EFFECT OF MODEL MISMATCH ON POLYNOMIAL ENVELOPE AND PHASE MODELING OF NONSTATIONARY SINUSOIDS

C.S. Ramalingam R. Kumaresan

Department of Electrical Engineering
University of Rhode Island
Kingston, RI 02881.
USA

ABSTRACT

In this paper we examine the effect of model mismatch when modeling a signal consisting of multiple non-stationary sinusoids. The envelopes and frequencies of the components were modeled as polynomials of low order over short intervals of time and the coefficients estimated using the least-squares error criterion. If the block length and model orders are not properly chosen, unacceptable errors occurred in the estimated frequency tracks. The errors tended to increase as the number of components increased. Using the simpler constant envelope, constant frequency sinewave model for short, heavily overlapping blocks and smoothing the resulting frequency tracks gave surprisingly good results when analyzing a long duration multicomponent signal. We conclude that the more complex polynomial model may not always yield the expected increase in accuracy in signal modeling.

1. INTRODUCTION

A number of naturally occurring and man-made signals have components whose envelopes and frequencies vary slowly with time. A typical example is voiced-speech. In this paper we concern ourselves with signals of the form

$$s[n] = \sum_{k=1}^{M} a_k[n] \exp j\psi_k[n] = \sum_{k=1}^{M} \left(\sum_{l=0}^{P_k} a_{kl}\, \alpha_{kl}[n] \right) \exp \left(j \sum_{r=0}^{Q_k} b_{kr}\, \beta_{kr}[n] \right), \quad (1)$$

where the non-negative envelope $a_k[n]$ has been expressed as a linear combination of the polynomials $\alpha_{kl}[n]$; similarly, $\psi_k[n]$ is a linear combination of $\beta_{kr}[n]$. The *instantaneous frequency* (IF) of the k-th component is $f_k[n] = (1/2\pi)(\psi_k[n] - \psi_k[n-1])$, and M is the total number of components present. We assume that $a_k[n]$ and $f_k[n]$ are only slowly varying so that they can be modeled quite accurately by polynomials of low order over short intervals of time. We further assume that $f_k[n] \neq f_l[n]$ whenever $k \neq l$. Given a signal $x[n]$ with components that can be modeled as in Eq. (1), our aim is to estimate the coefficients a_{kl} and b_{kr} by minimizing the mean squared error between the signal and the model. Observe that Eq. (1) reduces to the conventional sinewave model when the envelopes are constant and the phases are linear.

The classical FM signal can be thought of as a monocomponent polynomial phase signal and FM receivers can be viewed as signal analyzers. Cahn [1] showed that the performance of the phase-locked loop (PLL) can be improved if one allows for a processing delay. Tufts and Francis [2] showed that searching locally for peaks in the discrete Fourier transform of short, heavily overlapping blocks is a MAP estimator for the local frequency. For single component signals, these procedures yield good results. Recently, analysis of multicomponent signals has attracted a lot of attention. Kumaresan and Verma [3] and Liang and Arun [4] used rank reduction methods to get the parameters of chirp signals. Djurić and Kay [5] modeled the phase as a polynomial and estimated its parameters using a least-squares procedure. Peleg and his collaborators [6, and the references therein] introduced the Discrete Polynomial Transform to compute the parameters of single and multicomponent polynomial phase signals. In [7] Boashash has reviewed a number of instantaneous frequency estimation methods, including polynomial phase modeling. The most general polynomial envelope and phase model given in Eq. (1) was proposed, independently of the work in [8], by Friedlander and Francos [9]. They give details of the maximum-likelihood estimator of the polynomial coefficients and also derive relevant Cramér–Rao lower bounds.

Our main concern is not so much the study of the least-squares estimator's performance as the examination of the effects of model mismatch on the estimated frequency tracks. The motivation to examine such effects was provided by the difficulties we faced when modeling natural voiced speech. Except in the case of certain computer generated signals, model mismatch cannot be avoided. A simple example demonstrates that choosing inappropriate block sizes and model orders can result in unacceptable errors in the frequency tracks of the components. In contrast, smoothing the frequency tracks obtained from fitting the simpler constant envelope, constant frequency sinewave model gave surprisingly good results when analyzing a a long duration signal. But the heavy overlap can lead to excessive computation. We have proposed alternative methods [10–12] that

This work was supported by the Air Force Contract AFOSR F49620-93-I-0026. The first author was also supported by a URI Graduate Fellowship award.

