Biol Cybern (2005) 92: 380–392
DOI 10.1007/s00422-005-0566-y

ORIGINAL PAPER

Biological Cybernetics

Hualou Liang · Steven L. Bressler · Elizabeth A. Buffalo
Robert Desimone · Pascal Fries

Empirical mode decomposition of field potentials from macaque V4 in visual spatial attention

Received: 12 November 2004 / Accepted: 18 March 2005 / Published online: 19 May 2005

Abstract Empirical mode decomposition (EMD) has recently been introduced as a local and fully data-driven technique for the analysis of non-stationary time-series. It allows the frequency and amplitude of a time-series to be evaluated with excellent time resolution. In this article we consider the application of EMD to the analysis of neuronal activity in visual cortical area V4 of a macaque monkey performing a visual spatial attention task. We show that, by virtue of EMD, field potentials can be resolved into a sum of intrinsic components with different degrees of oscillatory content. Low-frequency components in single-trial recordings contribute to the average visual evoked potential (AVEP), whereas high-frequency components do not, but are identified as gamma-band (30–90 Hz) oscillations. The magnitude of time-varying gamma activity is shown to be enhanced when the monkey attends to a visual stimulus as compared to when it is not attending to the same stimulus. Comparison with Fourier analysis shows that EMD may offer better temporal and frequency resolution. These results support the idea that the magnitude of gamma activity reflects the modulation of V4 neurons by visual spatial attention. EMD, coupled with instantaneous frequency analysis, is demonstrated to be a useful technique for the analysis of neurobiological time-series.

H. Liang (✉)
School of Health Information Sciences,
University of Texas Health
Science Center at Houston,
7000 Fannin, Suite 600 Houston, TX 77030, USA
Tel: +1-713-5003914
Fax: +1-713-5003929
E-mail: hualou.liang@uth.tmc.edu

S.L. Bressler
Center for Complex Systems and Brain Sciences,
Florida Atlantic University,
Boca Raton, FL 33431, USA

E.A. Buffalo · R. Desimone
Laboratory of Neuropsychology,
National Institute of Mental Health,
National Institutes of Health,
Bethesda, MD 20892, USA

P. Fries
F.C. Donders Centre for Cognitive Neuroimaging,
University of Nijmegen,
6525 EK Nijmegen, The Netherlands

R. Desimone
Mc Govern Institute for Brain Research
at Massachusetts Institute of Technology,
Cambridge, MA02139, USA

1 Introduction

Spectral analysis has gradually become an indispensable tool in nearly every field of engineering and science (Jenkins and Watts 1968; Bendat and Piersol 1986; Percival and Walden 1993). To date, the Fourier transform (Oppenheim and Schafer 1989) is perhaps the most commonly used method for spectral analysis. It provides a general method for estimating the global power-frequency distribution (i.e. power spectrum) of a given random process, assuming that the process is stationary. Many processes encountered in real-world situations, however, are non-stationary. In the case of neurobiological time-series data, Fourier analysis (and its derived techniques) are often insufficient because the underlying processes are clearly non-stationary. Yet the analysis of neurobiological time-series frequently requires time–frequency representations (spectrograms) that indicate how the power spectrum changes over time.

The short-time Fourier transform (STFT) provides a simple and intuitive method for attaining a spectrogram (Oppenheim and Schafer 1989). By sliding a window along the time axis and repeatedly calculating the Fourier transform, a time–frequency distribution can be obtained. The spectrogram is an adequate description, however, only if the time-series is stationary. Moreover, even if the data are stationary within a time window, the spectrogram has the additional problem of limited frequency resolution. To obtain precise time information, a narrow time window is required, but the Fourier transform in a short window gives rise to low-frequency resolution, which is inversely proportional to the length of the data window. Characterization of

time–frequency energy concentrations are restricted by the Heisenberg uncertainty principle (Mallat 1998). Therefore, with the STFT, one cannot obtain both time-localized and frequency-localized information with good resolution.

Wavelet analysis (Daubechies 1992; Mallat 1998) has become popular in the past decade as a method for time-frequency representation. The Wavelet transform (WT) of a function $x(t)$ is defined as:

$$w(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\psi^* \left(\frac{t-b}{a}\right) \mathrm{d}t \qquad (1)$$

where a is the scale dilation parameter, b is the translation parameter and the function $\psi(t)$ is the mother wavelet (with * representing the complex conjugate). The mother wavelet used in this paper is the complex-valued Morlet wavelet, which is a complex sine wave modulated by a Gaussian envelope: $\psi(t) = \exp(i\omega_0 t)\exp(-t^2/2)$. A large ω_0 gives good frequency resolution at the expense of poor time resolution. A compromise value of 6 is used in the paper. The scale dilation parameter can be converted into the frequency by $f = \omega_0/(2\pi a)$. Variation of the scale dilation parameter, a, changes the frequency range of the basis function. The variation of the scale dilation and translation parameters yields a family of wavelet basis functions. At each frequency, the translation parameter b, may take a range of values corresponding to the time position along the data sample. A complete wavelet spectrum is obtained by using all appropriate values of a and b.

