

Seismic data decomposition into spectral components using regularized nonstationary autoregression

Sergey Fomel*, The University of Texas at Austin, Austin, TX, USA
sergey.fomel@beg.utexas.edu

Summary

Seismic data can be decomposed into nonstationary spectral components with smoothly variable frequencies and smoothly variable amplitudes. To estimate local frequencies, I use a nonstationary version of Prony's spectral analysis method defined with the help of regularized nonstationary autoregression (RNAR). To estimate local amplitudes of different components, I fit their sum to the data using regularized nonstationary regression (RNR). Shaping regularization ensures stability of the estimation process and provides controls on smoothness. Potential applications of the proposed technique include noise attenuation, seismic data compression, and seismic data regularization.

Introduction

Decomposing data into components has an immediate application in noise-attenuation problems in cases where signal and noise correspond to different components. The classic Fourier transform, Radon transform (Gardner and Lu, 1991), wavelet transform (Mallat, 2009), curvelet frame (Herrmann and Hennenfent, 2008), and seislet transform and frame (Fomel and Liu, 2010) are some examples of possible decompositions applicable to seismic data. A fundamental characteristic of seismic data is non-stationarity. In 1-D (time dimension), seismic data are nonstationary because of wave-attenuation effects. In 2-D and 3-D (time and space dimensions), non-stationarity is manifested by variable slopes of seismic events. The nonstationary character of seismic data can be captured by EMD (empirical mode decomposition) proposed by Huang et al. (1998). EMD has a number of important applications in seismic data analysis (Magrin-Chagnolleau and Baraniuk, 1999; Battista et al., 2007; Bekara and van der Baan, 2009). However, it remains "empirical" because its properties are not fully understood. Daubechies et al. (2011) recently proposed an EMD-like decomposition using the continuous wavelet transform and synchrosqueezing (Daubechies and Maes, 1996). Synchrosqueezing improves the analysis but remains an indirect method when it comes to extracting spectral attributes.

In this paper, I develop an efficient decomposition algorithm, which explicitly fits seismic data to a sum of oscillatory signals with smoothly varying frequencies and smoothly varying amplitudes. Such decomposition is close in properties to the one produced by EMD but with more explicit controls on the frequencies and amplitudes of different components and on their smoothness. The main tool is regularized nonstationary regression, or RNR (Fomel, 2009), a general method for fitting data to a set of basis functions with nonstationary coefficients. RNR was previously applied to time-frequency decomposition over a set of regularly sampled frequencies (Liu and Fomel, 2012). When the input signal is fitted to shifted versions of itself, RNR turns into regularized nonstationary autoregression or RNAR, which is related to adaptive prediction-error filtering. RNAR had previous applications in data regularization (Liu and Fomel, 2011) and noise removal (Liu et al., 2012). In this paper, I use it for spectral analysis and estimating different frequencies present in the data by a nonstationary extension of Prony's method of autoregressive spectral analysis (Marple, 1987; Bath, 1995). After the frequencies have been identified, I use RNR to determine local, smoothly varying amplitudes of different components.

Regularized Nonstationary Regression

Regularized nonstationary regression (Fomel, 2009) is based on the following simple model. Let $d(\mathbf{x})$ represent the data as a function of data coordinates \mathbf{x} , and $b_n(\mathbf{x})$, $n = 1, 2, \dots, N$ represent a collection of basis functions. The goal of *stationary* regression is to estimate coefficients a_n , $n = 1, 2, \dots, N$ such that the prediction error

$$e(\mathbf{x}) = d(\mathbf{x}) - \sum_{n=1}^N a_n b_n(\mathbf{x}) \quad (1)$$

is minimized in the least-squares sense. In the case of regularized *nonstationary* regression (RNR), the coefficients become variable,

$$e(\mathbf{x}) = d(\mathbf{x}) - \sum_{n=1}^N a_n(\mathbf{x}) b_n(\mathbf{x}), \quad (2)$$

and their resolution and variability is controlled by shaping regularization (Fomel, 2007). Shaping regularization applied to RNR amounts to linear inversion, $\mathbf{a} = \mathbf{M}^{-1} \mathbf{c}$, where \mathbf{a} is a vector composed of $a_n(\mathbf{x})$, the elements of vector \mathbf{c} are

$$c_i(\mathbf{x}) = \mathbf{S} \mathbf{b}_i^*(\mathbf{x}) d(\mathbf{x}), \quad (3)$$

the elements of matrix \mathbf{M} are

$$M_{ij}(\mathbf{x}) = \lambda^2 d_{ij} + \mathbf{S} \mathbf{b}_i^*(\mathbf{x}) b_j(\mathbf{x}) - \lambda^2 d_{ij} \quad (4)$$

