2 \approx a^2[n]G_1^2(e^{j\dot{\theta}[n]})$$
$$|y_2[n]|^2 \approx a^2[n]G_2^2(e^{j\dot{\theta}[n]}) = a^2[n]\dot{\theta}^2[n]G_1^2(e^{j\dot{\theta}[n]}).$$

The AM and FM are then estimated by

$$\hat{\dot{\theta}}[n] = \sqrt{\frac{|y_2[n]|^2}{|y_1[n]|^2}} \tag{5}$$

$$\hat{a}[n] = \frac{\sqrt{|y_1[n]|^2}}{G_1(e^{j\hat{\dot{\theta}}[n]})}. \tag{6}$$

C. Filter Choice and Transduction Error

We have the following constraints on the choice of $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$.

1) The filter outputs must be analytic to provide a unique AM–FM decomposition.
2) $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$ should be chosen to reduce the transduction error in (3).

The first constraint is satisfied by virtue of $G_1(e^{j\omega})$ being zero for $-\pi < \omega < 0$, i.e., the Hilbert transform operation is embedded within $G_1(e^{j\omega})$. To satisfy the second constraint, we analyze the transduction error from both a time-domain and a frequency-domain perspective.

1) Time Domain: In the time domain, error bounds for the transduction approximation have been derived by Bovik *et al.* [4]. Let $z[n]$ be the output of an LTI filter with frequency response $H(e^{j\omega})$ and impulse response $h[n]$ for input $s[n]$, i.e., $z[n] = s[n] * h[n]$, and let $y[n]$ be given by the transduction approximation (3). Then, the error defined as $\varepsilon[n] = |z[n] - y[n]|$ is bounded by

$$\varepsilon[n] \le \sum_{\substack{p\in\mathbb{Z}\\ p\neq 0}} |h[p]| \int_{n-p}^{n} \left(|\dot{a}(v)| + a_{\max}|p||\ddot{\theta}(v)|\right) dv \tag{7}$$

where $a_{\max} = \max_n a[n]$, and $a(v)$ and $\theta(v)$ are the continuous time signals corresponding to $a[n]$ and $\theta[n]$. The above upper bound suggests that a filter with small transduction error has energy in its impulse response concentrated about $n = 0$. For our objective of reducing transduction error in AM–FM

[1] For some signals, this definition of instantaneous frequency matches our intuition. Many signals, however, produce "paradoxical" results, in particular, signals that have multiple components [5]. To avoid problems with this definition, we redefine instantaneous frequency to be a *set* of phase derivatives with each phase derivative corresponding to a particular AM–FM component of the signal.

[2] Determining the appropriate number of sinusoids is a rather complex issue and not the topic of this paper. Therefore, we assume that the number of sinusoids is known *a priori* and that the components are always present.

estimation, we will now see, however, the importance of both a time- and a frequency-domain characterization of the error analysis.

2) Frequency Domain: Our frequency-domain approach to characterizing transduction error is similar to the "quasistationary" method used in early work on FM communication systems [22]. Consider a filter with a frequency response

$$H(e^{j\omega}) = \begin{cases} H_+(e^{j\omega}), & 0 \le \omega < \pi \\ 0, & -\pi \le \omega < 0 \end{cases} \tag{8}$$

where $H_+(e^{j\omega})$ is some arbitrary function. We assume that over the frequency range of the input signal, we can represent $H_+(e^{j\omega})$ as a sum of polynomials, i.e.,

$$H_+(e^{j\omega}) = \sum_{k=0}^{\infty} \omega^k f_k. \tag{9}$$

Since the output of a filter with shape ω^k is $(-j)^k(d^k/dn^k)x[n]$ [15], we can analyze the transduction error explicitly by looking at the transduction error that corresponds to each term in (9).[3]

For an input of the form $x[n] = a[n]e^{j\theta[n]}$, we have the filter outputs as shown in (9a) at the bottom of the page. We now make the following observations:

1) As the number of high-order terms in (9) decreases, the transduction error, in general, becomes less severe.
2) If $H(e^{j\omega}) = 1$ over the frequency range of the signal, there is no transduction error.
3) If $H(e^{j\omega}) = c\omega + d$ over the frequency range of the signal, there is no transduction error due to frequency modulation.

A few comments are in order with regard to the two different time- and frequency-domain perspectives on transduction error. Consider, as an example, the Hilbert transformer filter used to obtain an analytic signal. This method does not suffer from any transduction error, yet the time-domain analysis would lead us to believe that the Hilbert transformer suffers from significant transduction error. It is important, therefore, to use frequency-domain transduction error analysis in designing transduction filters. Nevertheless, we still desire that the impulse response of the filters remains short so that the AM–FM estimation algorithm operates on only a local portion of the signal. This condition will minimize transient effects due to any nonsmooth modulating functions that the signal may contain.

From the above analysis, the filters should have short impulse responses and have few high-order coefficients in their polynomial expansion.[4] This type of problem, i.e., optimizing with a tradeoff between multiple objectives, can be solved by an optimization technique known as the goal attainment method [8]. The goal attainment method solves the problem given by

$$\underset{\mathbf{c},\gamma}{\text{minimize}}\ \gamma \qquad \text{such that} \qquad \mathbf{F}(\mathbf{c}) - \mathbf{w}\gamma \le \mathbf{g} \tag{10}$$

where

$\mathbf{F}$ = vector of costs;
$\mathbf{w}$ = vector of weights;
$\mathbf{g}$ = vector of goals;
$\mathbf{c}$ = vector of unknown parameters;
γ = scale factor.

The advantage in using the goal attainment method is that it allows the objectives to be over- and under-achieved. For our filter design problem, our objectives[5] are

$$F_1(\mathbf{c}) = \min \sum_{p=R}^{\infty} c_p^2$$

$$F_2(\mathbf{c}) = \min \sum_{n=N+1}^{\infty} g_1^2[n]$$

$$F_3(\mathbf{c}) = \min \sum_{n=N+1}^{\infty} g_2^2[n]$$

[3] $(d^k/dn^k)x[n]$ corresponds to samples of $(d^k/dt^k)x(t)$, where $x(t)$ is a bandlimited continuous-time signal corresponding to $x[n]$.

[4] Prolate spheroidal wave functions satisfy a similar constraint in the time domain but an energy concentration constraint in frequency [23].

[5] Adding the constraints $F_5(c) = G_1^2(e^{j0}) + G_1^2(e^{j\pi}) = 0$ and $F_6(c) = ((d/d\omega)G_1(e^{j\omega})|_{\omega=0})^2 + ((d/d\omega)G_1(e^{j\omega})|_{\omega=\pi})^2 = 0$ significantly improved the performance because these constraints aid in making the frequency response "smooth." This reduces the transient response in the frequency domain when the impulse response is truncated to obtain an FIR filter.

$$\begin{aligned}
&\underline{\text{Filter}} && \underline{\text{Output}} \\
&1 \leftrightarrow y[n] = a[n]e^{j\theta[n]} \\
&\omega \leftrightarrow y[n] = \underbrace{\dot{\theta}[n]a[n]e^{j\theta[n]}}_{\substack{\text{Transduction} \\ \text{Approximation}}} - \underbrace{j\dot{a}[n]e^{j\theta[n]}}_{\varepsilon_{1[n]}} \\
&\omega^2 \leftrightarrow y[n] = \underbrace{\dot{\theta}^2[n]a[n]e^{j\theta[n]}}_{\substack{\text{Transduction} \\ \text{Approximation}}} - \underbrace{j\Big(\ddot{\theta}[n]a[n] + \dot{\theta}[n]\dot{a}[n]\Big)e^{j\theta[n]} - j\frac{d}{dn}\varepsilon_1[n]}_{\varepsilon_2[n]} \\
&\qquad\vdots \\
&\omega^n \leftrightarrow y[n] = \underbrace{\dot{\theta}^n[n]a[n]e^{j\theta[n]}}_{\substack{\text{Transduction} \\ \text{Approximation}}} - \underbrace{j\Big((n-1)\dot{\theta}^{n-2}[n]\ddot{\theta}[n]a[n] + \dot{\theta}^{n-1}[n]\dot{a}[n]\Big)e^{j\theta[n]} - j\frac{d}{dn}\varepsilon_{n-1}[n]}_{\varepsilon_n[n]}
\end{aligned} \tag{9a}$$

subject to the constraint

$$F_4(\mathbf{c}) = \|g_1^2[n]\|_2 = 1$$

with $\mathbf{F}(\mathbf{c}) = [F_1(\mathbf{c}) \; F_2(\mathbf{c}) \; F_3(\mathbf{c}) \; F_4(\mathbf{c})]$ and where $G_1(\omega) = \sum_{p=0}^{P} c_p(\omega - \pi/2)^p$, $g_1[n] = \int_0^{\pi} G_1(\omega)e^{-j\omega n}\,d\omega$, and $g_2[n] = \int_0^{\pi} \omega G_1(\omega)e^{-j\omega n}\,d\omega$. Given the symmetry of $g_1[n]$ and $g_2[n]$, only positive time values of these responses are used in the minimization. The summation lower bounds are selected as $R = 5$, which will concentrate the energy in the first five polynomial coefficients, and $N = 3$, which concentrates the energy of the impulse response in seven samples. We also chose $P = 20$ because using a higher order polynomial for $G_1(\omega)$ makes the optimization too computationally intensive. We want to satisfy F_1, F_2, and F_3 while forcing $F_4 = 1$. Therefore, our goal vector is

$$\mathbf{g} = [0\ 0\ 0\ 1].$$

The weight vector $\mathbf{w}$ specifies the relative degree of the under- or over-achievement of the goals. We chose the weight vector to be

$$\mathbf{w} = [1\ 1\ 1\ 0].$$

Referring back to (10), for $\mathbf{w} = [w_1 \; w_2 \; w_3 \; w_4]$, $w_4 = 0$ will force F_4 to achieve the value of 1. Choosing $w_1 = w_2 = w_3 = 1$ will cause the corresponding objectives to be under-achieved to an equal degree. The goal attainment method will find a solution that minimizes γ, which, again referring to (10), means that F_1, F_2, and F_3 are as close to zero as possible, and $F_4 = 1$. By using the goal attainment method, we have gained the ability to incorporate hard constraints and to trade off between conflicting goals in our optimization problem.

