

RESEARCH

Open Access

Basis approach to estimate the instantaneous frequencies in multicomponent AM-FM signals

Mohammadali Sebghati¹, Hamidreza Amindavar^{1*} and James A Ritcey²

Abstract

In this paper, an analytical approach to estimate the instantaneous frequencies of a multicomponent signal is presented. A non-stationary signal composed of oscillation modes or resonances is described by a multicomponent AM-FM model. The proposed method has two main stages. At first, the signal is decomposed into its oscillation components. Afterwards, the instantaneous frequency of each component is estimated. The decomposition stage is performed through the basis expansion exploiting orthogonal rational functions in the complex plane. Orthogonal rational bases are generalized to expand linear time-varying systems. To decompose the non-stationary signal, its equivalent time-varying system is sought. The time-varying poles of this system are required to construct appropriate basis functions. An adaptive data segmentation algorithm is provided for this purpose. The effect of noise is scrutinized analytically and evaluated experimentally to verify the robustness of the new method. The performance of this method in extraction of embedded instantaneous frequencies is asserted by simulations on both synthetic data and real-world audio signal.

Keywords: Instantaneous frequency estimation; AM-FM modeling; Multicomponent signals; Oscillation modes; Orthogonal rational functions

1 Introduction

Non-stationary signals which are a compound of constituents with time-varying amplitudes and frequencies can be characterized by amplitude-modulated frequency-modulated (AM-FM) models. This modeling is attended to express genuine signals in communications [1,2], acoustic and speech processing [3-6], biomedical signal processing [7], and image processing [8]. To estimate instantaneous amplitudes (IAs) and instantaneous frequencies (IFs) embedded in a multicomponent signal [9], the first step is decomposing it into oscillation components. This procedure is termed 'demodulation', 'separation' or 'decomposition' in the literature. Each component should represent a valid oscillation for which the definition of instantaneous frequency is physically meaningful [10].

All methods of multicomponent AM-FM signal decomposition and parameter estimation can be categorized into non-parametric and parametric methods. One class of

non-parametric approaches is based on the joint time-frequency processing. Widespread time-frequency distributions (TFDs) such as short-time Fourier transform (STFT), Wigner-Ville distribution, and Choi-Williams distribution are employed [11]. These methods are limited by the well-known compromise between time and frequency resolution. Moreover, the cross-terms appear troublesome. The energy separation algorithm (ESA) which uses a non-linear differential operator, called Teager-Kaiser energy operator (TKEO), is another method [12]. The energy separation algorithm which tracks the energy of the source producing the signal is originally applicable for monocomponent AM-FM signals. Nevertheless, they are modified for application in multicomponent cases by designing a bank of filters. Multiband-ESA (MESA) that consists of bandpass filtering followed by monocomponent energy separation is introduced based on this concept [13]. The separation performed by bandpass filtering is proper when components are spectrally far enough.

Huang et al. [14] proposed an iterative technique known as the empirical mode decomposition (EMD). This technique is an algorithmic way to extract oscillation modes embedded in the signal, named intrinsic mode functions

*Correspondence: hamidami@aut.ac.ir

¹Electrical Engineering Department, Amirkabir University of Technology, Tehran 15914, Iran

Full list of author information is available at the end of the article

(IMFs). Each IMF gives a valid IF which is estimated by applying the Hilbert transform (HT) [14]. The proposed algorithm for the implementation of EMD called the sifting process showed several drawbacks such as sensitivity to perturbation and mode mixing problem [15]. To overcome these difficulties, the original EMD is modified, and a new algorithm is developed which is named ensemble empirical mode decomposition (EEMD) [15]. This method is indeed the iteration of EMD for noise-added signals. In each iteration, controlled white noise is added to the data, and EMD is applied. Each individual trial may generate noisy results, but the noise is canceled out by taking the average of the results. Thus, true solution is the ensemble mean of enough trials. Although each trial produces a set of IMFs, the sum of IMFs is not necessarily an IMF. An empirical solution for this issue is suggested in [15]. EMD and EEMD methods suffer from the lack of analytic foundation. Some research has attempted to establish and improve analytical aspects of these empirical approaches [16]. In [17], an alternative algorithm for EMD is introduced based on iterating certain filters, such as Toeplitz filters. The results of iterative filtering are similar to those of the conventional sifting process. Although the authors of [17] do not claim superiority for their method, it lays down a mathematical framework for an alternative approach to EMD. The convergence of iterative filtering EMD is studied in [18]. A variant of EMD to decompose multiscale data is proposed in [19]. This work provides some theoretical understanding of EMD for a class of multiscale data and introduces two algorithms, Newton-Raphson-based EMD and ODE-based EMD, as the variations of the sifting process. The decomposition of multiscale data based on EMD is pursued in [20] inspired by the compressed sensing theory. The sparsest representation of multiscale data is sought within the largest possible dictionary constructed of IMFs. The problem is formulated as a non-linear L^1 optimization, and an iterative algorithm is proposed to solve it. Noise and perturbation in data may cause numerical instability in this method. Daubechies et al. developed a method which captures the philosophy of EMD and decomposes special functions in a defined class [21]. This method employs a combination of wavelet analysis and reallocation technique called synchrosqueezing transform which aim to sharpen a time-frequency representation. Synchrosqueezed wavelet transform is also investigated for signal sampling and denoising applications in multicomponent signal analysis [22]. In [23], an algorithm for AM-FM parameter estimation is proposed based on the iterated application of the Hilbert transform to amplitude envelopes obtained by adaptively low-pass filters. Furthermore, the IF of AM-FM components can be calculated by *a posteriori* adaptive segmentation of the acquired phase signal. Another iterative AM-FM decomposition

is suggested in [24] using the quasi-harmonic model (QHM) for quasi-harmonic signals such as voiced speech.

There are various parametric approaches to extract IFs of multicomponent signals. One common approach is based on signal segmentation, while some simplifying assumptions such as constant frequency seem logical in each segment. Then, an estimator is designed to estimate the model parameters segment by segment. In [25], the maximum windowed likelihood (MWL) criterion is used to estimate the AM-FM components. The high non-linearity of this method makes the necessary optimization difficult. Another parametric approach is based on the statistical modeling of the signal according to its statistical attributes and assumptions. Speech signals are statistically modeled as AM-FM signals, and the extended Kalman filter (EKF) is applied for demodulation [3]. The idea of EKF is also exploited in [26]. Polynomial phase signal (PPS) modeling is another parametric approach which is employed for AM-FM signals [27].