λ is a scaling coefficient, and \mathbf{S} represents a shaping (usually smoothing) operator. When inversion is implemented by an iterative method, such as conjugate gradients, we find that strong smoothing makes \mathbf{M} close to identity and easier (taking less iterations) to invert, whereas weaker smoothing slows down the inversion but allows for more details in the solution. This intuitively logical behavior distinguishes shaping regularization from alternative methods (Fomel, 2009).

Regularized nonstationary autoregression (RNAR) corresponds to the case of basis functions being causal translations of the data itself. In 1-D, this condition implies $b_n(t) = d(t - n\Delta t)$.

Autoregressive spectral analysis

Prony's method of data analysis was developed originally for representing a noiseless signal as a sum of exponential components (Prony, 1795). It was extended later to noisy signals, complex exponentials, and spectral analysis (Pisarenko, 1973; Marple, 1987; Bath, 1995). The basic idea follows from the fundamental property of exponential functions: $\exp[\alpha(t + \Delta t)] = \exp[\alpha t] \exp[\alpha \Delta t]$. In signal-processing terms, it implies that a time sequence $d(t) = A \exp[\alpha t]$ (with real or complex α) is predictable by a two-point prediction-error filter $(1, -\exp[\alpha \Delta t])$, or, in the Z-transform notation,

$$F_0(Z) = 1 - Z/Z_0, \quad (5)$$

where $Z_0 = \exp[-\alpha \Delta t]$. If the signal is composed of multiple exponentials,

$$d(t) \approx \sum_{n=1}^N A_n \exp(a_n t), \quad (6)$$

they can be predicted one by one by using a convolution of several two-point prediction-error filters:

$$F(Z) = (1 - Z/Z_1)(1 - Z/Z_2) \cdots (1 - Z/Z_N) = 1 + a_1 Z + a_2 Z^2 + \cdots + a_N Z^N, \quad (7)$$

where $Z_n = \exp[-\alpha_n \Delta t]$. This observation suggests the following three-step algorithm:

1. Estimate a prediction-error filter from the data by determining filter coefficients a_1, a_2, \dots, a_N from the least-squares minimization of

$$e(t) = d(t) - \sum_{n=1}^N a_n d(t - n\Delta t). \quad (8)$$

2. Writing the filter as a Z polynomial (equation 7), find its complex roots Z_1, Z_2, \dots, Z_N . The exponential factors $\alpha_1, \alpha_2, \dots, \alpha_N$ are determined then as $\alpha_n = -\log(Z_n) / \Delta t$.
3. Estimate amplitudes A_1, A_2, \dots, A_N of different components in equation 6 by linear least-squares fitting.

Prony's method can be applied in sliding windows, which was a technique developed by Russian geophysicists (Gritsenko et al., 2001) and applied to identifying low-frequency seismic anomalies (Mitrofanov et al., 1998). I propose to extend it to a smoothly nonstationary analysis with the following modifications:

1. Using RNAR, the filter coefficients a_n become smoothly-varying functions of time $a_n(t)$, which allows the filter to adapt to nonstationary changes in the data.
2. At each instance of time, roots of the corresponding Z polynomial also become functions of time $Z_n(t)$. I apply a robust, eigenvalue-based algorithm for root finding (Toh and Trefethen, 1994). The instantaneous frequency of different $f_n(t)$ is determined directly from the phase of different roots:

$$f_n(t) = -\text{Re} \left[\arg \left(\frac{Z_n(t)}{2\rho \Delta t} \right) \right]. \quad (9)$$

3. Finally, using RNR, I estimate smoothly-varying amplitudes of different components $A_n(t)$. The nonstationary decomposition model for a complex signal $d(t)$ is thus

$$d(t) \approx \sum_{n=1}^N \hat{a}_n d_n(t), \text{ where } d_n(t) = A_n(t) \exp[i f_n(t)] \quad (10)$$

and the local phase $\phi_n(t)$ corresponds to time integration of the instantaneous frequency determined in Step 2:

$$f_n(t) = 2\rho \int_0^t f_n(t) dt. \quad (11)$$

For ease of analysis, real signals can be transformed to the complex domain by using analytical traces (Taner et al., 1979).