Although the WT appears similar to the STFT, there are three basic differences: (1) the basis functions in the WT are not limited to the sinusoidal waves, but rather are chosen by the user to address various problems of time–frequency resolution; (2) in contrast to the STFT, which uses a single analysis window width, the WT principally uses short windows at high frequencies and longer windows at low frequencies; and (3) when the scale dilation factor, a, in the WT is changed, the duration and the bandwidth of the basis function are both simultaneously changed, yet the original shape of the mother wavelet is retained. Nonetheless, wavelet analysis is also limited by the fundamental uncertainty principle, in which both time and frequency cannot simultaneously be resolved with the same precision. Moreover, the analysis results depend on the choice of mother wavelet, which is arbitrary and may not be optimal for the time-series being analyzed.

Even with an optimized joint time-frequency localization (Gabor 1946), the tradeoff between time and frequency energy concentrations, as bounded by the uncertainty principle, is unavoidable. Neither the STFT nor the WT can simultaneously provide both good frequency and time resolution. The notion of the instantaneous frequency (Boashash 1992) has been proposed as a natural solution to the problem of representing the frequency of a time-series at each time instant without the knowledge of the values at other times. The Hilbert transform (Bendat and Piersol 1986; Oppenheim and Schafer 1989), among others (Potamianos and Maragos 1994), has been routinely performed to compute the instantaneous frequency. It may be possible, therefore, to obtain good time and frequency resolution by the Hilbert transform. There is a problem with the use of instantaneous frequency, however, in that it only provides one value at each time, whereas in reality, time-series data usually contain many intrinsic frequencies. Thus, separation of these many intrinsic frequencies in a time-series is required in order to obtain a meaningful, well-defined instantaneous frequency.

A common solution to this problem has been to perform bandpass filtering of the time-series, and then to use the Hilbert transform to extract the instantaneous frequency or amplitude for each passband of interest (Freeman 2004a,b). This solution may not be optimal, however, since the choice of passbands is typically arbitrary, and the resulting instantaneous frequencies or amplitudes may be difficult to interpret, particularly for a wide passband (see below). The recently proposed empirical mode decomposition (EMD) (Huang et al. 1998a) is a new method that decomposes the time-series into narrow-band components, namely intrinsic mode functions (IMFs), by empirically identifying the physical time scales intrinsic to the data. The peculiar property of each IMF having a single local frequency is particularly suitable to computing the instantaneous frequency. The EMD method, therefore, may be viewed as a pre-processing procedure that is necessary before the Hilbert transform can be used to obtain instantaneous frequencies. EMD is based on direct extraction of the signal energy associated with various intrinsic time scales. The technique adaptively decomposes non-stationary time-series into a set of intrinsic oscillatory modes. These components allow the calculation of a meaningful multi-component instantaneous frequency by virtue of the Hilbert transform. Thus, one can potentially localize events in both time and frequency, even in non-stationary time-series.

In this paper, we examine the use of EMD to study neuronal activity in visual cortical area V4 of a macaque monkey performing a visual spatial attention task (Fries et al. 2001). A brief report on part of this work appeared recently in an article by Liang et al. (2005a). This article is organized as follows. In Sect. 2, we present the basic methodology of EMD and the Hilbert transform, each with intuitive examples. The combined use of the EMD method with the Hilbert transform provides a general approach to non-stationary, nonlinear time-series analysis. In Sect. 3, we consider two simulations, where we know exactly the time-series composition, to verify the implementation and to demonstrate its advantages. The first example is designed such that the time-series components significantly overlap in both time and frequency, whereas the second is focused on a quantitative examination of the performance of the EMD method. In Sect. 3, we also explore the application of EMD to the analysis of field potential data from macaque cortical area V4 during visual spatial attention. Section 4 concludes the paper with discussion of some practical issues relating to EMD applications.

2 Methods

2.1 Empirical mode decomposition

Empirical mode decomposition is a general nonlinear, non-stationary time-series analysis method. The EMD method was initially proposed for the study of ocean waves (Huang et al. 1998a), and found immediate application in biomedical engineering (Huang et al. 1998b; Liang et al. 2000; Balocchi et al. 2004). The major advantage of EMD is that the basis functions are derived directly from the time-series itself. Hence the analysis is adaptive, in contrast to Fourier analysis, where the basis functions are linear combinations of fixed sine and cosine waves.

The central idea of EMD time-series analysis is a sifting process to decompose a time-series into a set of IMFs having well-defined instantaneous frequencies by empirically identifying the physical time scales intrinsic to the time-series. The instantaneous frequency is defined by the time lapse between successive extrema. A time-series must satisfy two criteria to be an IMF: (1) the number of extrema and the number of zero crossings are either equal or differ at most by one; and (2) the mean of its upper and lower envelopes equals zero. The first criterion is similar to a narrow-band requirement. The second criterion modifies a global requirement to a local one, and is necessary to ensure that the instantaneous frequency will not have unwanted fluctuations as induced by asymmetric waveforms. The time-series must have at least two extrema – one maximum and one minimum to be successfully decomposed into IMFs.