The interpretations and estimation of instantaneous frequencies embedded in a multicomponent signal have been controversial [28]. Three different approaches are proposed to estimate the IFs after the decomposition stage. In the first approach, the Hilbert transform is exploited to get the analytic signal whose phase is differentiated to find the IF [10]. The energy operator (TKEO) is utilized in the second approach [12], and the third one is based on TFD [11,29]. Different definitions of IF are considered in these approaches, and consequently, their results are not necessarily equivalent. In [28] and [30], the different definitions and estimation methods are compared and discussed. The main contribution of this paper is to develop a novel approach based on the expansion of time-varying systems by orthogonal rational functions. The method introduced in [31] is extended and improved to be applied as the essence of the new method for IF estimation. An adaptive segmentation procedure in the proposed algorithm allows us to estimate the IFs locally. The decomposition is performed using orthogonal rational functions.

2 Problem statement

Multicomponent signals are first introduced in [9]. A multicomponent AM-FM model describes a non-stationary signal as the combination of oscillation terms with time-varying amplitudes and frequencies:

$$x(t) = \sum_{k=1}^{N_c} A_k(t)e^{j\theta_k(t)}, \quad (1)$$

where N_c is the number of components. $A_k(t)$ and $\theta_k(t)$ are time-varying envelope and time-varying phase of the

k th component respectively, and the instantaneous frequency denoted by $f_k(t)$ is defined from $\theta_k(t)$:

$$f_k(t) = \frac{d\theta_k(t)}{dt}. \quad (2)$$

The general model in (1) can be interpreted as the signal expansion by a generalized complex exponential basis, which are exponential functions with time-varying amplitudes and frequencies. The decomposition of the multicomponent AM-FM signal is investigated through this point of view. Therefore, we are going to find an appropriate basis to expand the non-stationary signal $x(t)$:

$$x(t) = \sum_{k=1}^K c_k g_k(t). \quad (3)$$

The functions $\{g_k; k = 1, 2, \dots, K\}$ should represent the oscillation modes in signal, for which the instantaneous frequency is meaningfully definable. Accordingly, each term represents a valid IF of the multicomponent signal. The main idea to attain such decomposition is expanding the corresponding system of the AM-FM signal in the complex plane. Since the transfer function of a realistic linear system has a rational representation, it can be expanded by orthogonal rational functions in the complex z -plane. Returning back to the time-domain, each rational function is equivalent to a generalized exponential basis and represents one valid oscillation term or resonance. To perform this procedure, we should specify the generating system of the AM-FM signal. The corresponding system of a non-stationary signal is modeled by a linear time-varying (LTV) system [32]. LTV models have been applied to describe non-stationary signals [33,34]. Our proposed method is developed based on this approach of modeling. Let us consider the discrete-time AM-FM signal $x[n]$, obtained by time sampling of $x(t)$ at the rate of f_s . Its generating system is modeled as a LTV system, which can be described by a bivariate function to characterize the input-output linear relationship [32]. Hence, a bivariate discrete-time impulse response, $h[m, n]$, is considered, where n and m are two independent time instants, representing the time variable of the signal and the time variable of the system, respectively. Taking the Z -transform of $h[m, n]$ with respect to the time variable of the signal, $H(m, z)$ is obtained which denotes the generating system of $x[n]$. The orthogonal rational basis has been investigated for the decomposition of linear time-invariant (LTI) systems [35] and should be generalized to expand the time-varying generating system of the AM-FM signal as follows:

$$H(m, z) = \sum_{k=1}^K C_k(m) G_k(m, z), \quad (4)$$

where $\{G_k(m, z); k = 1, 2, \dots, K\}$ is a rational basis. Fortunately, to find the orthogonal rational functions for system expansion, it is not necessary to have the system's transfer function with all the details. The knowledge of poles or logical assumptions about them are sufficient to extract a proper basis [35].

3 Orthogonal rational functions

The decomposition step in the proposed method is indeed an expansion of the AM-FM signal, which is accomplished through the incorporating expansion of the equivalent time-varying system by orthogonal rational functions. Generally, the knowledge about the poles is sufficient as *a priori* information to describe the desired space being spanned by a rational basis. Let us consider a set of M time-varying poles, $\{\xi_k[m], k = 1, \dots, M\}$. We can make a first-order IIR transfer function by each pole. So, a basis set is constructed including all specified poles, but not orthogonal. The Blaschke products [35] formed by these poles are two-dimensional functions,

$$B_0(m, z) = 1, \\ B_k(m, z) = \prod_{i=1}^k \frac{1 - \bar{\xi}_i[m]z}{z - \xi_i[m]}, \quad k = 1, 2, \dots, M. \quad (5)$$

Applying the Gram-Schmidt procedure on these rational functions with respect to z , two-dimensional functions are obtained:

$$G_k(m, z) = \frac{\sqrt{1 - |\xi_k[m]|^2}}{z - \xi_k[m]} \prod_{i=1}^{k-1} \left(\frac{1 - \xi_i[m]^* z}{z - \xi_i[m]} \right), \quad (6) \\ k = 1, \dots, M.$$

This is the same routine for finding Takenaka-Malmquist functions [36]. Now, these functions are generalized to two-dimensional functions for the LTV system expansion. The resultant functions in (6) are orthogonal with respect to z in the complex domain. The inner product of each pair of these functions is a function of time, m :

$$d_{kl}[m] = \langle G_k(m, z), G_l(m, z) \rangle_z \\ = \frac{1}{2\pi j} \oint_C G_k(m, z) G_l^* \left(m, \frac{1}{z^*} \right) \frac{dz}{z}. \quad (7)$$

Utilizing the Cauchy integral implies that $d_{kl}[m]$ is zero at each snapshot, m , for $k \neq l$. Taking the inverse Z -transform of $G_k(m, z)$ produces $g_k[m, n]$. Since the Z -transform and its inverse are homomorphic transforms, these functions would preserve orthogonality with respect to n . Two problems remain to be studied. At first, the underlying time-varying poles of the corresponding LTV

system should be determined. Secondly, univariate terms should be extracted from bivariate functions, $g_k[m, n]$, to express the oscillation modes of $x[n]$. These issues are resolved simultaneously by adaptive segmentation which is addressed in the following section.

4 Time-varying modeling

The concept of poles and zeros are also generalized for linear time-varying systems. Concerning the stability and behavior of the LTV systems, several definitions of poles or eigenvalues of such systems have been proposed [37], depending on the characterization method of the LTV system. The notion of time-varying poles in this paper is founded on the time-varying autoregressive model. Parametric models for LTI systems can be generalized for LTV ones by imposing time-varying parameters on the model. The AM-FM signal, $x[n]$, is modeled by a time-varying autoregressive (TVAR) of order M :

$$x[n] = - \sum_{m=1}^M a_m[n] x[n-m] + v[n]. \quad (8)$$

$\{a_m[n], m = 1, \dots, M\}$ are the time-varying parameters, and $v[n]$ is the zero-mean innovation process, also addressed as a modeling error. The most general case of this model is where the parameters are completely uncorrelated at each time sample. Therefore, each time sample of $x[n]$ would be represented by M unknown coefficients; hence, it is not a practical approach. Based on a common practical assumption, the non-stationary signal is approximately regarded locally stationary or quasi-stationary. This assumption implies that the parameters of the TVAR model are correlated, and the coefficients are supposed to be constant in subintervals of the total time span, referred to as segments. This model is called a block stationary AR model [34]. For multicomponent AM-FM signals whose IAs and IFs are slowly time-varying or piecewise-constant, the segmentation strategy is applicable. By virtue of this assumption, a real multicomponent AM-FM signal over its support is considered as a superposition of temporarily more limited signals with constant frequencies. These intervals can generally have various lengths, and different methods from fixed-length windowing to adaptive segmentation algorithms are introduced to determine the borders of the segments [23]. In the proposed method, the segmentation is performed adaptively from the aspect of TVAR parameter estimation.