0-7803-2431-5/95 $4.00 © 1995 IEEE

use the smoothed IFs of the components; they incorporate Costas's principle of Residual Signal Analysis (RSA) [13]. Riedel [14] uses similar ideas, but computes the IF through kernel estimators.

2. ESTIMATING THE MODEL PARAMETERS

In Eq. (1) the only unknowns are the envelope and phase polynomial coefficients $\{a_{kl}\}$ and $\{b_{kr}\}$; we have to choose them so as to minimize $E = \sum_{n=0}^{N-1} |s[n] - x[n]|^2$. The coefficients should be constrained to be real-valued.

The following definitions lead to succinct expressions: The Hadamard product '$*$' of two $m \times n$ matrices A and B is defined as the element-by-element multiplication of their corresponding entries [15]. Let

$$\begin{aligned}
\tilde{\mathbf{s}} &= (s[0]\ s[1]\ \cdots\ s[N-1])^T \\
\tilde{\mathbf{x}} &= (x[0]\ x[1]\ \cdots\ x[N-1])^T \\
\boldsymbol{\alpha}_{kl} &= (\alpha_{kl}[0]\ \alpha_{kl}[1]\ \cdots\ \alpha_{kl}[N-1])^T \\
\boldsymbol{\beta}_{kl} &= (\beta_{kl}[0]\ \beta_{kl}[1]\ \cdots\ \beta_{kl}[N-1])^T \\
\boldsymbol{\psi}_k &= \sum_{r=0}^{Q_k} b_{kr}\, \boldsymbol{\beta}_{kr} \\
\mathbf{W}_k &= \left[\boldsymbol{\alpha}_{k0} * e^{j\boldsymbol{\psi}_k}\ \boldsymbol{\alpha}_{k1} * e^{j\boldsymbol{\psi}_k}\ \cdots\ \boldsymbol{\alpha}_{kP_k} * e^{j\boldsymbol{\psi}_k}\right] \\
\tilde{\mathbf{W}} &= (\mathbf{W}_1\ \mathbf{W}_2\ \cdots\ \mathbf{W}_M) \\
\mathbf{a}_k &= (a_{k0}\ a_{k1}\ \cdots\ a_{kP_k})^T \\
\mathbf{a} &= \left(\mathbf{a}_1^T\ \mathbf{a}_2^T\ \cdots\ \mathbf{a}_M^T\right)^T \\
\mathbf{b}_k &= (b_{k0}\ b_{k1}\ \cdots\ b_{kQ_k})^T \\
\mathbf{b} &= \left(\mathbf{b}_1^T\ \mathbf{b}_2^T\ \cdots\ \mathbf{b}_M^T\right)^T \quad . \qquad (2)
\end{aligned}$$

Using the above definitions, the error E can be written as $E = \left\|\tilde{\mathbf{W}}\mathbf{a} - \tilde{\mathbf{x}}\right\|^2$. Because the coefficients have to be real-valued, E is written in a slightly modified form before minimization. Let

$$\begin{aligned}
\mathbf{W} &= \begin{bmatrix} \mathrm{Re}\{\tilde{\mathbf{W}}\} \\ \mathrm{Im}\{\tilde{\mathbf{W}}\} \end{bmatrix} , \\
\mathbf{x} &= \begin{bmatrix} \mathrm{Re}\{\tilde{\mathbf{x}}\} \\ \mathrm{Im}\{\tilde{\mathbf{x}}\} \end{bmatrix} , \\
\mathbf{s} &= \begin{bmatrix} \mathrm{Re}\{\tilde{\mathbf{s}}\} \\ \mathrm{Im}\{\tilde{\mathbf{s}}\} \end{bmatrix} ,
\end{aligned}$$

where $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ denote the real and imaginary parts. If the following error function is minimized, then we can be assured that the coefficients will be real-valued:

$$E = \|\mathbf{W}\mathbf{a} - \mathbf{x}\|_2^2 \quad . \qquad (3)$$

The problem of finding $\mathbf{a}$ and $\mathbf{b}$ that minimize E is well known [16]; the expression for E can be written such that it depends explicitly on $\mathbf{b}$ only. Unfortunately, this dependence is nonlinear and finding the global minimum of E is, in general, very difficult. The usual procedure is to conduct a coarse search in the parameter space to get a good initial guess and then use a gradient descent algorithm to reach the nearest minimum, which will be the global minimum only if the initial guess is good. The dimension of the error surface E is $\sum_{k=1}^{M} Q_k$; the difficulty of finding its global minimum and the computational complexity grow rapidly with increasing dimension.