Given these two definitive requirements for an IMF, the sifting process for extracting IMFs from a given time-series $x(t)$ is described as follows:

1. Two smooth splines are constructed connecting all the maxima and minima of $x(t)$ to get its upper envelope, $x_{\text{up}}(t)$, and its lower envelope, $x_{\text{low}}(t)$; the extrema can be simply found by determining the change of sign of the derivative of the time-series. Once the extrema are identified, all the maxima are connected by a cubic spline line as the upper envelope. The procedure is repeated for the local minima to produce the lower envelope. All the data points should now be covered by the upper and lower envelopes.
2. The mean of the two envelopes is subtracted from the data to get their difference $d(t) = x(t) - (x_{\text{up}}(t) + x_{\text{low}}(t))/2$
3. The process is repeated for $d(t)$ until the resulting signal $c_1(t)$, the first IMF, satisfies the criteria of an intrinsic mode function.

The residue $r_1(t) = x(t) - c_1(t)$ is then treated as a new time-series subject to the sifting process as described above, yielding the second IMF from $r_1(t)$. The procedure continues until either the recovered IMF or the residual time-series are too small (in the sense of the integrals of their absolute values), or the residual time-series has no turning points. Once all of the wavelike IMFs have been extracted, the final residual component represents the DC component containing the overall trend of the time-series.

At the end of this process, the time-series $x(t)$ can be expressed as follows:

$$x(t) = \sum_{j=1}^{N} c_j(t) + r_N(t) \tag{2}$$

where N is the number of IMFs, and $r_N(t)$ denotes the final residue, which can be interpreted as the trend of the time-series. The $c_j(t)$ are nearly orthogonal to each other, and all have nearly zero means. Due to this iterative procedure, none of the sifted IMFs is derived in closed analytical form.

By the nature of the decomposition procedure, the technique decomposes a time-series into N fundamental components, each with a distinct time scale. More specifically, the first component has the smallest time scale which corresponds to the fastest time variation of data. As the decomposition process proceeds, the time scale increases, and hence, the mean frequency of the mode decreases. Since the decomposition is based on the local characteristic time scale of the time-series to yield adaptive basis, it is applicable to nonlinear and non-stationary data analysis.

A simple example is shown in Fig. 1 to illustrate the idea of EMD decomposition. The analyzed time-series (Fig. 1a) is composed of a linear chirp (a sinusoidal wave that increases in frequency linearly over time) and a nonlinear triangular waveform. Figure 1b plots the same time-series (thin solid line) with its upper and lower envelopes (dashed lines), as well as the mean of the envelopes (thick line). The final IMF components derived from the EMD are shown in Fig. 1c–e, where the first IMF (Fig. 1c) and the second IMF (Fig. 1d) are, respectively, identified as the linear chirp and the triangular waveform, and the last component (Fig. 1e) is the negligible residual. With the presence of the non-harmonic triangular waveform, any harmonic analysis method, such as the STFT or WT, would produce a much less compact and physically less meaningful decomposition (Rilling et al. 2003). This example, though simple, underscores the potentially "non-harmonic" nature of EMD.

In practice, the time-series remaining after a certain number of iterations does not carry significant physical information, because, if sifting is carried on to an extreme, it could result in a pure frequency modulated signal of constant amplitude. To avoid this, it is typical to stop the sifting process by limiting the standard deviation, computed from two consecutive sifting results, which is usually set between 0.2 and 0.3. By construction, the number of extrema decreases when going from one residual to the next, and the whole decomposition is guaranteed to be completed with a finite number of modes.

2.2 Hilbert transform

Once all the IMFs' are determined, the instantaneous frequency of each IMF at each time point can readily be obtained by the Hilbert transform (Bendat and Piersol 1986; Oppenheim and Schafer 1989). The Hilbert transform has been widely used to obtain the analytic signal associated with