4.1 Segmentation procedure

The entire signal of N samples is segmented into L blocks with various lengths:

$$x_\ell[n] = x[n], \quad n_{\ell-1} \leq n < n_\ell, \quad (9)$$

where $\ell = 1, \dots, L$ and $n_0 = 0$. The TVAR coefficients are supposed to be constant in each segment. The mean square error in the ℓ th segment is given by

$$J_\ell = \frac{1}{n_\ell - n_{\ell-1}} \sum_{n=n_{\ell-1}}^{n_\ell-1} \left| x[n] + \sum_{m=1}^M a_{\ell,m} x[n-m] \right|^2, \quad (10)$$

where $\{a_{\ell,m}, m = 1, \dots, M\}$ are the TVAR coefficients of the ℓ th segment. The boundaries of each segment are determined such that the error J_ℓ remains under a specified threshold. The segmentation algorithm operates as follows. At the start of each stage, the length of the current segment (say ℓ) is considered the minimum possible length, equal to the order of the TVAR model, i.e., $n_\ell = n_{\ell-1} + M$. The TVAR coefficients, $a_{\ell,m}$, are estimated by the recursive least squares (RLS) technique, and the error in (10) is computed. If it is still greater than the pre-specified threshold, the length of the segment increases by one sample, and the calculations are repeated. This procedure continues by one-sample increment in each stage until the error falls below the threshold. Now, the boundaries and the length of the current segment are determined, and the procedure starts over the next time sample for another segment establishment. This algorithm runs through the entire signal repeatedly and stops at the end of the data batch. The question arises here about the threshold setting and how it can affect the accuracy of the IF estimation. This issue is scrutinized in the succeeding subsection separately. Once the TVAR parameters are estimated, the corresponding time-varying poles, denoted by $\{\xi_k[m], k = 1, \dots, M\}$, are obtained by applying the Z-transform of (8) with respect to n .

4.2 Error analysis

It is noteworthy to mention the relation between the error caused by segmentation and the error of IF estimation. This analysis leads us to select a reliable error threshold in the adaptive segmentation procedure. Let us consider a discrete-time AM-FM component:

$$x[n] = A[n] e^{j\theta[n]}. \quad (11)$$

The error in the instantaneous phase imposed by the TVAR modeling in each segment, denoted by ϵ_θ , sets off an error in signal:

$$\widehat{x}[n] = A[n] e^{j(\theta[n] + \epsilon_\theta[n])}. \quad (12)$$

For very small phase errors, the following approximation is considered using the Maclaurin series:

$$e^{j(\theta[n] + \epsilon_\theta[n])} \approx \left(1 - \frac{\epsilon_\theta^2[n]}{2} + j\epsilon_\theta[n] \right) e^{j\theta[n]}. \quad (13)$$

Substituting this approximation in (12), we have

$$e[n] = x[n] - \widehat{x}[n] = \varepsilon[n] A[n] e^{j\theta[n]}, \quad (14)$$

where

$$\varepsilon[n] = \frac{\epsilon_\theta^2[n]}{2} - j\epsilon_\theta[n]. \quad (15)$$

The error $e[n]$ whose instantaneous amplitude is absolute error of modeling is also an AM-FM signal:

$$|e[n]| = |\varepsilon[n] A[n]|. \quad (16)$$

The error in phase is now transduced to the error of amplitude. Let us define a time-varying threshold denoted by $\eta[n]$ such that $|e[n]|$ is restrained lower than it, i.e., $|e[n]| < \eta[n]$. If we substitute $|e[n]|$ by (16), the following inequality holds:

$$|\varepsilon[n]| < \frac{\eta[n]}{|A[n]|}. \quad (17)$$

So, the absolute error of phase depends on the signal envelope. This means that for a fixed threshold, where $\eta[n]$ is constant over the entire signal, larger phase errors can occur when IA becomes smaller. Therefore, the threshold should vary adaptively, adjusted to the envelope of the observed signal. In other words, the locally normalized error for each segment is a proper threshold. Since the IA evolves slowly, its mean or minimum amount during the segment can be utilized for normalization. The normalized threshold is denoted by $\bar{\eta}$ for brevity:

$$\bar{\eta} = \frac{\eta[n]}{\text{mean}\{|A[n]|\}}. \quad (18)$$

Thus, the inequality (17) is practically used as the following one:

$$|\varepsilon[n]| < \bar{\eta}. \quad (19)$$

The square of $|\varepsilon[n]|$ is obtained from (15):

$$|\varepsilon[n]|^2 = \frac{\epsilon_\theta^4[n]}{4} + \epsilon_\theta^2[n]. \quad (20)$$

Exploiting this relation in the inequality (19) and performing some mathematical reformulations result in a bound for the phase error:

$$|\varepsilon_\theta[n]|^2 < 2 \left(\sqrt{1 + \bar{\eta}^2} - 1 \right). \quad (21)$$

When $\bar{\eta} \rightarrow 0$, the right-hand side of the above inequality is approximately equal to $\bar{\eta}^2$. Keeping the phase error (ϵ_θ) under control, the error of IF is consequently controlled. By definition (2), IF is the derivative of instantaneous phase, which is a difference equation in the discrete-time situation:

$$\omega[n] = \frac{\theta[n] - \theta[n-1]}{T_s}, \quad (22)$$

where $\omega[n] = 2\pi f[n]$ is the instantaneous frequency in radian per second, and T_s denotes the sampling time. In the worst case, the maximum phase errors of two consecutive instants accumulate. Thus, the maximum error of IF is $2\epsilon_\theta f_s$. For example, if $\bar{\eta} = 10^{-3}$, then from (21), the maximum phase error is almost 10^{-3} , and the absolute error of IF is at most 0.2% of the sampling frequency. This error can be controlled by arbitrary selection of $\bar{\eta}$. A smaller threshold leads to wider segments, in which the assumption of constant frequency is no longer respected. Our experiments verified that the condition of piecewise-constant frequency for slowly varying IFs is satisfied for $\bar{\eta}$ in the order of $10^{-3} \sim 10^{-2}$.