The frequency responses of the filters obtained using the goal attainment method are shown in Fig. 2(a). In Fig. 2(b) and (c), the energy of the impulse responses is seen to be concentrated in seven samples around $n = 0$. The coefficients of the polynomial of expansion of $G_1(\omega)$ about $\omega = \pi/2$ is shown in Fig. 3, where it can be seen that only the first five coefficients are significant. It is important to again observe that the "transduction gain," i.e., the extent to which FM is transduced to AM, comes from the relation between the two filters $G_2(e^{j\omega}) = \omega G_1(e^{j\omega})$ and not the shape of the filters. This is one of the strengths of the algorithm because it allows for the choice of filters that minimize transduction error, with the above filter relation preventing both filters from being flat in the frequency domain.

C. Inverse Modulation

The performance of the algorithm presented in the previous sections degrades as the rate and extent of AM and FM increases because as the modulation is increased, the transduction approximation becomes less accurate. To reduce the extent of the AM and FM, we use the estimates obtained from the AM–FM estimation algorithm of the previous section to invert the modulation and then reapply the algorithm to the new demodulated signal. Since the initial estimates are not exact, the demodulated signal still contains some modulation. Iterating this procedure, however, reduces the remaining modulation. The block diagram of this system is shown in Fig. 4.

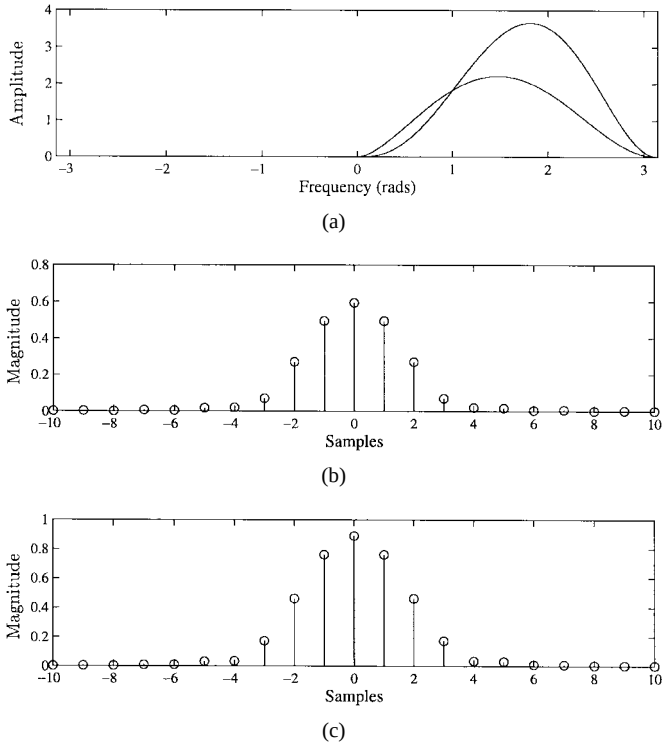


Fig. 2. Frequency and impulse response of $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$. (a) Frequency response, $G_1(e^{j\omega})$ (lower), $G_2(e^{j\omega})$ (upper). (b) Magnitude of the impulse response of $g_1[n]$. (c) Magnitude of the impulse response of $g_2[n]$.

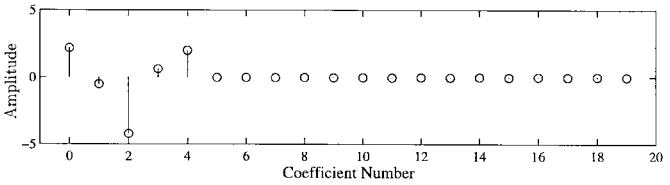


Fig. 3. Coefficients of the polynomial expansion of $G_1(\omega)$.

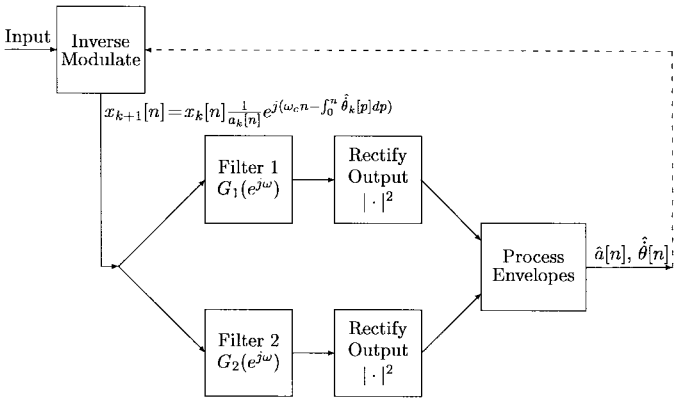
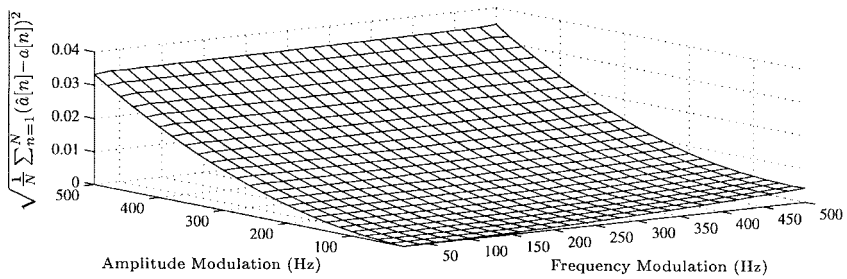
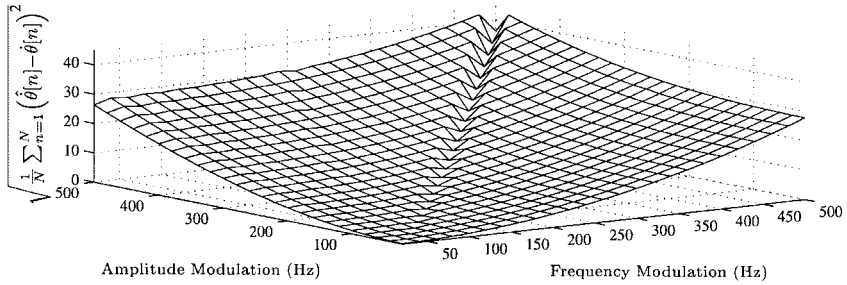


Fig. 4. Block diagram of the single-component AM-FM estimation algorithm with iterative refinement.

We denote by $x_k[n]$ the input on the kth iteration and by $\hat{a}_k[n]$ and $\hat{\theta}_k[n]$ the amplitude and frequency estimates corresponding to $x_k[n]$. On the initial iteration, $x_0[n] = x[n]$,



(a)



(b)

Fig. 5. Average root-mean-square error in amplitude and frequency estimates of transduction-based AM–FM estimation algorithm. The frequency is in Hertz with a 10 000-Hz sampling rate.

$\hat{a}_0[n] = 1$, and $\hat{\dot{\theta}}_0[n] = 0$. The inverse modulation procedure is described by

$$x_{k+1}[n] = \frac{1}{\hat{a}_k[n]}\, x_k[n] e^{-j\int_0^n \hat{\dot{\theta}}_k[p]dp} e^{j\omega_c n}$$
$$= \left(\prod_{i=1}^{k} \frac{a[n]}{\hat{a}_i[n]}\right) e^{j\left(\theta[n] - \sum_{m=1}^{k}\int_0^n \hat{\theta}_m[p]\,dp\right)} \cdot e^{jk\omega_c n}$$

where the term $e^{j\omega_c n}$ is necessary to shift the spectrum of the signal back up to the passband of the filters $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$. The overall phase and amplitude estimates after k iterations are given by

$$\hat{\dot{\theta}}[n] = \sum_{i=1}^{k} \hat{\dot{\theta}}_i[n] - k\omega_c$$

and

$$\hat{a}[n] = \prod_{i=1}^{k} \hat{a}_i[n].$$

The procedure can be summarized as follows.

1) Estimate the modulation of $x_k[n]$ with the estimation algorithm of previous sections.
2) Divide $x_k[n]$ by the amplitude estimate $a_k[n]$.
3) Obtain a phase estimate $\hat{\theta}_k[n]$ by numerically integrating $\hat{\dot{\theta}}_k[n]$ using the trapezoidal rule [24].
4) Invert the FM and multiply by $e^{-j\hat{\theta}_k[n]}$. For real $x[n]$, the FM inversion, however, is constrained so that the negative frequencies of the real signal are not modulated to positive frequencies as described in step (5).
5) Modulate $x_k[n]$ back up to the passband of the filters $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$. Because the filters were optimized about $\pi/2$, the modulation frequency ω_c is selected as close as possible to $\pi/2$ without modulating negative frequencies to positive frequencies (see the Appendix).
6) Stop at the current iteration or repeat the procedure with the new "inverse modulated" signal.