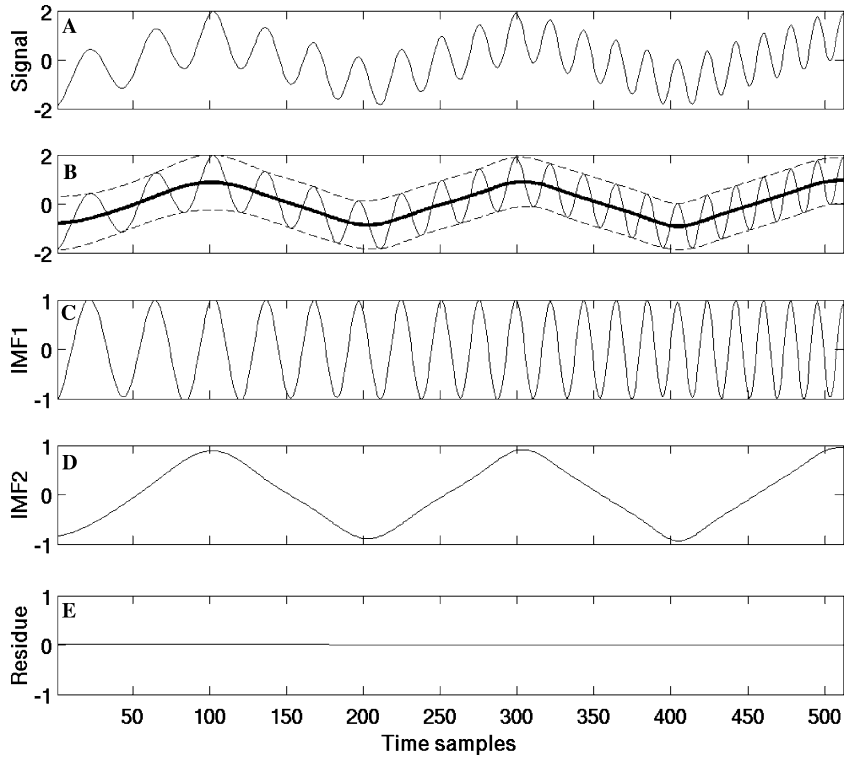

Fig. 1 A simple example of the empirical mode decomposition (EMD) of a two-component signal. **a** The analyzed signal consisting a linear chirp and a nonlinear triangular waveform. **b** Upper and lower envelopes (*dashed lines*) of the original signal and its point-by-point mean of the envelopes (*thick line*). **c–e** Three IMFs from the signal. The EMD ends up with essentially two IMFs, each corresponding to the linear chirp and the triangular waveform

a real signal $x(t)$, and consequently, the instantaneous envelope and phase functions. The instantaneous frequency that is of interest in our present report can be derived from the instantaneous phase.

Given an arbitrary time-series $x(t)$, the corresponding analytic signal is defined as:

$$z(t) = x(t) + iH[x(t)] = a(t)\exp[i\theta(t)]\,, \tag{3}$$

where $a(t)$ and $\theta(t)$ are the instantaneous amplitude and phase of the analytic signal $z(t)$, and the imaginary part $H[x(t)]$ is the Hilbert transform of $x(t)$:

$$H[x(t)] = \frac{1}{\pi} P\left[\int_{-\infty}^{\infty} \frac{x(u)}{t-u}\mathrm{d}u\right] \tag{4}$$

where the notation P indicates the Cauchy principal value of the integral (Oppenheim and Schafer 1989; Huang et al. 1998a).

The instantaneous phase can thus be derived from the analytic signal, which still retains the information content of the original real signal. The instantaneous frequency can then be obtained from the instantaneous phase as:

$$\omega(t) = \frac{d\theta(t)}{dt} \tag{5}$$

To summarize this process, given an IMF $c_j(t)$ we first compute its Hilbert transform $\boldsymbol{H}[\boldsymbol{c_j(t)}]$, and then find its phase through the combination of $c_j(t)$ and $\boldsymbol{H}[\boldsymbol{c_j(t)}]$. The instantaneous frequency of the IMF is finally obtained as the derivative of the instantaneous phase with respect to time. The Hilbert transform is applied to the IMFs produced by EMD because if it were applied to an arbitrary wide-band time series, it could produce negative frequencies, which bear no relationship to the real oscillations in the time-series (Huang et al. 1998a). Direct application of the Hilbert transform to any given time-series is thus of little practical value. To obtain meaningful and well-behaved instantaneous frequencies, the time-series to be analyzed must have no riding waves, and must be locally symmetrical about its mean as defined by the envelopes of local extrema. Such time-series are not common in real world data. With the advent of the EMD, it becomes possible to take full advantage of the Hilbert transform.

To appreciate the advantage of the Hilbert transform, we designed a simple simulation where a sine wave with abrupt frequency shifts around its midpoints was generated (Fig. 2a). As this already met the criteria for being an IMF, the Hilbert transform was directly applied to the signal. The resulting graph (Fig. 2b) depicts the transition points and distinguishes between the two frequencies with precision. As a comparison, the results from the STFT (spectrogram, Fig. 2c) and the WT (scalogram, Fig. 2d) of the same data both show poor frequency and time localizations for the frequency changes. Note that some frequency oscillations in the Hilbert spectrum are due to the discontinuity in the data.

In summary, the EMD allows time-series to be represented by intrinsic mode functions, to which the Hilbert transform can then be applied. The resulting Hilbert spectrum