Figure 1a shows a benchmark signal from Liu et al. (2011), which consists of two nonstationary components with smoothly varying (parabolic) frequencies. The corresponding time-frequency analysis over a range of regularly sampled frequencies is shown in Figure 1b. Two instantaneous frequencies were extracted at Step 2 of the algorithm from a time-varying, three-point prediction-error filter (Figure 1c). They correspond precisely to the two components present in the synthetic signal. Finally, Step 3 separates these components (Figure 2.)

The cost of the proposed decomposition is $O(N N_t N_{iter})$, where N_t is the number of time samples and N_{iter} is the number of shaping iterations (typically between 10 and 100). This is significantly faster than the $O(N_t^2 N_{iter})$ cost of time-frequency decomposition for a regularly sampled range of frequencies.

Although the examples of this paper use only 1-D analysis, the proposed technique is directly applicable to analyzing variable slopes of 2-D and 3-D seismic events, where the analysis applies to different frequency slices in the f - x domain (Canales, 1984; Spitz, 1999, 2000).

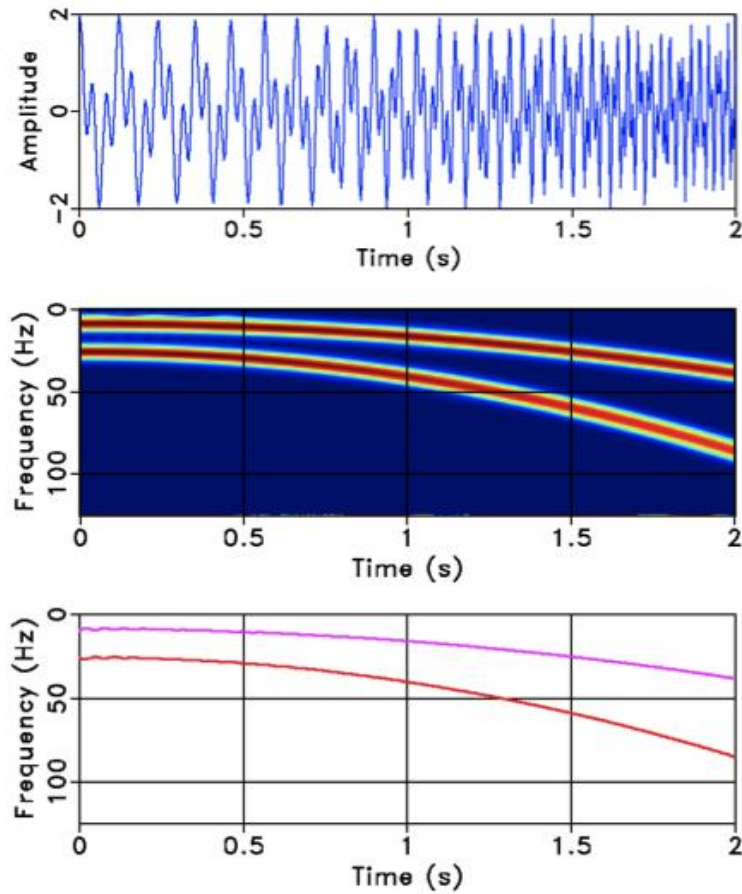


Figure 1: (a) Test signal composed of two variable-frequency components. (b) Time-frequency decomposition. (c) Instantaneous frequencies estimated by RNAR.