5 Estimation framework

The main algorithm of IF estimation takes the extracted time-varying poles to construct $G_k(m, z)$ in (6). Then, bivariate functions, $\{g_k[m, n], k = 1, \dots, M\}$, are produced by taking the inverse Z-transform. Now, the one-dimensional basis is extracted from the existing bivariate functions to achieve a one-dimensional expansion for $x[n]$ as in (3). The basis $g_k[n]$ is constructed by the concatenation of truncated pieces of bivariate functions, $g_k[m, n]$, based on the result of the segmentation procedure:

$$g_k[n] = \sum_{\ell=1}^L W_\ell[n] g_k[m_\ell, n], \quad (23)$$

where L is the number of total segments, and $W_\ell[n]$ is an arbitrary weighting window over the ℓ th segment. It is supposed that during this block, the corresponding pole remains equal to $\xi_k[m_\ell]$. Each resultant function, $g_k[n]$, is a valid oscillation mode for which IF is definable. Thus, the estimation of the embedded IFs is achieved through the IF estimation of the extracted functions. To estimate the IF, linear regression of the phase for each segment is computed by applying the weighted linear least squares technique on a first-order polynomial model. The abrupt changes in the phase which can affect the IF estimation severely are an important issue. While the consecutive segments derived from different rows of $g_k[m, n]$ are concatenated, there may be some phase discontinuities over the resultant bases in the junctions of segments. Such discontinuities in the phase trajectory cause serious deficiency in the IF estimation which appears as spikes over the resultant IFs. To remedy this problem, a proper data window such as the Hamming window is chosen as $W_\ell[n]$ in (23), which controls the effect of borderline samples. The Hamming window is commonly utilized as an analysis window in audio and speech processing [24,27].

When the signal is contaminated by noise, the time-varying poles estimated from noisy observations are

misplaced. Thus, the estimated IF incurs more error due to the error in the estimation of poles imposed by the noise. This issue is alleviated by increasing the order of the TVAR model. Each resonance of a clean signal is represented by a pair of time-varying poles; hence, the order of the TVAR model (M) is twice the number of components (N_c). Nonetheless, to improve the estimation of time-varying poles in the presence of noise, we should have $M > 2N_c$. Therefore, extraneous poles appear besides the valid poles. A minimum distance classifier [38] is applied to assign the poles of the resonances in each segment and distinguish them from the invalid poles. The perturbation of the poles due to the estimation error of TVAR coefficients is investigated mathematically in the succeeding section. The steps of the proposed algorithm are summarized as follows:

1. Adaptive segmentation of the AM-FM signal based on TVAR modeling and computation of underlying time-varying poles.
2. Assignment of the poles to the components by minimum distance classifier.
3. Employing the time-varying poles to construct oscillation terms, $g_k[n]$.
4. Fitting a linear model to the phase of each segment of $g_k[n]$ to estimate the IF.

In this novel method of IF extraction, the Hilbert transform that is a global operator is not employed. Additionally, a linear model is applied to the phase of components segment by segment in spite of differentiating throughout. It makes the proposed method less sensitive to phase changes. Therefore, the adaptive segmentation is advantageous for both decomposition and frequency estimation.

6 Pole perturbation

The coefficients of the TVAR model (8), estimated through the RLS technique, are affected by noise. The perturbation in these coefficients leads to the perturbation in the resultant time-varying poles. Let $p(\ell, z)$ be the polynomial of AR model in the ℓ th segment whose roots are time-varying poles:

$$p(\ell, z) = \sum_{i=0}^M a_{\ell,i} z^{-i} = \prod_{k=1}^M (1 - \xi_k[m_\ell] z^{-1}), \quad (24)$$

where the coefficients are normalized, i.e., $a_{\ell,0} = 1$. If the perturbation $\Delta a_{\ell,i}$ occurs in the i th coefficient, $a_{\ell,i}$, the polynomial (24) changes to

$$\tilde{p}(\ell, z) = p(\ell, z) + \Delta a_{\ell,i} z^{-i}. \quad (25)$$

The roots of this new polynomial differ from $\xi_k[m_\ell]$ by $\Delta \xi_k[m_\ell]$, which may be real or complex. Let us denote the perturbed roots by $\tilde{\xi}_k[m_\ell] = \xi_k[m_\ell] + \Delta \xi_k[m_\ell]$:

$$\tilde{p}(\ell, \tilde{\xi}_k[m_\ell]) = p(\ell, \tilde{\xi}_k[m_\ell]) + \Delta a_{\ell,i} (\tilde{\xi}_k[m_\ell])^{-i} = 0. \quad (26)$$

From now on, the k th pole and its perturbation are denoted by ξ_k and $\Delta \xi_k$, respectively, and their arguments are neglected for brevity. When $\Delta \xi_k \rightarrow 0$, the above equation is simplified by employing Taylor's expansion:

$$p'(\ell, \xi_k) \Delta \xi_k + \Delta a_{\ell,i} \xi_k^{-i} = 0, \quad (27)$$

where $p'(\ell, z)$ is the first derivative of $p(\ell, z)$ with respect to z . This equation expresses the linear relation between perturbation in poles and perturbation in coefficients:

$$\Delta \xi_k = -\frac{\xi_k^{-i}}{p'(\ell, \xi_k)} \Delta a_{\ell,i}. \quad (28)$$

Considering $\Delta a_{\ell,i}$ and $\Delta \xi_k$ as random variables, their variances, respectively $\sigma_{a_{\ell,i}}^2$ and $\sigma_{\xi_k}^2$, are related linearly,

$$\sigma_{\xi_k}^2 = \left| \frac{\xi_k^{-i}}{p'(\ell, \xi_k)} \right|^2 \sigma_{a_{\ell,i}}^2. \quad (29)$$

Obtaining $p'(\ell, \xi_k)$ from (24), the ratio of the variances is given by

$$\frac{\sigma_{\xi_k}^2}{\sigma_{a_{\ell,i}}^2} = \frac{1}{\left| \xi_k^{i-1} \prod_{\substack{j=1 \\ j \neq k}}^M (1 - \xi_k \xi_j^{-1}) \right|^2}. \quad (30)$$

This ratio indicates the sensitivity of the poles to the perturbation of the coefficients. Smaller ratio represents less sensitivity or more robustness of pole. When all poles are inside the unit circle, i.e., $|\xi_k| < 1, k = 1, \dots, M$, the ratio in (30) takes its maximum value for $i = M$. This means that the perturbation in the coefficient of the highest order in the polynomial has the most effect on misplacing the poles. This maximum value for each pole is defined as its variance ratio:

$$\gamma_k = \frac{\sigma_{\xi_k}^2}{\sigma_{a_{\ell,M}}^2} = \frac{1}{\left| \xi_k^{M-1} \prod_{\substack{j=1 \\ j \neq k}}^M (1 - \xi_k \xi_j^{-1}) \right|^2}. \quad (31)$$

γ_k is the variance ratio of ξ_k which indicates the robustness of this pole. This parameter only depends on the positions of the underlying poles, which are determined by AM-FM signal characteristics.