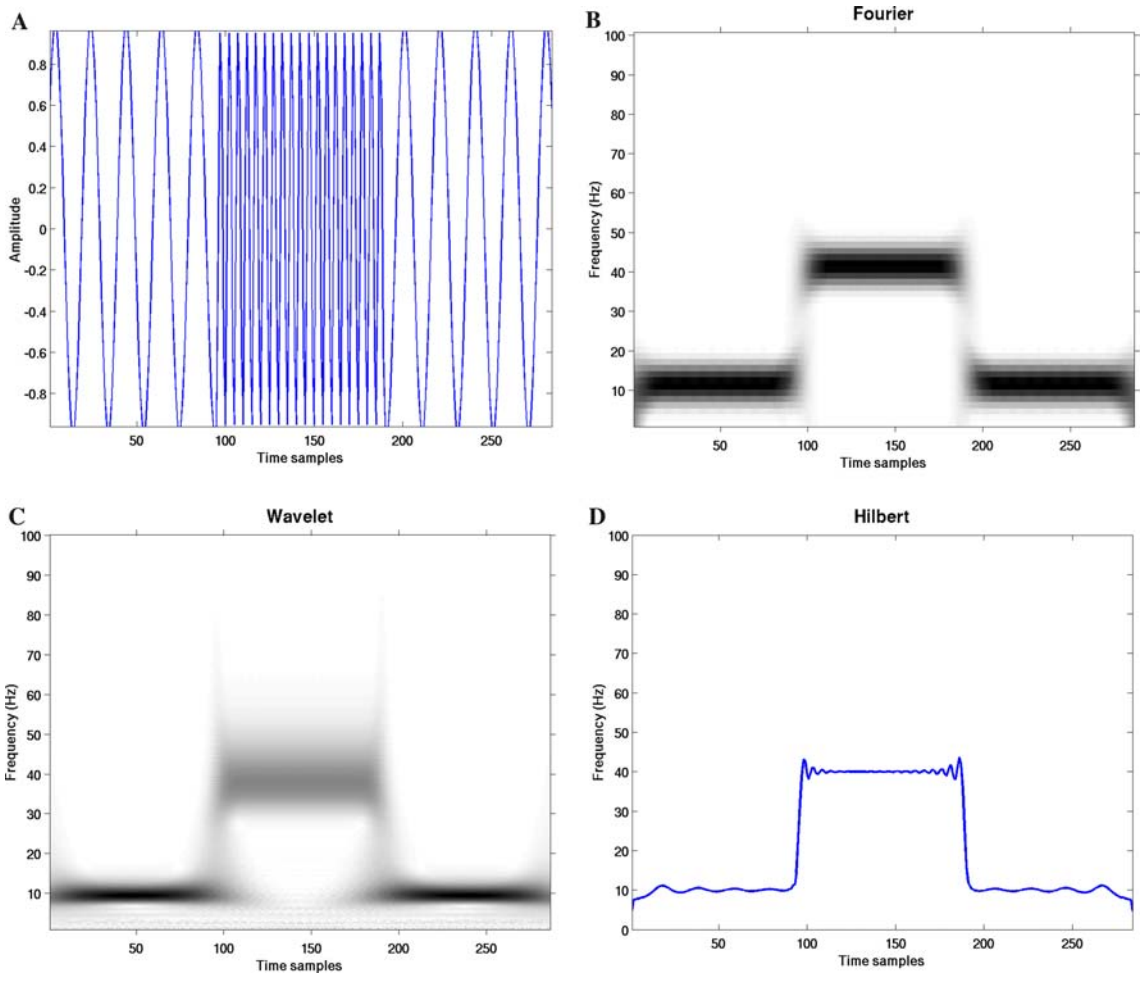

Fig. 2 **a** A sine wave with abrupt changes of frequency around its midpoints. **b–d** Comparison between the short-time Fourier transform **b** the Morlet wavelet transform **c** and the Hilbert transform **d** The Hilbert spectrum reveals much sharper frequency and time localization compared to both the short-time Fourier transform and the Morlet wavelet transform. Note that some ripples in the Hilbert spectrum are due to the discontinuity in the data

enables us to represent the amplitude and the instantaneous frequency as functions of time in a three-dimensional plot. The combination of the EMD method with the Hilbert transform offers a potentially powerful analysis technique, namely the Hilbert–Huang transform (HHT) (Huang et al. 1998a). The HHT method provides not only a more precise definition of particular events in time-frequency space than Fourier-based analysis, but also more physically meaningful interpretations of the underlying dynamic processes.

3 Results

3.1 Simulations

A further series of computer simulations was conducted to verify our implementation of the method. The first simulation demonstrated that the EMD functions can act as adaptive time-variant filters. The analyzed time-series (Fig. 3a, top graph) consisted of two frequency-modulated components, each a sinusoidal wave that initially increased and then decreased in frequency linearly over time, but with the components significantly overlapping in both time and frequency as evident in the spectrogram (Fig. 3b, left). The EMD decomposition of the time-series gave eight IMF components (Fig. 3a). The first two IMFs were clearly identified as the two frequency-modulated components contained in the original composite time-series, as confirmed by their spectrograms shown in Fig. 3b (middle and right). In no way can these two components be separated by pre-defined sub-band filtering.