In steps (4) and (5), it is important to observe that even if a portion of the negative frequency spectrum does get modulated to positive frequencies (see the Appendix), the algorithm still produces useful results. This is because the filters $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$ are close to zero near $-\pi$ and 0. Any negative-frequency spectral components that are modulated to the low or high positive frequency range are approximately filtered out.

D. Examples

To quantify the performance of the single-component AM–FM estimation algorithm, we show the error for a group of 625 signals which are described as

$$x_{k,l}[n] = \left[1 + 0.5\sin\left(\frac{2\pi k n}{500}\right)\right] \cdot \cos\left(\left(\frac{\pi n}{2} + \frac{50}{l}\cos\left(\frac{2\pi l n}{500}\right)\right)\right) \quad (11)$$

where $k = 2, 4, \ldots, 50$, and $l = 2, 4, \ldots, 50$. If these signals were obtained from sampling continuous-time signals at a rate of 10 000 samples/s, the continuous-time set of signals would have sinusoidal amplitude modulation varying from 0.5 to 1.5 over a range of frequencies from 20–500 Hz. The FM varies from 2000–3000 Hz around a carrier of 2500 Hz, and the rate of the frequency modulation ranges from 20–500 Hz. The root-mean-square error is shown in Fig. 5. Plots of the estimates are given in [25].

In comparing performance with the iterative inverse modulation method, we used the same signal set described in (11)

TABLE I
AVERAGE OF THE ROOT-MEAN-SQUARE ERROR OVER THE SIGNAL SET IN (11) USING THE ITERATIVE INVERSE MODULATION METHOD. (FREQUENCY ERROR CALCULATIONS DONE IN TERMS OF RADIANS)

Iteration:	1	2	3	4	5
Avg. Mean Sq. AM Estimate Error	1.27×10^{-2}	6.29×10^{-3}	4.96×10^{-3}	3.59×10^{-3}	2.75×10^{-3}
Avg. Mean Sq. FM Estimate Error	1.09×10^{-2}	4.58×10^{-3}	3.95×10^{-3}	2.84×10^{-3}	2.19×10^{-3}

and found that the iterative method results in error reduction by a factor of 10. Table I lists the root-mean-square error averaged over the signal set after 1–5 iterations.

E. Additive Noise

If the input signal is corrupted with another signal, then the frequency estimate becomes

$$\hat{\dot{\theta}}[n] = \frac{\left|a[n]\dot{\theta}[n]G_1(e^{j\dot{\theta}[n]})e^{j\theta[n]} + \mathbf{g}_2^t\mathbf{w}\right|}{\left|a[n]G_1(e^{j\dot{\theta}[n]})e^{j\theta[n]} + \mathbf{g}_1^t\mathbf{w}\right|} \tag{12}$$

where "t" denotes matrix transpose, and where $\mathbf{g}_1$ and $\mathbf{g}_2$ are vectors denoting the impulse responses of $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$, and $\mathbf{w}$ is a vector denoting a disturbance signal. The impulse responses are assumed to fall in the interval $[-N, N]$, and thus, all vectors are of length $2N+1$.

To more simply quantify the performance of the algorithm in the presence of a disturbance, we assume an amplitude-bounded disturbance vector and derive upper and lower bounds on the frequency estimate. Explicitly, we assume

$$\max_n |w[n]| < \frac{\epsilon}{||\mathbf{g}_1||_1}$$

where $||\mathbf{g}_1||_1 = \sum_{n=-N}^{N} |g_1[n]|$. Equation (12) is upper bounded by

$$\begin{aligned}
&\hat{\dot{\theta}}[n] \\
&\le \frac{\left|a[n]\dot{\theta}[n]G_1(e^{j\dot{\theta}[n]})e^{j\theta[n]} + \dot{\theta}[n]\mathbf{g}_1^t\mathbf{w}\right| + \left|(\mathbf{g}_2^t - \dot{\theta}[n]\mathbf{g}_1^t)\mathbf{w}\right|}{\left|a[n]G_1(e^{j\dot{\theta}[n]})e^{j\theta[n]} + \mathbf{g}_1^t\mathbf{w}\right|} \\
&= \dot{\theta}[n] + \frac{\left|(\mathbf{g}_2^t - \dot{\theta}[n]\mathbf{g}_1^t)\mathbf{w}\right|}{\left|a[n]G_1(e^{j\dot{\theta}[n]})e^{j\theta[n]} + \mathbf{g}_1^t\mathbf{w}\right|} \\
&\le \dot{\theta}[n] + \frac{\left|(\mathbf{g}_2^t - \dot{\theta}[n]\mathbf{g}_1^t)\mathbf{w}\right|}{\left|a[n]G_1(e^{j\dot{\theta}[n]})\right| - |\mathbf{g}_1^t\mathbf{w}|}
\end{aligned} \tag{13}$$

where, to avoid a singularity in (13), we assume that the disturbance is small enough so that

$$\left|\mathbf{g}_1^t\mathbf{w}\right| < \left|a[n]G_1(e^{j\omega})\right|. \tag{14}$$

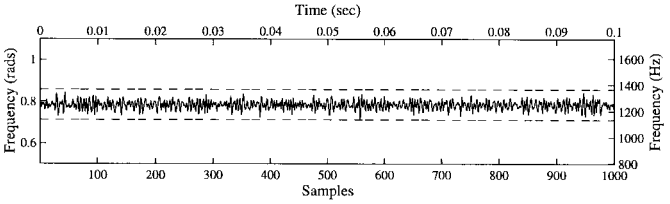


Fig. 6. Frequency estimate of a constant amplitude, constant frequency signal with a disturbance vector of amplitude corresponding to $\epsilon = 0.1$ (SNR $= 15$ dB). The upper and lower bounds are illustrated by dashed lines.

The upper bound can be written in terms of ϵ by first writing

$$\begin{aligned}
|\mathbf{g}_i^t\mathbf{w}| &= \left|\sum_{n=-N}^{N} g_i[n]w[n]\right| \\
&\le \sum_{n=-N}^{N} |g_i[n]w[n]| \\
&= \sum_{n=-N}^{N} |g_i[n]||w[n]| \\
&\le \frac{\epsilon}{||\mathbf{g}_1||_1} \sum_{n=-N}^{N} |g_i[n]| \\
&= \epsilon\frac{||\mathbf{g}_i||_1}{||\mathbf{g}_1||_1}.
\end{aligned}$$

The upper bound then becomes

$$\hat{\dot{\theta}}[n] \le \dot{\theta}[n] + \frac{\epsilon||\mathbf{g}_2 - \dot{\theta}[n]\mathbf{g}_1||_1/||\mathbf{g}_1||_1}{a[n]G_1(e^{j\dot{\theta}[n]}) - \epsilon}.$$

A lower bound derived in a similar manner is given by

$$\hat{\dot{\theta}}[n] \ge \dot{\theta}[n] - \frac{\epsilon||\mathbf{g}_2 - \dot{\theta}[n]\mathbf{g}_1||_1/||\mathbf{g}_1||_1}{a[n]G_1(e^{j\dot{\theta}[n]}) - \epsilon}.$$

These bounds are somewhat loose and, as can be seen in (13), become less tight as $\dot{\theta}[n]$ becomes large. An example of the frequency estimate of a constant amplitude, constant frequency signal, when there is a disturbance vector with a maximum amplitude corresponding to $\epsilon = 0.10$ (SNR $= 15$ dB), is shown in Fig. 6. The upper and lower bounds are illustrated by dashed lines.

IV. TWO-COMPONENT AM–FM ESTIMATION

We begin this section with showing that with the amplitude envelopes of the outputs of filters $G_1(\omega)$ and $G_2(\omega)$, it is possible to essentially reduce the two-component estimation problem to two single-component estimation problems and that this allows a unique solution. The resulting two single sinusoids represent the beating that occurs in each filter's output between the two sinusoids in the input. This beating phenomenon is exploited to form estimates of the AM and FM of the input components.

A. Reduction to Two Single-Component Problems

Consider a signal of the form

$$x[n] = a_1[n]\cos(\theta_1[n]) + a_2[n]\cos(\theta_2[n]). \tag{15}$$

If $x[n]$ is passed through the filters $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$, the filter outputs can be approximated as

$$\begin{aligned} y_1[n] \approx\, & a_1[n]G_1\left(e^{j\dot{\theta}_1[n]}\right)e^{j\theta_1[n]} \\ & + a_2[n]G_1\left(e^{j\dot{\theta}_2[n]}\right)e^{j\theta_2[n]} \end{aligned} \tag{16}$$

and

$$\begin{aligned} y_2[n] \approx\, & a_1[n]G_2\left(e^{j\dot{\theta}_1[n]}\right)e^{j\theta_1[n]} \\ & + a_2[n]G_2\left(e^{j\dot{\theta}_2[n]}\right)e^{j\theta_2[n]} \\ =\, & a_1[n]\dot{\theta}_1[n]G_1\left(e^{j\dot{\theta}_1[n]}\right)e^{j\theta_1[n]} \\ & + a_2[n]\dot{\theta}_2[n]G_1\left(e^{j\dot{\theta}_2[n]}\right)e^{j\theta_2[n]}. \end{aligned} \tag{17}$$