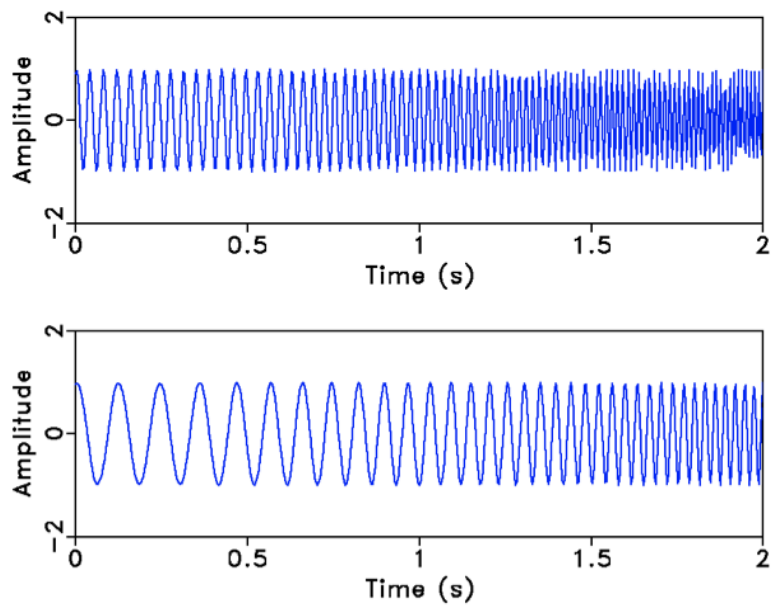


Figure 2: Decomposition of the signal from Figure 1a into two spectral components using RNR.

Example

To illustrate performance of the proposed approach, I use a 2-D section from a land seismic survey (Figure 7a), analyzed previously by Liu and Fomel (2013). I choose a three-point prediction-error filter to highlight the two most significant data components. The two estimated spectral components are shown in the bottom of Figure 7, and the residual is shown in Figure 7b. The instantaneous frequencies of the two components and their amplitudes are shown in Figure 8. Comparing frequency and amplitude attributes from different components, a low-frequency anomaly (a zone of attenuated high frequencies) in the top-left part of the section becomes apparent. This anomaly might indicate presence of gas (Castagna et al., 2003).

Conclusions

I have presented a constructive approach to decomposing seismic data into spectral components with smoothly variable frequencies and smoothly variable amplitudes. The output of the algorithm is close to that of empirical model decomposition (EMD) but with a more explicit control on parameters and more direct access to instantaneous-frequency and amplitude attributes. The main tool for the task is regularized nonstationary regression (RNR), which is applied twice: first to estimate local frequencies by autoregression (RNAR) and then to estimate local amplitudes. Although all examples shown in this paper use only 1-D analysis, the proposed technique is also applicable to analyzing 2-D and 3-D variable-slope seismic events in the f - x domain. Potential applications include noise attenuation, seismic data compression, and seismic data regularization.

Acknowledgements

I would like to thank Sergey Gritsenko for teaching me Prony's method more than 20 years ago. I am also grateful to Yihua Cai, Guochang Liu, Yang Liu, Karl Schleicher, and Lexing Ying for inspiring discussions.

This paper was previously presented at SEG 2012 Annual Meeting (Fomel, 2012). It is submitted to GeoConvention by invitation.

References

- Bath, M., 1995, *Modern Spectral Analysis with Geophysical Applications*: Soc. of Expl. Geophys. (Edited by Robert E. Wiley).
- Battista, B. M., C. Knapp, T. McGee, and V. Goebel, 2007, Application of the empirical mode decomposition and Hilbert-Huang transform to seismic reflection data: *Geophysics*, **72**, H29–H37.
- Bekara, M., and M. van der Baan, 2009, Random and coherent noise attenuation by empirical mode decomposition: *Geophysics*, **74**, V89–V98.
- Canales, L. L., 1984, Random noise reduction: 54th Ann. Internat. Mtg. Soc. of Expl. Geophys., Session:S10.1.
- Castagna, J. P., S. Sun, and R. W. Siegfried, 2003, Instantaneous spectral analysis: Detection of low-frequency shadows associated with hydrocarbons: *The Leading Edge*, **22**, 120–127.
- Daubechies, I., J. Lu, and H.-T. Wu, 2011, Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool: *Applied and Computational Harmonic Analysis*, **30**, 243–261.
- Daubechies, I., and S. Maes, 1996, A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models, *in Wavelets in medicine and biology*: CRC Press, 527–546.
- Fomel, S., 2007, Shaping regularization in geophysical-estimation problems: *Geophysics*, **72**, R29–R36.
- , 2009, Adaptive multiple subtraction using regularized nonstationary regression: *Geophysics*, **74**, V25–V33.
- , 2012, Seismic data decomposition into spectral components using regularized nonstationary autoregression: 82nd Annual International Meeting, SEG, <http://dx.doi.org/10.1190/segam2012-1416.1>