Our second simulation attempted to quantitatively examine the performance of the EMD method. We created a composite time-series consisting of four components, each amplitude-modulated and frequency-modulated in the gamma (30–90 Hz), beta (13–30 Hz), alpha (8–12 Hz) and theta (4–7 Hz) frequency bands, to represent typical frequency components found in field potential data. Specifically, each component was generated by the following rule:

$$s_i(t) = [1 + a_i(t)] \sin[2\pi f_i t + \theta_i(t)], \quad i = 1, \ldots, 4 \qquad (6)$$

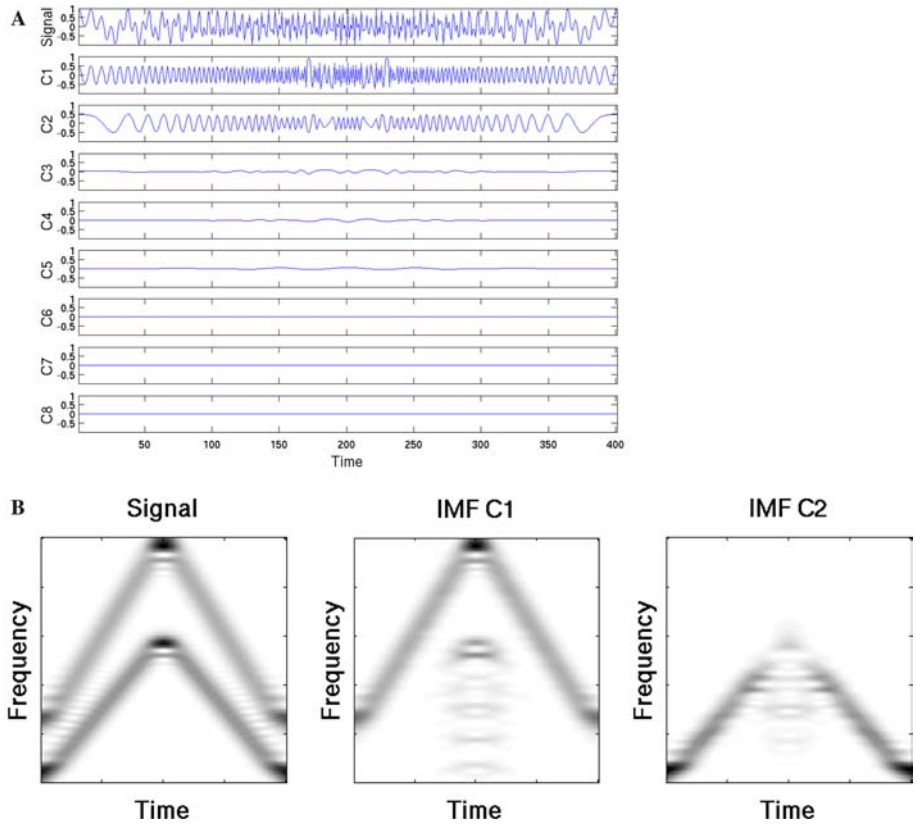

Fig. 3 EMD of a two-component signal. **a** The analyzed signal (*top panel*) and its eight IMFs (the remaining plots). The composite consists of two frequency-modulated components, each being sinusoidal wave that initially increases and then decreases in frequency linearly over time, but with both significantly overlapping in both time and frequency. **b** The spectrograms, from the left to the right, of the composite signal and of the first two IMFs from the EMD. The time–frequency analysis reveals that two components in the composite signal, overlapped in both time and frequency, are effectively separated and identified as the first two IMFs by the EMD

where f_i were the carrier frequencies, and $a_i(t)$ and $\theta_i(t)$ were obtained from white noise low-pass filtered at 0.05 Hz and 0.005 Hz, respectively. Each component thus consisted of a noisy, narrow-band oscillation around a central frequency at 40, 20, 10 and 5 Hz, respectively, and modulated in both amplitude and frequency.

The composite time-series and its EMD decomposition are represented in Fig. 4a. The time–frequency representations of the composite time-series based on the STFT, the WT, and the HHT are shown in Fig. 4b successively from the left to the right. While they all show similar energy–frequency distributions, both the STFT and the WT spectra suffer from the uncertainty principle and fail to reveal the frequency variation. The HHT, on the other hand, not only shows the frequency modulation but also gives a much sharper definition of the energy.

To quantify the performance of the EMD method, we used the root-mean-square error (RMSE) as the merit measure:

$$\mathrm{RMSE} = \sqrt{\sum_t \frac{(\hat{x}(t) - x(t))^2}{T}} \tag{7}$$

where $\hat{x}(t)$ is the decomposed IMF component, $x(t)$ is the original time-series, and T is the total length of the time-series. Fig. 4c shows the derived first four IMFs from the EMD (solid lines) correspondingly overlain on the original time-series components (dashed lines). The computed IMFs are seen to correspond well to the original time-series components, a correspondence that is confirmed by the small RMSE values (the top row in Table 1). As a comparison, the RMSE values were also computed for individual signal components reconstructed from the WT. We see from Table 1 that the EMD clearly gives better performance.

To illustrate the excellent performance of HHT for tracking the changes in frequency, we take the 40 Hz component as an example to compare its known modulated frequency and its instantaneous frequencies derived from the HHT and the WT. The results are shown in Fig. 4d. Quantitative comparison by computing RMSE revealed that the HHT indeed has smaller error measure (0.0264) than that of WT (0.0449).