3. SIMULATION RESULTS

We investigated model mismatch by considering examples containing up to three components. The envelope and frequency of each component are polynomials of degree two or less. We deliberately introduced model mismatch and examined whether or not the errors are within acceptable limits. More detailed results can be found in [8].

In Example 3.1, the envelope and *frequency* tracks were such that they were a linear combination of the same set of polynomials. In particular, the Gram polynomials [17] $p_k[n]$ were chosen because of their good numerical properties. The space of the phase functions $\boldsymbol{\psi}_k$ is spanned by $p_0[n]$ and the cumulative sums of $p_0[n]$, $p_1[n]$, ... In the figures that follow, the frequency tracks—rather than the phase tracks—are plotted. The sampling frequency in these examples was 16 kHz.

Example 3.1 Three Components: Quadratic Frequency, Linear Envelope Model. A signal containing three components was generated. The envelopes of all three components were quadratic and identical. The frequency tracks of the first two components were quadratic while the third was linear. They were nominally harmonic and spaced approximately 250 Hz apart. The signal was 289 samples long ($\approx$ 18 ms). Over the first 145 samples, each component was modeled by a quadratic frequency and linear envelope. The gradient descent algorithm was initialized at the true parameter values. The estimated frequency tracks are the solid lines shown in Figures 1(a)—1(c). The dashed lines represent the true frequency tracks. The maximum error in the frequency track of the second component is nearly 50 Hz. The estimated envelope tracks are shown in Figure 1(d); the thick line represents the true value, while the others are the estimates.

Example 3.2 Constant Envelope, Constant Frequency Modeling of Nonstationary Sinusoids. To see if the constant envelope, constant frequency model could be used successfully to analyze nonstationary components, we generated a 235 ms long signal containing four nominally harmonic components with varying frequency; the envelopes were also time-varying. Blocks of 145 samples were modeled by four sinewaves of constant envelope and frequency; the estimated parameters were assigned to the middle sample; the overlap between the blocks was 144 samples. The unsmoothed frequency track of the third component is shown in Figure 2(a); the result of smoothing it by an FIR lowpass filter is shown in Figure 2(b). It agrees remarkably well with the true frequency track.

4. DISCUSSION

When the signal contained only one component, model mismatch in either envelope or phase (or both) resulted in tracks that matched closely with the least-squares fit to the true tracks. However, when there are two or more components and incorrect model orders are specified, there can be significant errors in the estimated frequency tracks, as illustrated in Example 3.1. There, choosing wrong orders for the envelopes of the two components resulted in a maximum error of nearly 50 Hz in the estimated frequency tracks. When the block size was increased to 289, the errors due to model mismatch decreased; the maximum error in the frequency tracks obtained by initializing the gradient descent algorithm at the true parameter values was about 6 Hz. These errors can be expected to increase as the number of components increases. In general, the difficulty of finding the global minimum quickly grows with dimensionality of the error surface.

When the frequency spacing between the components was increased, the errors in the estimated tracks caused by wrong model orders, not surprisingly, tended to decrease. Another property of polynomial envelope and phase modeling that seems to be true based on simulation experiments is that choosing wrong orders for the envelopes has a more severe effect on the estimated frequency tracks than vice-versa. We modeled the signal in Example 3.1 by components that had quadratic envelopes—the correct order—while providing only linear frequencies for the first two components (their true tracks being quadratic); the third component's true and assigned orders matched; the block size was 145 samples. In this case the estimated frequency tracks almost coincided with tracks obtained as least-squares fit to the true ones, despite the three components being closely spaced. This is in sharp contrast with the 50 Hz or so maximum error in the second component's estimated frequency track when there was a mismatch in the model order for the envelope.