7 Experimental evaluation

The proposed method is implemented on both synthetic and real-world signals, and its performance is compared to the results of two previously introduced methods.

7.1 Synthetic data

A two-component AM-FM signal is considered as below:

$$x(t) = \sum_{k=1}^2 \exp(\alpha_k t) \exp[j2\pi(\beta_{k,3}t^3 + \beta_{k,2}t^2 + \beta_{k,1}t)]. \quad (32)$$

The parameters of the first and the second components are $\alpha_1 = -2$, $\beta_{1,1} = 10^3$, $\beta_{1,2} = 5 \times 10^3$, $\beta_{1,3} = -20 \times 10^3$ and $\alpha_2 = -5$, $\beta_{2,1} = 5 \times 10^3$, $\beta_{2,2} = -3 \times 10^3$, $\beta_{2,3} = 0$. The two following IFs are embedded in this AM-FM signal:

$$\begin{aligned} f_1(t) &= -60t^2 + 10t + 1 \text{ kHz}, \\ f_2(t) &= -6t + 5 \text{ kHz}. \end{aligned} \quad (33)$$

The signal in (32) is sampled at $T_s = 50 \mu\text{s}$ intervals, so the sampling frequency is $f_s = 20 \text{ kHz}$. Since the absolute values of the instantaneous frequencies increase over time, a limited span of signal is observed to avoid aliasing. The algorithm is run over $N = 2,000$ samples of the signal. The corresponding system has four time-varying complex poles $\{\xi_1[m], \dots, \xi_4[m]\}$ in conjugate pairs. There are two conjugate pairs, and each pair represents one oscillation component. Consequently, four complex functions, $g_1[n], \dots, g_4[n]$, are extracted whose phases yield desired IFs. $g_1[n]$ and $g_2[n]$ are common resonances, which means that their IFs have equal absolute values, but opposite signs. $g_3[n]$ and $g_4[n]$ determine the second IF likewise. The adaptive segmentation procedure divides the data batch into $L = 45$ segments with different lengths. The threshold of error in the segmentation procedure is set to 0.001. The coefficients of the TVAR model and, accordingly, the time-varying poles are estimated through the RLS algorithm with the forgetting factor of 0.98. Figure 1 demonstrates the estimated IFs denoted by \hat{f}_i ; $i = 1, 2$ besides the original IFs. \hat{f}_1 and \hat{f}_2 track the original IFs very closely. The step-like variation, produced by segmentation, is obvious. For quantitative evaluation, the relative mean absolute error (MAE) of IF estimation is computed for \hat{f}_1 and \hat{f}_2 , and illustrated in Tables 1 and 2, respectively.

The proposed method is compared with two previous methods, EEMD-HT and QHM. The EEMD-HT is a non-parametric method which decomposes the components by the EEMD algorithm and estimates the IFs of resultant IMFs utilizing the Hilbert transform [14]. The QHM is a parametric method which has been appraised on speech signals [24]. The result of decomposition of the original signal by the EEMD procedure is displayed in Figure 2. The relative standard deviation of added noise is 0.2, and the ensemble number for each run is 100. Nine IMFs are derived while there are just two embedded oscillation modes. The first and the second IMFs are expected

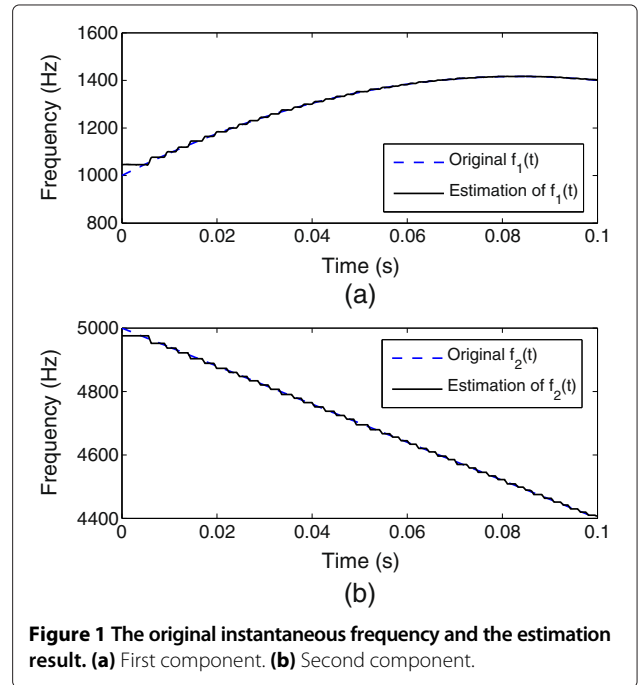


Figure 1 The original instantaneous frequency and the estimation result. (a) First component. (b) Second component.

oscillation modes, and the others are false IMFs generated due to deficiencies in the iterative algorithm. EEMD initially extracts faster oscillations. Thus, the first IMF has higher frequency and represents our second component, and reversely, the second IMF corresponds to the first component. The valid IMFs are selected and assigned subjectively or based on the result of estimation [15]. In this simulation, the estimated IFs of the first and the second IMFs are closer to f_2 and f_1 , respectively. MAEs of the estimation of IFs are recorded in Tables 1 and 2. The estimation errors of QHM are also provided in these tables for comparison. The analysis window of 64 samples and hop step of one sample is considered for this algorithm.

The simulation is repeated in the presence of additive white-Gaussian noise, and the errors for different SNRs are depicted in the same tables. Each value is the average of 100 iterations. The proposed method outperforms the other algorithms, especially in stronger noise. These results verify the robustness of the proposed algorithm. Although the QHM has smaller error for clean signal, it is more sensitive to noise. The EEMD, which is more robust than EMD, is still affected by the perturbations on

Table 1 Relative MAE (%) for the estimation of the first instantaneous frequency, $f_1(t)$