The square of the amplitude envelopes of the filter outputs are given by

$$\begin{aligned} |y_1[n]|^2 \approx\, & a_1^2[n]G_1^2\left(e^{j\dot{\theta}_1[n]}\right) + a_2^2[n]G_1^2\left(e^{j\dot{\theta}_2[n]}\right) \\ & + 2a_1[n]a_2[n]G_1\left(e^{j\dot{\theta}_1[n]}\right)G_1\left(e^{j\dot{\theta}_2[n]}\right) \\ & \cdot\cos(\theta_2[n] - \theta_1[n]) \end{aligned} \tag{18}$$

and

$$\begin{aligned} |y_2[n]|^2 \approx\, & a_1^2[n]G_2^2\left(e^{j\dot{\theta}_1[n]}\right) \\ & + a_2^2[n]G_2^2\left(e^{j\dot{\theta}_2[n]}\right) + 2a_1[n]a_2[n] \\ & \cdot G_2\left(e^{j\dot{\theta}_1[n]}\right)G_2\left(e^{j\dot{\theta}_2[n]}\right)\cos(\theta_2[n] - \theta_1[n]) \\ =\, & a_1^2[n]\dot{\theta}_1^2[n]G_1^2\left(e^{j\dot{\theta}_1}[n]\right) + a_2^2[n]\dot{\theta}_2^2[n]G_1^2\left(e^{j\dot{\theta}_2}[n]\right) \\ & + 2a_1[n]a_2[n]\dot{\theta}_1[n]\dot{\theta}_2[n]G_1\left(e^{j\dot{\theta}_1[n]}\right)G_1\left(e^{j\dot{\theta}_2[n]}\right) \\ & \cdot\cos(\theta_2[n] - \theta_1[n]). \end{aligned} \tag{19}$$

Treating the first two terms of $|y_1[n]|^2$ and $|y_2[n]|^2$ as "noise,"[6] and denoting them as

$$\nu_1[n] = a_1^2[n]G_1^2\left(e^{j\dot{\theta}_1[n]}\right) + a_2^2[n]G_1^2\left(e^{j\dot{\theta}_2[n]}\right) \tag{20}$$

$$\nu_2[n] = a_1^2[n]\dot{\theta}_1^2[n]G_1^2\left(e^{j\dot{\theta}_1[n]}\right) + a_2^2[n]\dot{\theta}_2^2[n]G_1^2\left(e^{j\dot{\theta}_2[n]}\right) \tag{21}$$

we are left with two single AM–FM sinusoids with additive noise

$$\begin{aligned} |y_1[n]|^2 \approx\, & 2a_1[n]a_2[n]G_1\left(e^{j\dot{\theta}_1[n]}\right)G_1\left(e^{j\dot{\theta}_2[n]}\right) \\ & \cdot\cos(\theta_2[n] - \theta_1[n]) + \nu_1[n] \end{aligned}$$

and

$$\begin{aligned} |y_2[n]|^2 \approx\, & 2a_1[n]a_2[n]\dot{\theta}_1[n]\dot{\theta}_2[n]G_1\left(e^{j\dot{\theta}_1[n]}\right)G_1\left(e^{j\dot{\theta}_2[n]}\right) \\ & \cdot\cos(\theta_2[n] - \theta_1[n]) + \nu_2[n]. \end{aligned}$$

We can think of the envelopes $|y_1[n]|^2$ and $|y_2[n]|^2$ as sinusoids due to the *beating* between the two filter input components. The AM envelope of these sinusoidal components is determined in part by the cross product of the original envelopes, and the FM of each component is given by the difference, or spacing, between the input frequencies.[7]

B. AM–FM Separation

Our goal now is to separate from $|y_1[n]|^2$ and $|y_2[n]|^2$ the AM and FM of each sinusoidal component of the input. Because $|y_1[n]|^2$ and $|y_2[n]|^2$ are essentially single-component functions, they can be processed with the single-component AM–FM estimation method. With $|y_1[n]|^2$ passed to the single-component estimation algorithm, the algorithm gives a unique amplitude and frequency estimate because the signal is made analytic. The unique amplitude function is given by

$$\hat{a}_{|y_1|^2}[n] = 2a_1[n]a_2[n]G_1\left(e^{j\dot{\theta}_1[n]}\right)G_1\left(e^{j\dot{\theta}_2[n]}\right) \tag{22}$$

and the unique instantaneous frequency estimate by

$$\hat{\dot{\theta}}_{|y_1|^2}[n] = \dot{\theta}_2[n] - \dot{\theta}_1[n]. \tag{23}$$

Similarly, for $|y_2[n]|^2$, we get a unique amplitude estimate

$$\hat{a}_{|y_2|^2}[n] = 2a_1[n]a_2[n]\dot{\theta}_1[n]\dot{\theta}_2[n]G_1\left(e^{j\dot{\theta}_1[n]}\right)G_1\left(e^{j\dot{\theta}_2[n]}\right)$$

and the same instantaneous frequency estimate

$$\hat{\dot{\theta}}_{|y_2|^2}[n] = \dot{\theta}_2[n] - \dot{\theta}_1[n].$$

We can now work with these AM and FM functions to determine the AM and FM of the input components.

The product of the individual instantaneous frequencies can be obtained from

$$\dot{\theta}_2[n]\dot{\theta}_1[n] = \frac{\hat{a}_{|y_2|^2}[n]}{\hat{a}_{|y_1|^2}[n]}. \tag{24}$$

Solving for $\dot{\theta}_2[n]$ in (23) and substituting this into (24) gives

$$\dot{\theta}_1[n]\left(\dot{\theta}_1[n] + \hat{\dot{\theta}}_{|y_1|^2}[n]\right) = \frac{\hat{a}_{|y_2|^2}[n]}{\hat{a}_{|y_1|^2}[n]}$$

$$\dot{\theta}_1^2[n] + \hat{\dot{\theta}}_{|y_1|^2}[n]\dot{\theta}_1[n] - \frac{\hat{a}_{|y_2|^2}[n]}{\hat{a}_{|y_1|^2}[n]} = 0$$

which has two possible solutions

$$\dot{\theta}_1[n] = \frac{1}{2}\left(-\hat{\dot{\theta}}_{|y_1|^2}[n] \pm \sqrt{\hat{\dot{\theta}}^2_{|y_1|^2}[n] + 4\frac{\hat{a}_{|y_2|^2}[n]}{\hat{a}_{|y_1|^2}[n]}}\right).$$

Since the desired frequency is positive, the correct solution is

$$\dot{\theta}_1[n] = \frac{1}{2}\left(-\hat{\dot{\theta}}_{|y_1|^2}[n] + \sqrt{\hat{\dot{\theta}}^2_{|y_1|^2}[n] + 4\frac{\hat{a}_{|y_2|^2}[n]}{\hat{a}_{|y_1|^2}[n]}}\right). \tag{25}$$

Substituting this result into (23), we obtain

$$\dot{\theta}_2[n] = \frac{1}{2}\left(\hat{\dot{\theta}}_{|y_1|^2}[n] + \sqrt{\hat{\dot{\theta}}^2_{|y_1|^2}[n] + 4\frac{\hat{a}_{|y_2|^2}[n]}{\hat{a}_{|y_1|^2}[n]}}\right). \tag{26}$$

[6] In Section IV-B, we assume these noise terms are negligible. We will discuss when this treatment is valid in Section IV-C.

[7] If the crossterms are not present, the equations are not meaningful; we would want to detect this situation and fall back to the single-component algorithm.

This solution always assigns the lower valued frequency estimate to $\dot{\theta}_1[n]$ and the higher valued frequency estimate to $\dot{\theta}_2[n]$.[8] It is now possible to estimate the AM functions using the instantaneous frequency estimates and the magnitude of the filter outputs. First, we multiply $|y_1[n]|^2$ of (18) by $\dot{\theta}_1\dot{\theta}_2$ and then subtract the result from $|y_2[n]|^2$ of (19), which results in

$$\begin{aligned}|y_2[n]|^2 - \dot{\theta}_1[n]\dot{\theta}_2[n]|y_1[n]|^2 &= a_1^2[n]G_1^2\left(e^{j\dot{\theta}_1[n]}\right)\left(\dot{\theta}_1^2[n] - \dot{\theta}_1[n]\dot{\theta}_2[n]\right) \\ &\quad + a_2^2[n]G_1^2\left(e^{j\dot{\theta}_2[n]}\right)\left(\dot{\theta}_2^2[n] - \dot{\theta}_1[n]\dot{\theta}_2[n]\right). \quad (27)\end{aligned}$$

From (22), we have

$$a_2^2[n]G_1^2\left(e^{j\dot{\theta}_2[n]}\right) = \frac{(\hat{a}_{|y_1|^2}[n])^2}{4a_1^2[n]G_1^2(e^{j\dot{\theta}_1[n]})}. \quad (28)$$

Dropping the time subscript so that the equations are less cumbersome and combining (27) with (28), we have

$$\begin{aligned}|y_2|^2 - \dot{\theta}_1\dot{\theta}_2|y_1|^2 = a_1^2G_1^2\left(e^{j\dot{\theta}_1}\right)\left(\dot{\theta}_1^2 - \dot{\theta}_1\dot{\theta}_2\right) \\ + \frac{(\hat{a}_{|y_1|^2})^2}{4a_1^2G_1^2(e^{j\dot{\theta}_1})}\left(\dot{\theta}_2^2 - \dot{\theta}_1\dot{\theta}_2\right).\end{aligned}$$

Rearranging terms gives

$$\begin{aligned}&\left[a_1^2G_1^2(e^{j\dot{\theta}_1})\right]^2 - \frac{|y_2|^2 - \dot{\theta}_1\dot{\theta}_2|y_1|^2}{\dot{\theta}_1^2 - \dot{\theta}_1\dot{\theta}_2} \\ &\cdot\left[a_1^2G_1^2(e^{j\dot{\theta}_1})\right] - \frac{(\hat{a}_{|y_1|^2})^2\dot{\theta}_2}{4\dot{\theta}_1} = 0.\end{aligned}$$

The above equation has two roots:

$$a_1^2G_1^2(e^{j\dot{\theta}_1}) = \frac{1}{2}\left(\frac{|y_2|^2 - \dot{\theta}_1\dot{\theta}_2|y_1|^2}{\dot{\theta}_1^2 - \dot{\theta}_1\dot{\theta}_2} \pm \sqrt{\left(\frac{|y_2|^2 - \dot{\theta}_1\dot{\theta}_2|y_1|^2}{\dot{\theta}_1^2 - \dot{\theta}_1\dot{\theta}_2}\right)^2 + \frac{(\hat{a}_{|y_1|^2})^2\dot{\theta}_2}{\dot{\theta}_1}}\right). \quad (29)$$

From (25) and (26), $\dot{\theta}_1$ and $\dot{\theta}_2$ are both positive; therefore, it is always true that

$$\frac{(\hat{a}_{|y_1|^2})^2\dot{\theta}_2}{\dot{\theta}_1} > 0$$

[8]If the frequencies cross, the frequency track that corresponded to $\dot{\theta}_1[n]$ will then correspond to $\dot{\theta}_2[n]$ and vice versa. A similar switching will occur for the amplitude estimates.