Modeling short blocks of data as constant envelope, constant frequency sinewaves and estimating their parameters does surprisingly well. The block size is not that critical a parameter in this approach. The idea is to overlap the blocks heavily and assign the estimates to the middle sample, similar to that proposed in [2] for one component. However, when there are many components, the estimated frequency tracks are quite noisy; *smoothing* them appropriately gives tracks that are quite close to the true ones. In contrast, unacceptable errors might still result due to model mismatch when using the more sophisticated polynomial model. One has to choose the proper block size, an accurate model, and have good initial guesses to reach the global minimum for each block. Our simulations indicate that polynomial modeling need not always give the increased accuracy that one expects from the more complex model.

5. CONCLUSIONS

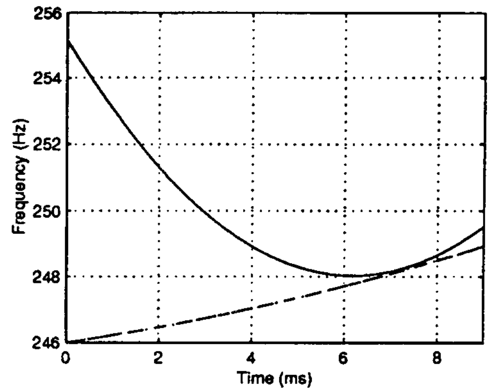
The main point that is made by the examples in this paper is that the more sophisticated polynomial model does not automatically give the increased accuracy expected of it. It is rather sensitive to model mismatch; proper block sizes have to be chosen and good initial guesses are required to reach the global minimum for each block. Otherwise, significant errors can result. If there is noise, the situation can be expected to get only worse. In contrast, the simpler constant envelope, constant frequency model with appropriate post-processing gives good results.

6. REFERENCES

[1] C. H. Cahn, *IEEE Trans. Info. Theory*, vol. 20, pp. 50–58, Jan. 1974.

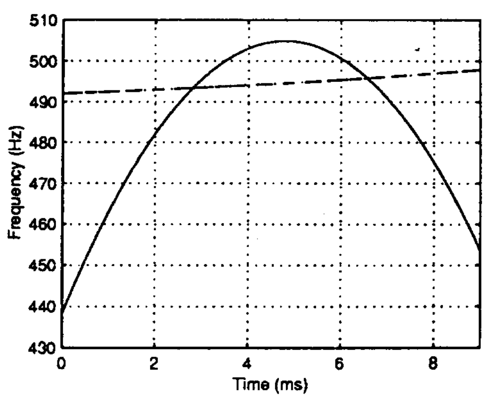
[2] D. W. Tufts and J. T. Francis, *IEEE Trans. Info. Theory*, vol. 23, pp. 742–751, Nov. 1977.

[3] R. Kumaresan and S. Verma, in *Proc. of the 21st Asilomar Conf. Sig., Syst., and Comp.*, (Pacific Grove, CA), pp. 555–558, Nov. 1987.



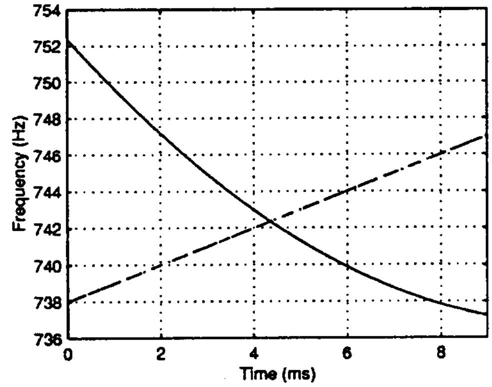
(a)

Figure 1: The signal contained three components with identical envelopes that were quadratic; the first two frequency tracks were also quadratic, while the third was linear. Each model component had linear envelope and quadratic frequency. The gradient descent procedure was initialized at the true values. (a) Estimated frequency track of the first component (solid line); also shown is the true track (dashed-dotted line).

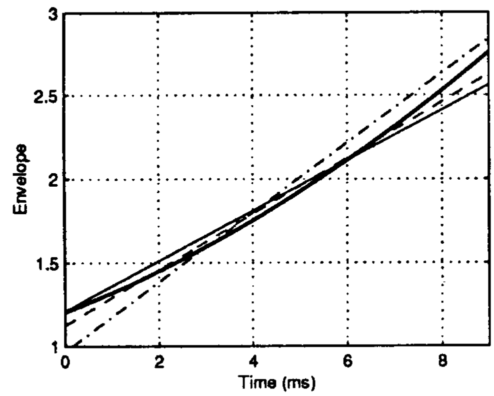


(b)

Figure 1: (b) True and estimated frequency tracks of the second component.