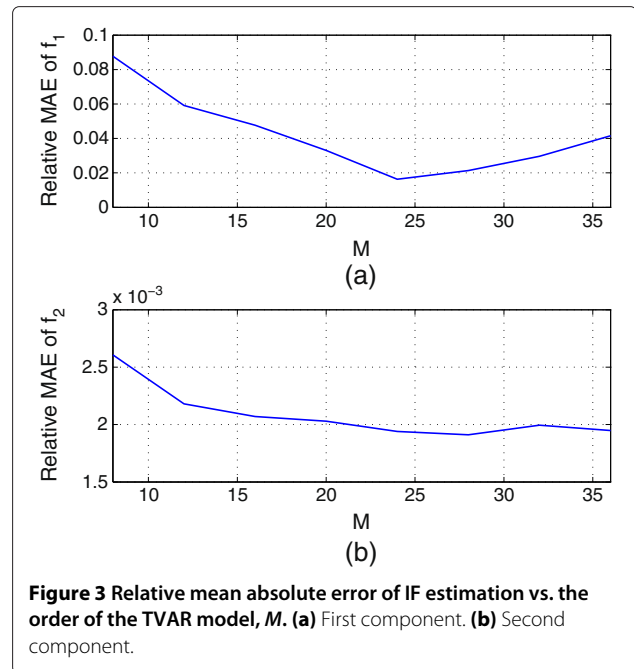
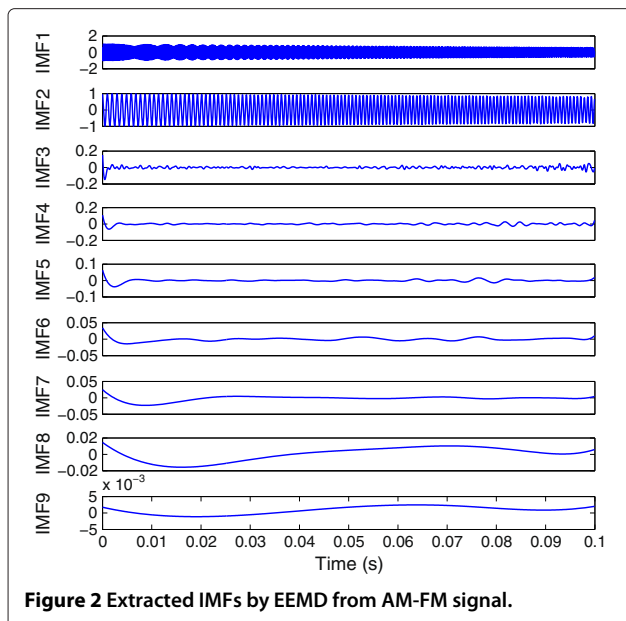
SNR	EEMD-HT	QHM	Proposed method
∞	0.57	0.06	0.53
30 dB	1.49	0.83	1.62
10 dB	79.68	44.03	40.74

Table 2 Relative MAE (%) for the estimation of the second instantaneous frequency, $f_2(t)$

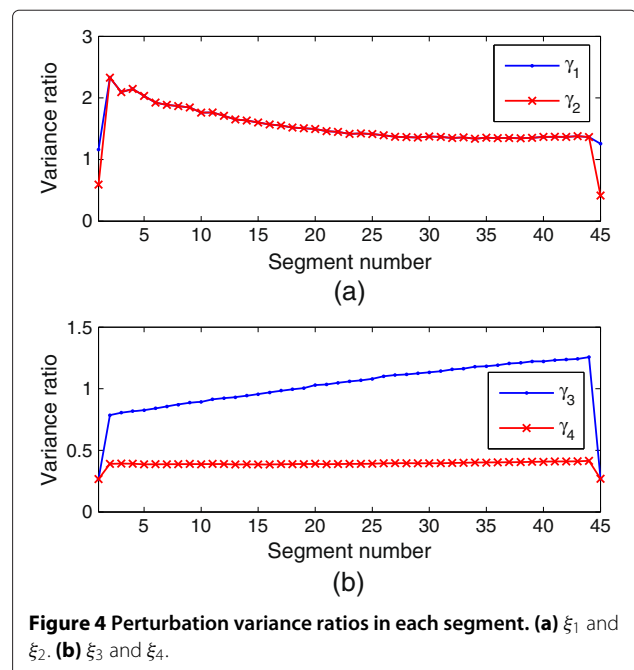
SNR	EEMD-HT	QHM	Proposed method
∞	0.17	0.03	0.09
30 dB	0.46	0.24	0.19
10 dB	21.54	8.30	3.23

the amplitude of the data. Therefore, EEMD-HT has the worst performance in the presence of noise. For the clean signal ($\text{SNR} = \infty$), the order of the TVAR model, M , is equal to 4. Since higher error is imposed on pole estimation for stronger noise, M should increase to alleviate this problem. We have $M = 24$ for $\text{SNR} = 30$ dB and $\text{SNR} = 10$ dB. Figure 3 depicts the MAE of IF estimation with respect to M for both components. It demonstrates that the error decreases remarkably as the order increases. However, there are more poles for higher orders which raise the error of classification routine and consequently the total error of estimation. Thus, there is a compromise to select the optimum order. In Figure 3, MAEs of \hat{f}_1 and \hat{f}_2 are minimum at $M = 24$ and $M = 28$, respectively.

The effect of noise is illustrated by the perturbation variance ratios, $\gamma_1, \dots, \gamma_4$ (31) for each segment in Figure 4. The variance ratios of the poles in a conjugate pair are very close to each other or even equal in some segments. γ_3 and γ_4 , corresponding to ξ_3 and ξ_4 or the second resonance, are smaller than γ_1 and γ_2 . This means that the poles corresponding to the first oscillation component are more sensitive to the noise. This observation justifies why the MAE for \hat{f}_1 (Table 1) deteriorates more than the MAE of \hat{f}_2 (Table 2) due to the noise.



The proposed method can distinguish components and estimate embedded close IFs, even when they have crossovers. To illustrate this capability, the simulation of a similar two-component signal is repeated, except that the frequency of the second component, f_2 , is reduced by 4.8 kHz, i.e., $\beta_{2,1} = 1,200$ in (32). In this case, two IFs are closer and intersect each other. Both IFs are depicted in Figure 5. The proposed method is implemented in the same condition, and the results of estimation are exhibited



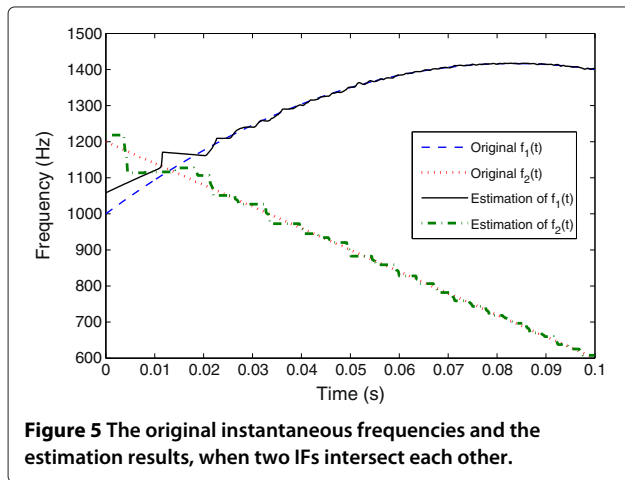


Figure 5 The original instantaneous frequencies and the estimation results, when two IFs intersect each other.

in Figure 5, besides the original IFs. This method can track both IFs, and the intersection in their trajectories is handled properly. The degradation of estimation, especially around the crossover junction, is obvious by comparing Figures 5 and 1, but the trajectories are not missed and the estimated IFs are improved with time. This advantage arises from the step of pole classification in the proposed algorithm.