3.2 Applications of EMD to V4 cortical field potentials

To our knowledge, previous applications of EMD have primarily involved single time-series records. In neurobiology, it is common to collect hundreds, even thousands, of repeated trials of subjects performing the same task. To make the

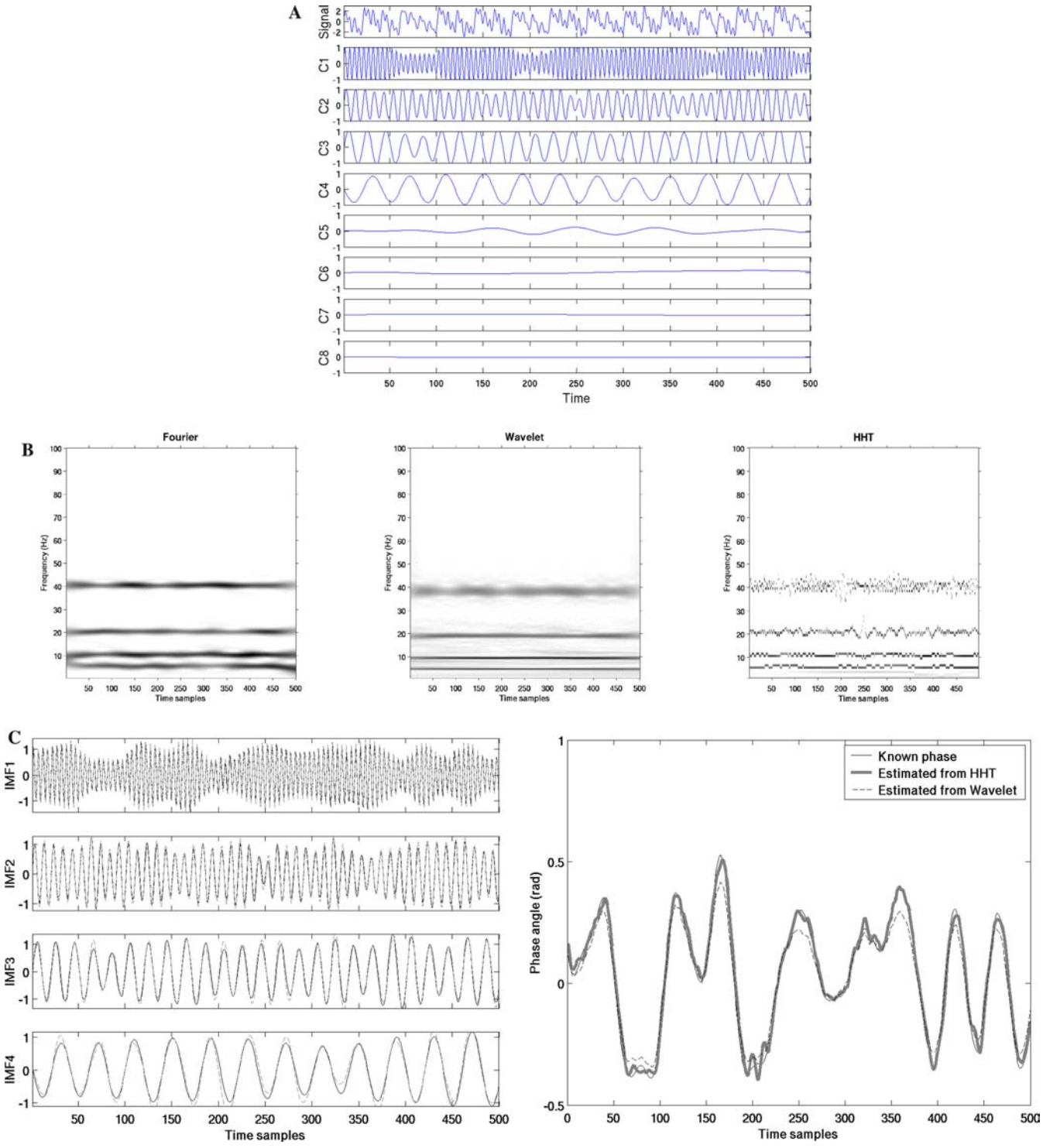

Fig. 4 EMD of a four-component signal. **a** The analyzed signal (*the very top panel*) is composed of four narrow-band oscillations around central frequencies, respectively, at 40, 20, 10 and 5 Hz, and each modulated in both amplitude and frequency (see text for details). Also shown are eight IMF components derived from the EMD decomposition, in which the first four IMFs are identified as the original signal components. **b** Comparison between the short-time Fourier transform, the Morlet wavelet transform, and the HHT of the composite signal (*from the left to the right*). Both the frequency and amplitude variations are clearly visible in the Hilbert spectrum. **c** The first four IMFs from the EMD (*solid lines*) and the original signal components (*dashed lines*). Quantitative comparisons reveal that both reach excellent agreement. **d** The known modulated frequency of the 40 Hz component (*thin line*) and its instantaneous frequencies derived from HHT (*fat line*) and WT (*dashed line*)

Table 1 Comparison of RMSE values between HHT and WT for four individual components centered at 40, 20, 10 and 5 Hz

	40 Hz	20 Hz	10 Hz	5 Hz
EMD	0.0701	0.1031	0.1301	0.1487
Wavelet	0.1167	0.1955	0.2136	0.2771

EMD technique directly applicable to such neurobiological data, we have developed strategies for analyzing multiple-trial recordings. Specifically, we introduce here two working procedures to illustrate the application of EMD in the analysis of cortical field potentials. The first application deals with the automatic identification of high-frequency components in the EMD, whereas the second application focuses on the automatic extraction of low-frequency components in single-trial recordings that contributed to the average visual evoked potential (AVEP).