- Fomel, S., and Y. Liu, 2010, Seislet transform and seislet frame: *Geophysics*, **75**, V25–V38.
- Gardner, G. H. F., and L. Lu, eds., 1991, *Slant-stack processing*: Society of Exploration Geophysicists. Issue 14 of *Geophysics* reprint series.
- Gritsenko, S. A., S. Fomel, and V. S. Chernyak, 2001, Filtering using Prony's method: *Geofizika*, 24–26 (in Russian).
- Herrmann, F. J., and G. Hennenfent, 2008, Non-parametric seismic data recovery with curvelet frames: *Geophysical Journal International*, **173**, 233–248.
- Huang, N. E., Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, 1998, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis: *Proceeding of the Royal Society of London Series A*, **454**, 903–995.
- Liu, G., X. Chen, J. Du, and K. Wu, 2012, Random noise attenuation using f - x regularized nonstationary autoregression: *Geophysics*, **77**, V61–V69.
- Liu, G., S. Fomel, and X. Chen, 2011, Time-frequency analysis of seismic data using local attributes: *Geophysics*, **76**, P23–P34.
- Liu, Y., and S. Fomel, 2011, Seismic data interpolation beyond aliasing using regularized nonstationary autoregression: *Geophysics*, **76**, V69–V77.
- , 2013, Seismic data analysis using local time-frequency decomposition: *Geophysical Prospecting*, DOI: 10.1111/j.1365-2478.2012.01062.x.
- Magrin-Chagnolleau, I., and R. Baraniuk, 1999, Empirical mode decomposition based frequency attributes: 69th Ann. Internat. Mtg, Soc. of Expl. Geophys., 1949–1952.
- Mallat, S., 2009, *A wavelet tour of signal processing: The sparse way*: Academic Press.
- Marple, S. L., 1987, *Digital spectral analysis with applications*: Prentice-Hall.
- Mitrofanov, G., Z. Zhan, and J. Cai, 1998, Using of the Proni transform of Chinese seismic data: 68th Ann. Internat. Mtg, Soc. of Expl. Geophys., 1157–1159.
- Pisarenko, V. F., 1973, The retrieval of harmonics from a covariance function: *Journal of the Royal Astronomical Society*, **33**, 347–366.
- Prony, R., 1795, *Essai expérimental et analytique*: *Annuaire de l'École Polytechnique*, **1**, 24.
- Spitz, S., 1999, Pattern recognition, spatial predictability, and subtraction of multiple events: *The Leading Edge*, **18**, 55–58.
- , 2000, Model-based subtraction of multiple events in the frequency-space domain: 70th Ann. Internat. Mtg, Soc. of Expl. Geophys., 1969–1972.
- Taner, M. T., F. Koehler, and R. E. Sheriff, 1979, Complex seismic trace analysis: *Geophysics*, **44**, 1041–1063.
- Toh, K., and L. Trefethen, 1994, Pseudozeros of polynomials and pseudospectra of companion matrices: *Mathematika*, **68**, 403–425.