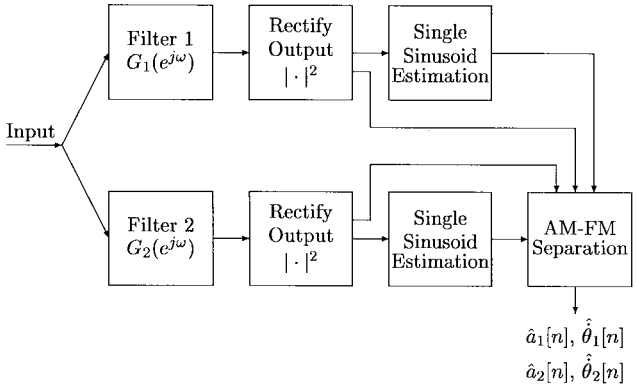


Fig. 7. Block diagram of two-sinusoid AM–FM estimation algorithm.

and thus, it is always true that

$$\sqrt{\left(\frac{|y_2|^2 - \dot{\theta}_1\dot{\theta}_2|y_1|^2}{\dot{\theta}_1^2 - \dot{\theta}_1\dot{\theta}_2}\right)^2 + \frac{(\hat{a}_{|y_1|^2})^2\dot{\theta}_2}{\dot{\theta}_1}} > \left|\frac{|y_2|^2 - \dot{\theta}_1\dot{\theta}_2|y_1|^2}{\dot{\theta}_1^2 - \dot{\theta}_1\dot{\theta}_2}\right|.$$

Therefore, one of the roots of (29) is always negative, and the other is always positive. The correct root is the positive root as shown in (30) at the bottom of the page. We can now obtain $a_2[n]$ from (22) and (30) as

$$a_2 = \frac{\hat{a}_{|y1|^2}}{2a_1G_1\left(e^{j\dot{\theta}_1}\right)G_1\left(e^{j\dot{\theta}_2}\right)}. \quad (31)$$

A block diagram of the two-sinusoid AM–FM estimation algorithm is shown in Fig. 7.

C. Validity of the Approximations

The two terms $\nu_1[n]$ and $\nu_2[n]$ described in (20) and (21) will be referred to as *self generated noise* (SGN). These SGN contributions consist primarily of low-frequency components and thus are effectively eliminated at the input of the single-sinusoid algorithm by the filters $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$.[9] Nevertheless, if the two signals are not stationary, some energy of the SGN passes through the filters and results in error in the estimates. It is difficult to quantify the allowable frequency range of the SGN because there are many factors that influence the sensitivity to the SGN. For example, if the two sinusoids are close in frequency, then $e^{j(\theta_2[n]-\theta_1[n])}$ is at a low frequency near the SGN. In this case, when $|y_i[n]|^2$ passes through $G_i(e^{j\omega})$, the SGN is not reduced relative to the single AM–FM term. Consequently, we use the general

[9]When there is no AM or FM modulation, the noise terms are constant and therefore completely eliminated.

$$a_1 = \frac{\sqrt{\frac{1}{2}\left(\frac{|y_2|^2 - \dot{\theta}_1\dot{\theta}_2|y_1|^2}{\dot{\theta}_1^2 - \dot{\theta}_1\dot{\theta}_2} + \sqrt{\left(\frac{|y_2|^2 - \dot{\theta}_1\dot{\theta}_2|y_1|^2}{\dot{\theta}_1^2 - \dot{\theta}_1\dot{\theta}_2}\right)^2 + \frac{(\hat{a}_{|y_1|^2})^2\dot{\theta}_2}{\dot{\theta}_1}}\right)}}{G_1\left(e^{j\dot{\theta}_1}\right)} \quad (30)$$

guideline that the frequency separation of the two sinusoids should be greater than the frequencies of the SGN. In other words, if we represent $\nu_1[n]$ and $\nu_2[n]$ as $a_1\cos(\phi_1[n])$ and $a_2\cos(\phi_2[n])$, respectively, then the guideline can be stated as $\dot{\theta}_2[n] - \dot{\theta}_1[n] > \max(\dot{\phi}_1[n], \dot{\phi}_2[n])$.

The presence of SGN also requires a modification in the single-component estimation algorithm. Recall that for a real input to the single-component algorithm, the inverse modulation algorithm modulates the signal to a frequency ω_c that is as close to $\pi/2$ as possible. The closeness of ω_c to $\pi/2$ was constrained to avoid modulating the negative frequencies to positive frequencies (see the Appendix). In the two-component case, we are even more constrained because we must prevent the SGN from being modulated to the passband of the filters, which would violate the assumption that most of the energy of the SGN is at low frequencies and therefore significantly reduced by the filters $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$. In a later section, we describe a modification to the two-component algorithm that significantly reduces the SGN. Once the SGN has been reduced, we can use the inverse modulation techniques in the embedded single-component algorithm.

D. Examples of the Two-Component AM-FM Estimation Algorithm

In this section, examples of the performance of the two-component AM–FM estimation algorithm are presented to demonstrate its capabilities and limitations.[10] As a first example, we use the signal

$$x[n] = \cos\left(\frac{2.2\pi n}{5} + 15\sin\left[\frac{2\pi n}{200}\right]\right) + \cos\left(\frac{2.8\pi n}{5} + 15\sin\left[\frac{2\pi n}{200}\right]\right). \quad (32)$$

For each component, the frequency modulation is the same, and the only difference is the carrier frequencies. The results are shown in Fig. 8. Note that the spectrums of the two components overlap significantly. The signal in the next example has two components that cross in frequency

$$x[n] = \cos\left(\frac{\pi n}{10} + \frac{2\pi n^2}{10\,000}\right) + \cos\left(\frac{9\pi n}{10} - \frac{2\pi n^2}{10\,000}\right) \quad (33)$$

and results are shown in Fig. 9. Since the algorithm assigns the lowest frequency estimate to $\hat{\dot{\theta}}_1[n]$, it does not show the frequency cross. In addition, observe that the effect of the frequency crossing is local.

Our purpose in this paper is to present a method of separating AM and FM of a two-tone signal from the amplitude envelope of the output of two transduction filters, motivated by the speculation that frontend auditory filters may use such a scheme. Nevertheless, it is of interest to consider a comparison with an existing scheme. One such approach described in the introduction to this paper is based on tracking spectrogram ridges, i.e., peaks in the short-time Fourier transform magnitude (STFTM) [3]. Figs. 10 and 11 show the frequency

[10]These examples do not use the iterative techniques in the embedded single-component algorithms.

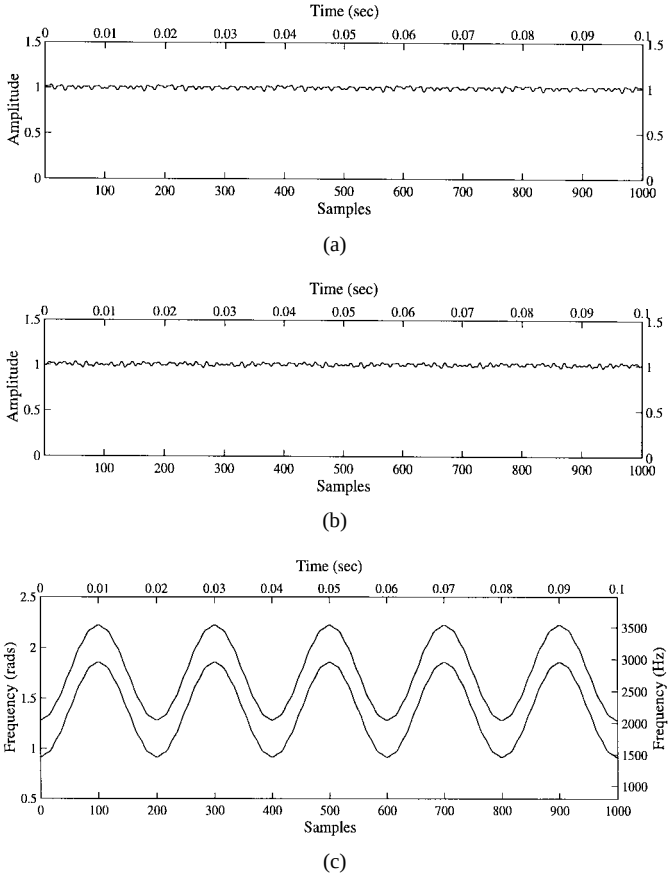


Fig. 8. AM and FM estimates of a two component signal with constant AM and sinusoidal FM. (a) Estimate of $a_1[n]$, mean square error: 2.48×10^{-4}, max deviation: 3.95×10^{-2}. (b) Estimate of $a_2[n]$, mean square error: 2.94×10^{-3}, max deviation: 3.77×10^{-2}. (c) Estimates of $\dot{\theta}_1[n]$ and $\dot{\theta}_2[n]$, mean square error, $\dot{\theta}_1[n]$: 6.29×10^{-5} (rad/sample), mean square error, $\dot{\theta}_2[n]$: 1.00×10^{-4} (rad/sample), max deviation, $\dot{\theta}_1[n]$: 1.73×10^{-2} (rad), max deviation, $\dot{\theta}_2[n]$: 2.10×10^{-2} (rad).