7.2 Real-world signal

Acoustic studies have revealed that many natural acoustic signals such as the sounds of songbirds and oceanic mammals fit into the AM-FM model [27,39]. A duration of 68 ms of a song of two songbirds, a canary, and a kinglet, recorded at 22.05 kHz sampling frequency, is considered for the experiment. The spectrum of this signal estimated by STFT with sliding window of 128 samples, an overlap of 100 samples, and frequency resolution bin of 2 Hz is depicted in Figure 6. This signal is composed of two AM-

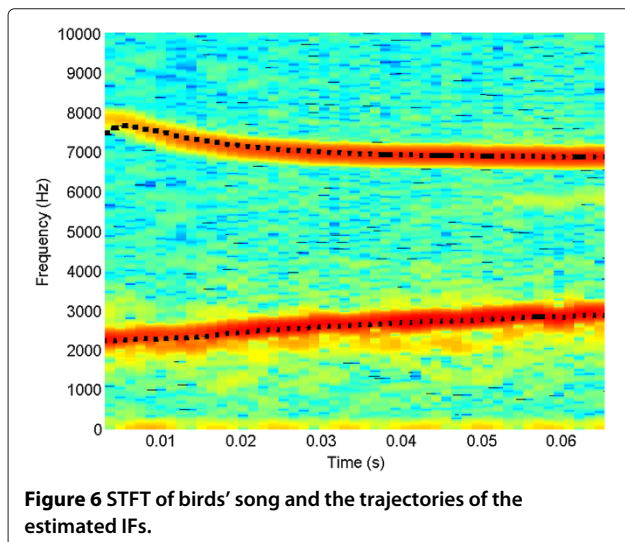


Figure 6 STFT of birds' song and the trajectories of the estimated IFs.

FM components and real-world disturbing signals such as ambient and instrumental noises. Four complex poles and consequently four bases are extracted which present two embedded dominant oscillations. Through the adaptive segmentation procedure, the signal is divided into $L = 76$ unequal blocks. The threshold of the segmentation procedure, $\bar{\eta}$, is set to 0.01. The extracted IFs are illustrated in the same figure of the STFT spectrum by black marks. The estimated IF tracks the frequency of dominant oscillations. In Figure 7, the IFs are demonstrated in addition to the results of EEMD-HT and QHM. Nine IMFs are extracted through EEMD, but the first and the second ones are valid resonances whose IFs are displayed. The relative standard deviation of added noise is set to 0.2, and the number of ensembles is 100. It takes several samples for EEMD-HT to resolve two IFs. Furthermore, the local variations and spikes in the results of EEMD-HT and QHM are noticeable. By contrast, the IFs of the proposed method are piecewise-constant, and no spike appears due to the segmentation. The lengths of the segments and consequently the variations of the resultant IFs (flatness or volatility) are controllable by the threshold selected in the segmentation procedure.

8 Conclusion

The analytical approach developed for extraction of oscillation modes of multicomponent non-stationary signals is founded on several facts which are summarized herein. Firstly, the decomposition of a multicomponent signal is considered as a signal expansion and investigated from this point of view. Secondly, the equivalent time-varying system of the non-stationary signal is modeled by TVAR and expanded by the orthogonal rational functions. Finally, oscillation modes are constructed, employing these bivariate rational functions, and their IFs are estimated. For this purpose, the evolution of the time-varying

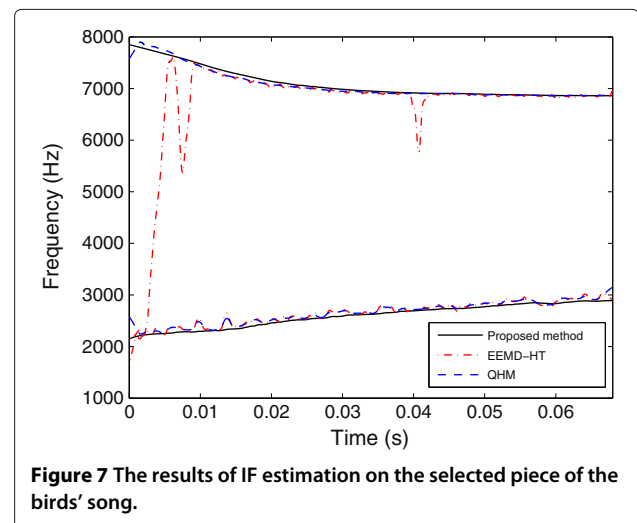


Figure 7 The results of IF estimation on the selected piece of the birds' song.

poles is considered piecewise-constant which imposes the same approximation in frequency estimation. The threshold parameter in the adaptive segmentation algorithm controls this approximation and governs the flatness of estimated IF against volatility.

The order of the TVAR model or equivalently the number of time-varying poles is another parameter which is set properly to confront the noise. Because this method utilizes the poles of the generating system of the AM-FM signal, the distortions on the amplitude, like noise, affect the estimation results mildly in comparison to those of the empirical methods. Although the noise degrades the estimation of underlying poles, it can be controlled by the order of the TVAR model. Simulations reveal the superiority of this method in the presence of noise. The controlling parameters in the proposed method yield some degree of freedom to adjust it for different realistic applications. Its capability to extract the embedded IFs in audio signals is illustrated.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Electrical Engineering Department, Amirkabir University of Technology, Tehran 15914, Iran. ²Electrical Engineering Department, University of Washington, Seattle, WA 98195, USA.