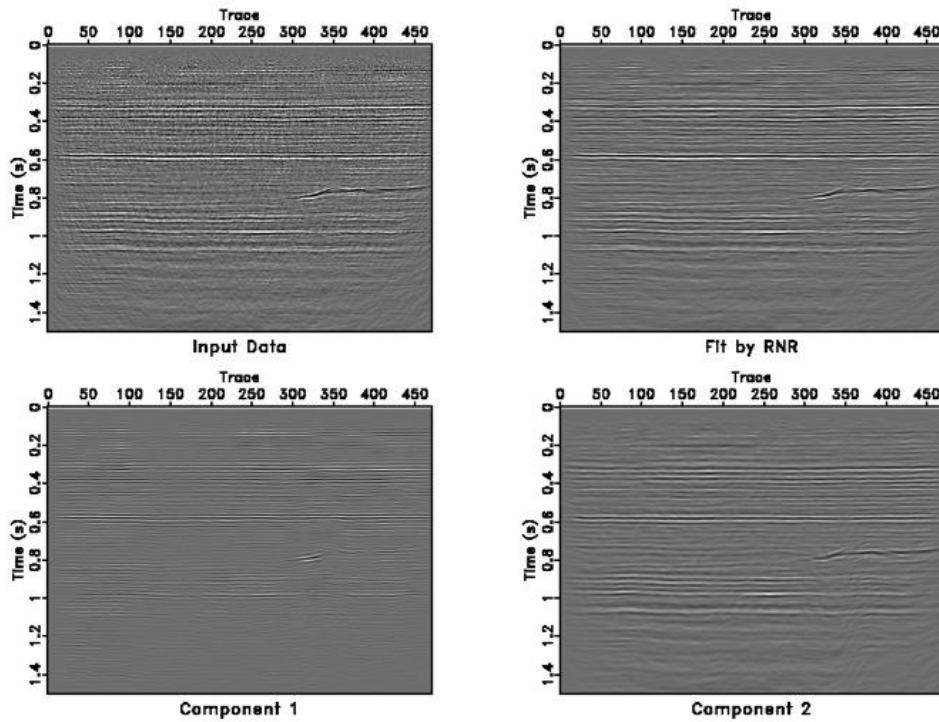


Figure 7: (a) 2-D seismic data section. (b) Residual after fitting data with two nonstationary spectral components shown in (c) and (d).

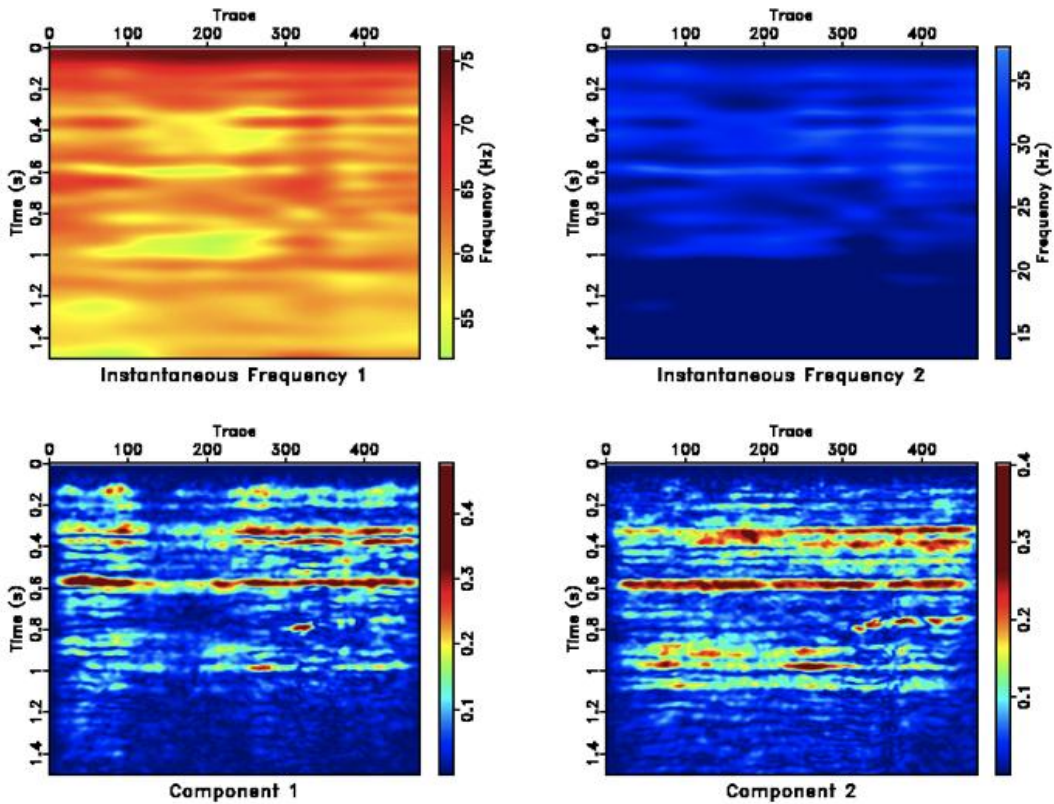


Figure 8: Top: instantaneous frequencies of high-frequency and low-frequency components from decomposition shown in Figure 7. Bottom: amplitudes of high-frequency and low-frequency components from decomposition shown in Figure 7.