estimation resulting from STFTM ridge tracking for the above two examples of this section. The STFTM is updated every sample, and a 10-ms window length was selected to give a reasonable tradeoff in time and frequency resolution. In the first case, we see that both algorithms have difficulty at the frequency crossover point; the ridge algorithm merges the two frequencies into one, whereas the transduction-based algorithm exhibits instability. For the sinusoidal FM case, the ridge algorithm has difficulty in resolving frequencies due to the smearing of the window Fourier transforms; a longer window in time, thus achieving an improved frequency resolution, degrades the estimate even further due to poor time resolution. A complete comparative study would involve more advanced time–frequency methods, including the Wigner distribution and its variations and other techniques alluded to the introduction. In addition, we can speculate that the inverse modulation approach of this paper might improve the ridge technique [3], as well as other AM–FM estimation methods.

We end this section with an example of an actual two-tone single consisting of the sum of two AM–FM tones from a guitar. Fig. 12 shows the spectrogram of the signal and the frequency estimates derived from our two-tone transduction-based algorithm.

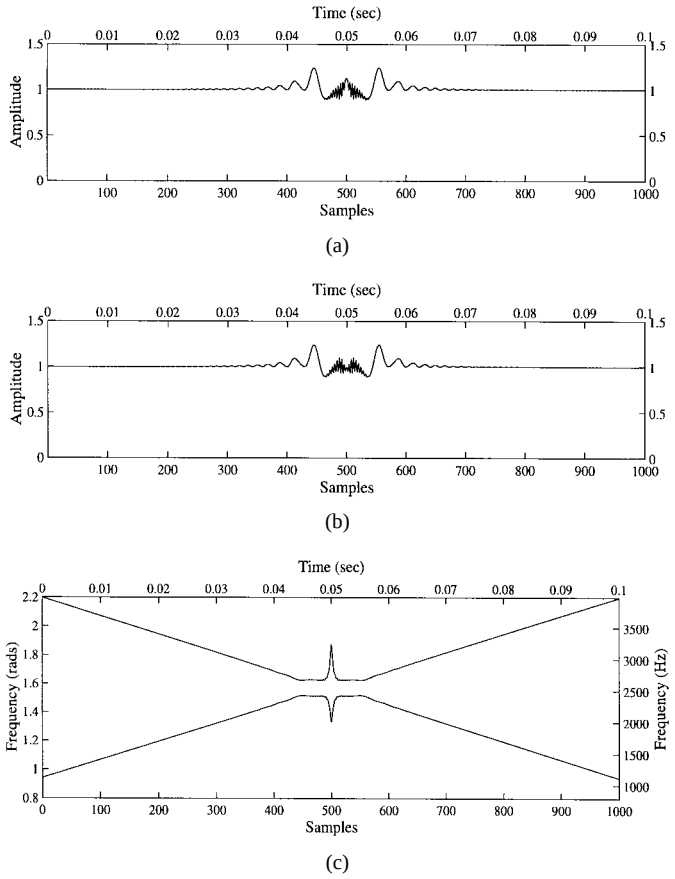


Fig. 9. Example in which the frequencies of the two components cross. (a) Amplitude estimate of first component. (b) Amplitude estimate of second component. (c) Frequency estimates.

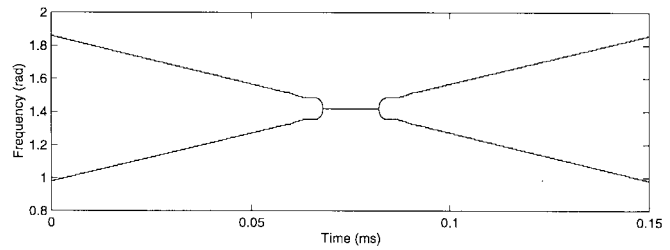


Fig. 10. Ridge tracking on signal with two crossing frequencies.

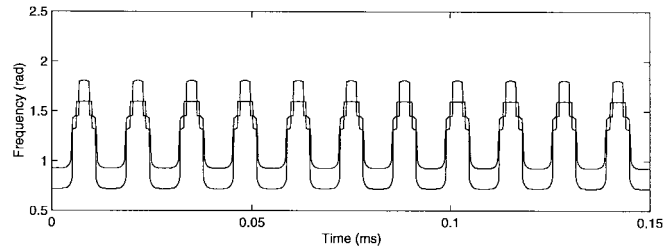


Fig. 11. Ridge tracking on signal with two frequencies with sinusoidal FM.

E. Improvements on the Two-Component Algorithm

We now make two improvements to the two-component algorithm. First, the estimates of $\hat{a}_1[n]$, $\hat{a}_2[n]$, $\hat{\dot{\theta}}_1[n]$, and $\hat{\dot{\theta}}_2[n]$ can be used to reduce the SGN. Second, the AM of each component can be inverted with time-varying filters to reduce the transduction error. Using both methods together, we reduce the overall mean-squared error by a factor of 10. This section briefly describes each technique.

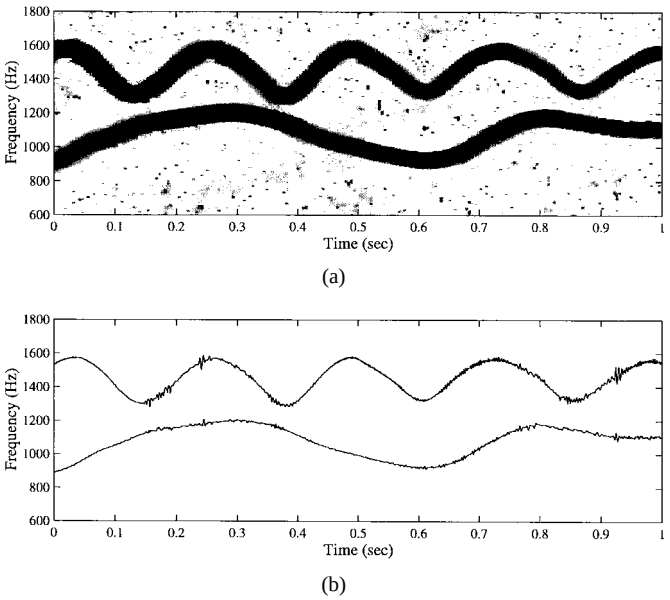


Fig. 12. FM estimates of AM–FM signal from a guitar. (a) Spectrogram of signal. (b) FM estimates of two-component algorithm.

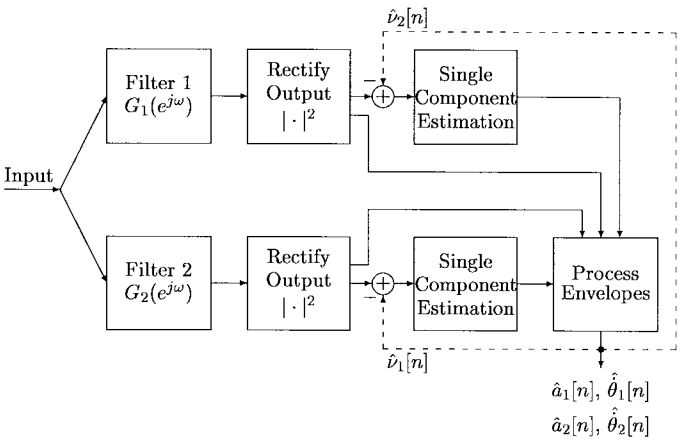


Fig. 13. Block diagram of two-component AM–FM estimation algorithm with SGN cancellation feedback.

1) Eliminating the SGN: The SGN can be estimated from the AM and FM estimates of a first pass with the two-component algorithm and then reduced on a second pass. The estimate of the SGN is calculated from (20) and (21) and is then subtracted from the output of filters $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$. Fig. 13 shows the block diagram with the SGN noise cancellation. With significant reduction of the SGN, we can use the single-sinusoid inverse modulation algorithm as it was described in Section IV.

To illustrate the effects of SGN cancellation, consider the signal

$$x[n] = \left[1 + 0.5\cos\left(\frac{2\pi n}{50}\right)\right]\cos\left(\frac{2\pi n}{5}\right) + \cos\left(\frac{3\pi n}{5}\right). \tag{34}$$

The effects of SGN cancellation for this signal are shown in Fig. 14. Further illustration of the benefits of SGN are given in [25].