(c)



(d)

Figure 1: (c) True and estimated frequency tracks of the third component. (d) Estimated envelopes of all three components. The thick line represents the true envelope.

[4] R. M. Liang and K. S. Arun, in *Proc. IEEE ICASSP-92*, (San Francisco, CA), pp. V:273–276, Mar. 1992.

[5] P. Djurić and S. M. Kay, *IEEE Trans. ASSP*, vol. 38, pp. 2118–2126, Dec. 1990.

[6] S. Peleg and B. Friedlander, in *Proc. 26th Asilomar Conf. Sig., Syst., and Comp.*, (Pacific Grove, CA), pp. 119–122, Oct. 1992.

[7] B. Boashash, *Proc. IEEE*, vol. 80, pp. 520–568, Apr. 1992.

[8] C. S. Ramalingam, *Analysis of Nonstationary Multicomponent Signals with Applications to Speech.* PhD thesis, University of Rhode Island, Kingston, RI, December 1994 (expected completion).

[9] B. Friedlander and J. Francos, *Proc. 27th Asilomar Conf. Sig., Syst., and Comp.*, (Pacific Grove, CA), pp. 431–435, Nov. 1993.

[10] R. Kumaresan and C. S. Ramalingam, in *Proc. 27th Asilomar Conf. Sig., Syst., and Comp.*, (Pacific Grove, CA), pp. 1041–1046, Nov. 1993.

[11] C. S. Ramalingam and R. Kumaresan, in *Proc. IEEE ICASSP-94*, (Adelaide, Australia), pp. I:473–476, Apr. 1994.

[12] R. Kumaresan, C. S. Ramalingam, and A. Rao, in *Proc. of the Seventh SSAP Workshop*, (Québec City, Canada), pp. 207–210, Jun. 1994.

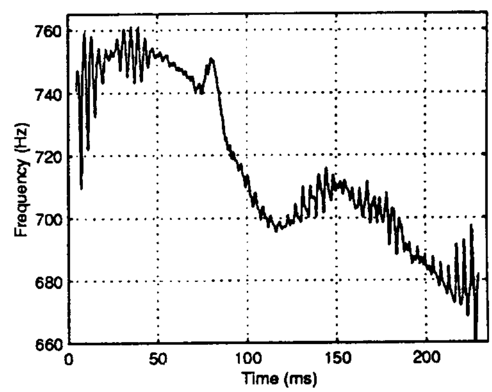
[13] J. P. Costas, in *Proc. of the First ASSP Workshop on Spectral Estimation*, (Hamilton, Canada), pp. 6.5.1–6.5.8, Aug. 1981.

[14] K. S. Riedel, *IEEE Trans. SP*, vol. 42, pp. 2644–2649, Oct. 1994.

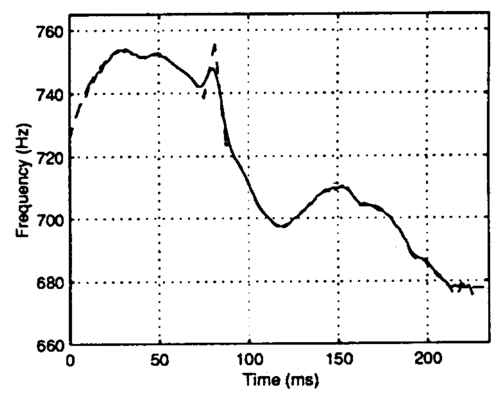
[15] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and its Applications.* New York, NY: John Wiley & Sons, 1971.

[16] R. Kumaresan, "Spectral Analysis," in *Handbook for Digital Signal Processing* (S. K. Mitra and J. F. Kaiser, eds.), New York, NY: John Wiley & Sons, 1993.

[17] F. B. Hildebrand, *Introduction to Numerical Analysis.* New York, NY: McGraw-Hill, second ed., 1974.



(a)



(b)

Figure 2: (a) The unsmoothed frequency track of the third component in Example 3.2 obtained by fitting a constant envelope, constant frequency model over short, heavily overlapping blocks. The estimated parameters were assigned to the middle sample. (b) The solid line is the result of smoothing the track shown in (a). The dashed line represents the true track.