Received: 22 October 2013 Accepted: 20 February 2014

Published: 5 March 2014

References

1. B Santhanam, P Maragos, Multicomponent AM-FM demodulation via periodicity-based algebraic separation and energy-based demodulation. *IEEE Trans Commun.* **48**(3), 473–490 (2000)
2. S Jang, P Loughlin, AM-FM interference excision in spread spectrum communications via projection filtering. *EURASIP J Appl. Signal Process.* **4**, 239–248 (2001)
3. S Lu, P Doerschu, Nonlinear modeling and processing of speech based on sums of AM-FM formant models. *IEEE Trans. Signal Process.* **44**(4), 773–782 (1996)
4. M Bester, P Collen, G Richard, B David, Estimation of frequency for AM-FM models using the phase vocoder framework. *IEEE Trans. Signal Process.* **56**(2), 505–517 (2008)
5. M Grimaldi, F Cummins, Speaker identification using instantaneous frequencies. *IEEE Trans. Audio Speech Lang. Process.* **16**(6), 1097–1111 (2008)
6. Y Kubo, S Okawa, A Kurematsu, K Shirai, Temporal AM-FM combination for robust speech recognition. *Speech Commun.* **53**, 716–725 (2011)
7. C Loizou, V Murray, M Pattichis, I Seimenis, M Pantziaris, C Pattichis, Multiscale amplitude-modulation frequency-modulation (AM-FM) texture analysis of multiple sclerosis in brain MRI images. *IEEE Trans. Inform. Technol. Biomed.* **15**(1), 119–129 (2011)
8. V Murray, P Rodriguez, M Pattichis, Multiscale AM-FM demodulation and image reconstruction methods with improved accuracy. *IEEE Trans. Image Process.* **19**(5), 1138–1152 (2010)
9. L Cohen, What is a multicomponent signal? Paper presented at the of IEEE international conference on acoustic, speech and signal processing (ICASSP) (USA, 23–26 Mar 1992), pp. 113–116
10. B Boashash, Estimating and interpreting the instantaneous frequency of a signal-part 1: fundamentals. *Proc. IEEE.* **80**(4), 520–538 (1992)
11. B Lovell, R Williamson, B Boashash, The relationship between instantaneous frequency and time-frequency representations. *IEEE Trans. Signal Process.* **41**(3), 1458–1461 (1993)
12. P Maragos, F Kaiser, T Quatieri, Energy separation in signal modulations with application to speech analysis. *IEEE Trans. Signal Process.* **41**(10), 3024–3051 (1993)
13. A Bovik, P Maragos, T Quatieri, AM-FM energy detection and separation in noise using multiband energy operator. *IEEE Trans. Signal Process.* **41**(12), 3245–3265 (1993)
14. N Huang, Z Shen, S Long, M Wu, H Shih, Q Zheng, N Yen, C Tung, H Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. London.* **454**, 903–995 (1998)
15. Z Wu, N Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Adv. Adaptive Data Anal.* **1**(1), 1–41 (2009)
16. E Deléché, J Lemoine, O Niang, Empirical mode decomposition: an analytical approach for sifting process. *IEEE Signal Process. Lett.* **12**(11), 764–767 (2005)
17. L Lin, Y Wang, H Zhou, Iterative filtering as an alternative algorithm for empirical mode decomposition. *Adv. Adaptive Data Anal.* **1**(4), 543–560 (2009)
18. Y Wang, Z Zhou, On the convergence of iterative filtering empirical mode decomposition. *Excursions in Harmonic Anal.* **2**, 157–172 (2013)
19. T Hou, M Yan, A variant of the EMD method for multi-scale data. *Adv. Adaptive Data Anal.* **1**(4), 483–516 (2009)
20. T Hou, Z Shi, Adaptive data analysis via sparse time-frequency representation. *Adv. Adaptive Data Anal.* **3**(1), 1–28 (2011)
21. I Daubechies, J Lu, H Wu, Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool. *Appl. Comput. Harmonic Anal.* **30**(2), 243–261 (2011)
22. S Meignen, T Oberlin, S McLaughlin, A new algorithm for multicomponent signals analysis based on synchrosqueezing: with an application to signal sampling and denoising. *IEEE Trans. Signal Process.* **60**(11), 5787–5798 (2012)
23. F Gianfeli, G Biagetti, P Crippa, C Turchetti, Multicomponent AM-FM representations: an asymptotically exact approach. *IEEE Trans. Audio Speech Lang. Process.* **15**(3), 823–837 (2007)
24. Y Pantazis, O Rosenc, Y Stylianou, Adaptive AM-FM signal decomposition with application to speech analysis. *IEEE Trans. Audio Speech Lang. Process.* **19**(2), 290–300 (2011)
25. S Gazor, R Rashidi Far, Adaptive maximum windowed likelihood multicomponent AM-FM signal decomposition. *IEEE Trans. Audio Speech Lang. Process.* **14**(2), 479–491 (2006)
26. W Pai, C Doerschuk, Statistical AM-FM models, extended Kalman filter demodulation, Cramer-Rao bounds and speech analysis. *IEEE Trans. Signal Process.* **48**(8), 2300–2313 (2000)
27. M Jabloun, F Leonard, M Vieira, N Martin, A New flexible approach to estimate the IA and IF of nonstationary signals of long-time duration. *IEEE Trans. Signal Process.* **55**(7), 3633–3644 (2007)
28. N Huang, Z Wu, S Long, K Arnold, Z Chen, K Blank, On instantaneous frequency. *Adv. Adaptive Data Anal.* **1**(2), 177–229 (2009)
29. N Khan, B Boashash, Instantaneous frequency estimation of multicomponent nonstationary signals using multiview time-frequency distributions based on the adaptive fractional spectrogram. *IEEE Signal Process. Lett.* **20**(2), 157–160 (2013)
30. D Vakman, On the analytic signal, the Teager-Kaiser energy algorithm, and other methods for defining amplitude and frequency. *IEEE Trans. Signal Process.* **44**(4), 791–797 (1996)
31. M Sebghati, H Amindavar, A novel analytical approach to orthogonal bases extraction from AM-FM signals. Paper presented at the IEEE international conference on acoustic, speech and signal processing (ICASSP) (Czech Republic, 22–27 May 2011), pp. 3820–3823
32. H D'Angelo, *Linear Time-Varying Systems, Analysis and Synthesis.* (Allyn and Bacon, Boston, 1970)
33. N Martin, An AR spectral analysis of non-stationary signals. *Signal Process.* **10**, 61–74 (1986)
34. J Hopgood, P Rayner, Blind single channel deconvolution using nonstationary signal processing. *IEEE Trans. Speech Audio Process.* **11**(5), 476–488 (2003)
35. P Heuberger, P Van den Hof, B Wahlberg, *Modeling and Identification with Rational Orthogonal Basis Functions.* (Springer, London, 2005)
36. B Ninness, F Gustafsson, A unifying construction of orthonormal bases for system identification. *IEEE Trans. Automatic Control.* **42**(4), 515–521 (1997)

37. R O'Brien Jr., P Iglesias, On the poles and zeros of linear time-varying systems. *IEEE T Circuits-I*. **48**(5), 565–577 (2001)
38. R Duda, P Hart, D Stork, *Pattern Classification*, 2nd edn. (Wiley, New York, 2000)
39. Q Meng, M Yuan, J Zhao, Empirical AM-FM decomposition of auditory signals. *J. Acoust. Soc. Am.* **131**(4) (2012). doi:10.1121/1.4708214

doi:10.1186/1687-4722-2014-8

Cite this article as: Sebghati et al.: Basis approach to estimate the instantaneous frequencies in multicomponent AM-FM signals. *EURASIP Journal on Audio, Speech, and Music Processing* 2014 **2014**:8.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