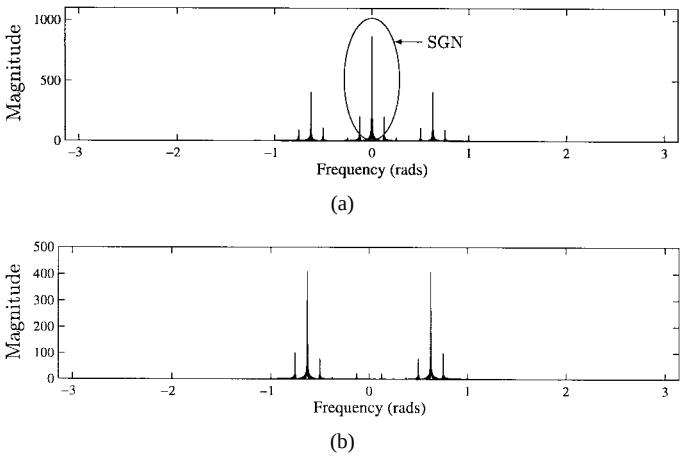


Fig. 14. (a) Magnitude of Fourier transform of $|g_1[n] * x[n]|^2$, which is the input to the single-component algorithm on the first pass (b) with SGN reduction on the second pass. The sequence $x[n]$ is given by (34)

2) Inverting the Amplitude Modulation: Inverting the modulation when the signal consists of two AM–FM components is not as straightforward as in the single-component case. The difficulty arises because we cannot operate on each signal component independently. This makes inverting the FM, in particular, very difficult. In order to keep the frequency separation between the two signals constant, we would need to adjust the sampling rate as a function of time to compress and stretch the spectrum of the signal. We do not address FM inversion for the two-component case in this paper.

To invert the AM, we use a pair of filters that have a linear frequency response for $\omega \in [0, \pi]$ and are unrestricted[11] for $\omega \in (-\pi, 0)$. At each time sample, using estimates from an initial pass through the two-component estimation algorithm, we weight the output of the two linear filters in such a way that the amplitudes of both components are scaled to a value of one. For example, suppose that at time n_0, $a_1[n_0] = 2$, $a_2[n_0] = 3$, $\dot{\theta}_1[n_0] = \pi/4$, and $\dot{\theta}_2[n_0] = 3\pi/4$. Then, if we pass this signal through a filter with a frequency response $H(e^{j\omega})$ such that $H(e^{j(\pi/4)}) = 1/2$ and $H(e^{j(3\pi/4)}) = 1/3$, the output of this filter at n_0 consists of the sum of two sinusoids: each sinusoid with an amplitude equal to one. With this procedure performed at each time sample, the amplitude modulation is approximately eliminated (assuming the modulation estimates are close to the true modulation). Implementation of the method is fully described in [25].

We again use the signal described in (34) to demonstrate the effect of AM inversion, together with SGN cancellation. The results are shown in Fig. 15. Adding the AM inversion has improved the estimate by reducing spectral sidelobes due to AM. The mean-squared-error for the AM and FM estimates is reduced overall by a factor of 10 [25].

V. Conclusion

In this paper, we first improved on a single-component AM–FM estimation algorithm that is based on the output envelope of two overlapping transduction filters. The improvement involved "inverting" the signal modulation and then applying the algorithm to the demodulated signal. We developed a frequency-domain approach to analyze and select filters with respect to transduction error. We also derived bounds on the frequency estimate for an amplitude-bounded noise disturbance. We then proposed a method to estimate the AM and FM functions when the input signal is the sum of two AM–FM sinusoids and showed that this problem can be essentially reduced to two single-component AM–FM estimation problems. In addition, we improved this algorithm by both canceling the self generated noise (SGN) and inverting the AM. We gave several examples demonstrating the algorithm under a wide range of AM–FM signals. Additional examples and evaluations can be found in [25].

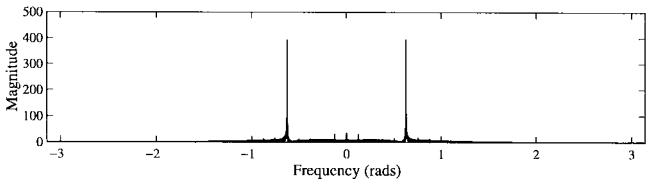


Fig. 15. Magnitude of Fourier transform of $|g_1[n] * x[n]|^2$ after SGN cancellation and AM inversion. The sequence $x[n]$ is given by (34).

There are several aspects of the algorithms presented that can be improved upon or that require further investigation. One of the key factors in choosing the particular relation between transduction filters was analytic simplicity. Therefore, although the filters are optimal according to our goal attainment approach, tradeoffs still exist in choosing this relation between overlapping transduction filters involving spectral shape and impulse response length as well as simplicity. In Section III, we showed experimentally that inverting the modulation with estimates of the AM and FM significantly reduced the estimation error. A current area of investigation is proving that this technique converges to the exact AM and FM functions. In the algorithms presented in this paper, we have used only two filters. A possible approach to reduce the error is to use many filters and exploit the redundancy to obtain improved AM–FM estimates. For example, we could average the estimates obtained from different filter sets. Such approaches, as was the theme of this paper, are motivated by the hypothesis that the auditory system, with multiple overlapping filters at various processing stages, may use similar mechanisms in robustly "perceiving" AM and FM in a signal [20]. Indeed, it has recently been shown that the transduction approach can be realized as a bank of cochlear bandpass filters followed by envelope detectors and shunting neural networks and that the resulting dynamical system is capable of robust AM–FM estimation in noisy environments and over a broad range of filter bandwidths and locations [2]. Shunting neural networks appear throughout the nervous system and are characterized by a form of ratio processing.

Finally, a problem we are currently considering is AM–FM estimation of multiple sinusoids. In this case, one approach is to use a filterbank of more than two filters, as illustrated in Fig. 16. Initial work in this area has shown that the shape of the filters must be nonlinear for any additional information to be gained by more then two filters, and that the approach

[11] The filters are unrestricted for $\omega \in (-\pi, 0)$ because when we compute the analytic signal, the negative frequencies of $x[n]$ are eliminated.

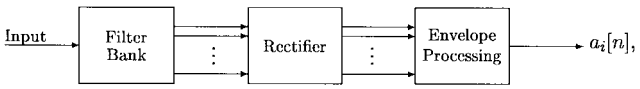


Fig. 16. General filterbank approach to AM–FM estimation.

used for the two-sinusoid case, i.e., reducing the problem to a lower order problem, is not applicable [25]. Rather, the multicomponent AM–FM problem has been formulated as a nonlinear estimation problem, and standard numerical approaches of the Newton iterative type were used to solve the set of nonlinear equations that result from taking the magnitude of the filter bank outputs. Examples for the single- and two-sinusoid case have been demonstrated using a five-filter filterbank, and an approach for more than two components was outlined. For the later case, solution uniqueness has yet to be established.

Generalization of the two-tone AM–FM estimation algorithm to multicomponent signals is currently being investigated in the particular context of speech applications where AM–FM representations have an extensive history. Perhaps the earliest application of AM–FM representations in speech processing was by Flanagan and Golden [7] in the phase vocoder that allowed bit-rate reduction by downsampling the AM and FM of each filterbank output, rather than the output signal itself, where the AM and FM had a smaller bandwidth than the output signal. This same concept has been applied more recently in a speech vocoder based on AM–FM resonances where the resonant AM and FM is downsampled rather than the bandpass filter output AM–FM [16]. Yet another application is in speaker and speech recognition, where a parameterized version of the AM and FM are used as features in recognition systems [10], [19]. AM–FM estimation of multiple tones, as well as of functions of the multiple components, will be especially important in these applications where sinusoidal components are often closely spaced.

APPENDIX
EFFECT OF NEGATIVE FREQUENCIES IN INVERSE MODULATION

In the inverse modulation algorithm of Section III-D, for a real signal $x[n]$, we restrict the function used to invert the FM so that negative frequencies are not modulated to positive frequencies. We denote $\dot{\theta}_+[n]$ as the positive FM function of $x[n]$ and $\dot{\theta}_-[n]$ as the negative FM function. For real $x[n]$

$$x[n] = \tfrac{1}{2}\, a[n]\left(e^{j\int_{-\infty}^{n} \dot{\theta}_+[p]\,dp} + e^{j\int_{-\infty}^{n} \dot{\theta}_-[p]\,dp}\right)$$

where $\dot{\theta}_+[n] = -\dot{\theta}_-[n]$. Demodulating with $\omega_c n - \hat{\theta}_+[n]$ results in

$$\begin{aligned}\tilde{x}[n] &= \tfrac{1}{2}\, a[n]\left(e^{j\int_{-\infty}^{n} \dot{\theta}_+[p]\,dp} + e^{j\int_{-\infty}^{n} \dot{\theta}_-[p]\,dp}\right) \\ &\quad \cdot e^{j(\omega_c n - \int_{-\infty}^{n} \hat{\theta}_+[p]\,dp)} \\ &\approx \tfrac{1}{2}\, a[n]\left(e^{j\omega_c n} + e^{j(\omega_c n + \int_{-\infty}^{n} (\dot{\theta}_-[p] - \hat{\dot{\theta}}_+[p])\,dp)}\right).\end{aligned}$$

To prevent the negative frequencies from being modulated into the positive frequency range, we must force

$$-\pi \le \dot{\theta}_-[n] - \hat{\dot{\theta}}_+[n] + \omega_c \le 0 \qquad \text{for all } n \tag{35}$$

where $\omega_c \in [0, \pi]$. A necessary condition for this inequality to be satisfied is that

$$|\dot{\theta}_-[n] - \hat{\dot{\theta}}_+[n]| < \pi \qquad \text{for all } n. \tag{36}$$

To satisfy this inequality, there are two options. We can scale $\hat{\dot{\theta}}_+[n]$ by some factor $0 < \alpha < 1$, or we can divide the signal into shorter sections in which the frequency variation of $x[n]$ is confined to a smaller region. With the first option, the frequency modulation would never be completely inverted. The second option assumes that the frequency modulation of $x[n]$ does not vary over a range of $\pi/2$ rad/sample within the duration of the impulse response of the filters $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$. This does not impose a significant constraint. For example, if the sampling frequency is 10 000 Hz, a signal that violated this constraint would have to sweep across 2500 Hz in approximately 1 ms. If a signal does violate this assumption, we can use both options, i.e., use a short block of the signal and scale the FM estimate by $0 < \alpha < 1$ to ensure that (36) is not violated. Therefore, we always use option two and, if necessary, combine it with option one.

Once we have satisfied (36), we can now choose ω_c. Because we have optimized the design of our transduction filter pair about the frequency $\omega = \pi/2$, we select ω_c as close to $\pi/2$ as possible while still satisfying (35) [25]. Assuming that (36) is satisfied, ω_c is determined by

$$\omega_c = \begin{cases} \min\limits_n(2\hat{\dot{\theta}}_+[n]), & \text{if } \min\limits_n(\hat{\dot{\theta}}_+[n]) < \dfrac{\pi}{4} \\ \dfrac{\pi}{2}, & \text{if } \min\limits_n(\hat{\dot{\theta}}_+[n]) > \dfrac{\pi}{4} \\ & \quad \text{and } \max\limits_n(\hat{\dot{\theta}}_+[n]) < \dfrac{3\pi}{4} \\ \pi - \max\limits_n(2\hat{\dot{\theta}}_+[n]), & \text{if } \max\limits_n(\hat{\dot{\theta}}_+[n]) \ge \dfrac{3\pi}{4}. \end{cases} \tag{37}$$

In deriving the constraint in (37), it was assumed that we knew the actual FM function. However, we have only an estimate of the FM function, and because this estimate is generally not exact, using it to demodulate the FM might violate the constraint in (36). In addition, the spectrum of an FM signal has spectral components outside the range of the instantaneous frequency. This implies that although $\dot{\theta}_-[n]$ is confined to $(-\pi, 0)$, $e^{j\dot{\theta}_-[n]}$ might have spectral components in the positive frequency range. Therefore, we must leave some room for error when we demodulate, thus giving the constraint

$$-\pi + \varepsilon < \dot{\theta}_-[n] - \hat{\dot{\theta}}_+[n] + \omega_c < 0 - \varepsilon \quad \text{for all } n \tag{38}$$

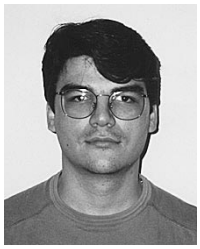
where $\varepsilon > 0$. Determining the best ε is a complex issue that involves finding a tight bound for the estimation error as well as determining the spectral range of the negative frequencies. We chose $\varepsilon = 0.06$ rad/s based on experimental results.

ACKNOWLEDGMENT

The authors wish to thank Dr. J. F. Kaiser of Duke University and the four anonymous IEEE reviewers for detailed and insightful comments on the manuscript.

REFERENCES

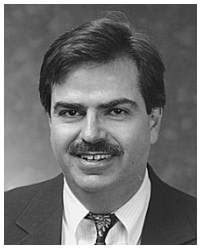
[1] E. H. Armstrong, "A method of reducing disturbances in radio signaling by a system of frequency modulation," *Proc. Inst. Radio Eng.*, vol. 24, pp. 689–740, May 1936.
[2] R. A. Baxter and T. F. Quatieri, "AM-FM separation using shunting neural networks," in *Proc. Int. Symp. Time-Freq. Time-Scale Anal.*, Pittsburgh, PA, 1998.
[3] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal," *Proc. IEEE*, vol. 80, pp. 519–568, Apr. 1992.
[4] A. C. Bovik, J. P. Havlicek, and M. D. Desai, "Theorems for discrete filtered modulated signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Minneapolis, MN, Apr. 1993, vol. 3, pp. 153–156.
[5] L. Cohen, *Time-Frequency Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1995.
[6] J. P. Costas, "Residual signal analysis," *Proc. IEEE*, vol. 68, pp. 1351–1352, Oct. 1980.
[7] J. L. Flanagan and R. M. Golden, "Phase vocoder," *Bell Syst. Tech. J.*, vol. 45, pp. 1493–1509, 1966.
[8] F. W. Gembicki, "Vector optimization for control with performance and parameter sensitivity indices," Ph.D. dissertation, Case Western Res. Univ., Cleveland, OH, 1974.
[9] D. M. Green, *An Introduction to Hearing.* New York: Wiley, 1976.
[10] C. R. Jankowski, "Fine structure features for speaker identification," Ph.D. disertation, Mass. Inst. Technol., Cambridge, June 1996.
[11] R. Kumaresan, A. G. Sadasiv, C. S. Ramalingam, and J. F. Kaiser, "Instantaneous nonlinear operators for tracking multicomponent signal parameters," in *Proc. IEEE 6th SP Workshop Stat. Signal Array Process.*, Victoria, B.C., Canada, 1992, pp. 404–407.
[12] 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, 1993.
[13] 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.
[14] R. M^{c} Eachern, "How the ear really works," in *Proc. IEEE-SP Int. Symp. Time-Freq. Time-Scale Anal.*, Victoria, B.C., Canada, Oct. 1992, pp. 437–440.
[15] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1989.
[16] A. Potiamanos and P. Maragos, "Applications of speech processing using AM-FM modulation model and energy operators," in *Proc. 7th Euro. Signal Process. Conf. EUSIPCO*, Edinburgh, U.K., 1994.
[17] T. F. Quatieri, T. E. Hanna, and G. C. O'Leary, "AM-FM separation using auditory-motivated filters," *IEEE Trans. Speech Audio Process.*, vol. 5, pp. 465–480, Sept. 1997; in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Atlanta, GA, May 1996.
[18] C. S. Ramalingam, "Analysis of nonstationary, multi-component signals with applications to speech," Ph.D. dissertation, Univ. Rhode Island, Kingston, Jan. 1995.
[19] J. Rouat and M. Garcia, "A prototype speech recognizer based on associative learning and nonlinear speech analysis," in *Proc. Workshop Comput. Scene Anal.*, Quebec City, P.Q., Canada, 1995.
[20] K. Saberi and E. R. Hafter, "A common neural code for frequency- and amplitude-modulated sounds," *Nature*, vol. 374, no. 6, pp. 537–539, Apr. 1995.
[21] B. Santhanam and P. Maragos, "Demodulation of discrete multicomponent AM-FM signals using periodic algebraic separation and energy demodulation," in *Proc. Int. Conf. Acoust., Speech Signal Process.*, 1997, pp. 2409–2412.
[22] M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques.* New York: McGraw-Hill, 1966.
[23] D. Slepian and H. O. Pollak, "Prolate spheroidal wave function, Fourier analysis and uncertainty," *Bell Syst. Tech. J.*, pp. 44–84, Jan. 1961.
[24] G. B. Thomas, *Calculus and Analytic Geometry*, 4th ed. Reading, MA: Addison-Wesley, 1968.
[25] W. P. Torres, "Estimation of signal modulation based on FM-to-AM transduction," Master's thesis, Mass. Inst. Technol., Dept. Elect. Eng. Comput. Sci., Cambridge, MA, May 1997.
[26] D. Vakman, "On the definition of concepts of amplitude, phase and instantaneous frequency of a signal," *Radio Eng. Electron. Phys.*, vol. 17, no. 5, pp. 754–759, 1972.
[27] W. A. Van Bergeijk, J. R. Pierce, and E. E. David, *Waves and the Ear*. Garden City, NY: Anchor, 1958.



Wade P. Torres (S'98) received the B.S. degree in electrical engineering and the B.S. degree in mathematics from Southern Illinois University, Carbondale, both in 1995. In 1997, he received the M.S. degree in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge. He is currently pursuing the Ph.D. degree in electrical engineering at MIT.

From 1995 to 1997, he was with MIT Lincoln Laboratory, Lexington, where he worked on time–frequency analysis of nonstationary signals. Since 1997, he has been with the Digital Signal Processing Group in the Research Laboratory of Electronics at MIT. His current research interests include chaos theory, time–frequency analysis, and digital signal processing in the music industry.

Mr. Torres is a member of Sigma Xi and Tau Beta Pi.



Thomas F. Quatieri (S'73–M'79–SM'87–F'98) received the B.S. degree (summa cum laude) from Tufts University, Medford, MA, in 1973 and the S.M., E.E., and Sc.D. degrees from the Massachusetts Institute of Technology (MIT), Cambridge, in 1975, 1977, and 1979, respectively.

He is currently a Senior Member of the Technical Staff at MIT Lincoln Laboratory, Lexington, where he is involved in digital signal processing for speech, audio, and underwater sound applications and in nonlinear signal processing. He has contributed many publications to journals and conference proceedings, written several patents, and coauthored chapters in numerous edited books. He holds the position of Lecturer at MIT, where he has developed the graduate course digital speech processing.

Dr. Quatieri is the recipient of the 1982 Paper Award of the IEEE Acoustics, Speech, and Signal Processing Society for the paper "Implementation of 2-D Digital Filters by Iterative Methods." In 1990, he received the IEEE Signal Processing Society's Senior Award for the paper "Speech Analysis/Synthesis Based on a Sinusoidal Representation," and in 1994, he won this same award for the paper "Energy Separation in Signal Modulations with Application to Speech Analysis," which was also selected for the 1995 IEEE W.R.G. Baker Prize Award. He was a member of the IEEE Digital Signal Processing Technical Committee, and from 1983 to 1992, he served on the steering committee for the biannual Digital Signal Processing Workshop. He was Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING in the area of nonlinear systems. He is also a member of Tau Beta Pi, Eta Kappa Nu, Sigma Xi, and the Acoustical Society of America.