Field potential data from visual cortical area V4 of a macaque monkey performing a visual spatial attention task were used to illustrate the usefulness of the EMD approach. Field potentials were simultaneously recorded from multiple V4 sites with overlapping receptive fields (RFs). The monkey fixated a central spot, and after a short delay, two stimuli were presented at equal eccentricity, one inside and one outside the RFs. On separate trials, the monkey was required to attend to the stimulus at one (target) location, and was rewarded for responding when the target changed color, ignoring changes at the other (distracter) location. Target and distracter color changes were equiprobable and uniformly distributed between 0.5 s and 5 s after stimulus onset. The result was two attention conditions: attention inside the RF vs. attention outside the RF. The analysis described here used field potentials from one V4 site on 300 trials correctly performed by a monkey whose attention was directed within the RF of that site, and another 300 trials with attention directed outside the RF.

A typical example of a single-trial field potential recording from area V4 in the macaque is shown in Fig. 5a, together with its IMFs derived from the EMD, and the instantaneous frequencies of IMF components. We can see from Fig. 5 that: first, strong gamma-band oscillations are observed to dominate the highest frequency (C1) component; and second, the instantaneous frequencies reveal clear frequency variation of each component as a function of time, reflecting the fact that the data are not stationary. Similar results were observed for other trials although there was variation from trial to trial in the number of components produced by EMD.

For the multiple-trial recordings, the number of IMF components obtained from individual trials was not the same, and hence, averaging the IMF components across trials was not possible. Two possible solutions were conceived. The first was to force the Hilbert spectra from individual trials to have the same number of bins in time–frequency space and then average them. Specifically, with the completion of the EMD decomposition, the Hilbert transform was performed for individual trials on all IMF components. Their corresponding Hilbert spectra were then derived and averaged across all the trials for the condition of attention inside the RF. An illustrative example of the resulting mean Hilbert spectrum from one V4 site is presented in Fig. 6a. Strong gamma activity, increasing over time, is clearly seen in the frequency range from 35 Hz to 70 Hz. A similar Hilbert spectrum was observed for attention outside the RF condition, but with a lower level of gamma power. Direct comparison between conditions is observed by contrasting either the spectrum at 2500 ms post-stimulus (Fig. 6c) or power at 50 Hz as a function of time for the two conditions (Fig. 6d). The statistical significance, indicated by the error bars, was assessed by the bootstrap resampling procedure (Efron and Tibshirani 1993). Gamma power when the monkey directed attention within the RF was significantly greater than that for attention outside the RF. This finding further confirms and extends previous findings (Fries et al. 2001) by revealing the detailed time course of task-related changes in gamma power.

For comparison, the STFT was performed on the same data, and the resulting average spectrogram is shown in Fig. 6b. Both the average Hilbert spectrum and the average spectrogram show general agreement about the concentration of gamma-band energy in time and frequency. However, the Hilbert spectrum gives a sharper and more refined definition of the energy contour, whereas the spectrogram spreads energy over a much wider frequency range. It is evident from Fig. 6a that the Hilbert spectrum clearly depicts fluctuations of the gamma frequencies (35–70 Hz) over time.

An alternative to the averaged Hilbert spectrum is to select IMF component with a particular frequency of interest from individual trials. The procedure is as follows: (1) EMD is first performed on individual trials; (2) the decomposed IMFs from each trial are then subject to the Fourier transform; (3) the selection of the IMF component pertinent to the frequency of interest (e.g. gamma activity) is obtained by identifying a dominant peak in the specific frequency range (e.g. the gamma band); and (4) averaging the selected IMF components from individual trials is finally achieved. In this procedure, the residual component is not subjected to the Fourier transform as it represents the mean or the trend in the data. By following this procedure to extract the high-frequency components, that is, the gamma activity, from individual trails, similar results to that shown in Fig. 6d were obtained.

The next application involved the selection of low-frequency IMF components in single-trial recordings in order to identify components corresponding to the AVEP. In the case of fine-to-coarse partial reconstruction, the estimated signal $\hat{x}(t)$ can be specifically written as:

$$\hat{x}(t) = \sum_{j=1}^{K} c_j(t) \tag{8}$$

where $K < N$, N is the total number of IMFs in the data. Considering IMFs statistics that each $c_j(t)$ has local zero mean of the signal, we designed a three-step procedure (Liang et al. 2005b) to identify the slow varying trend in the data: (1) the evolution of the mean of $\hat{x}(t)$ as a function of K is computed; (2) one sample t-test is employed to determine when the mean significantly departs from zero; and (3) once K is

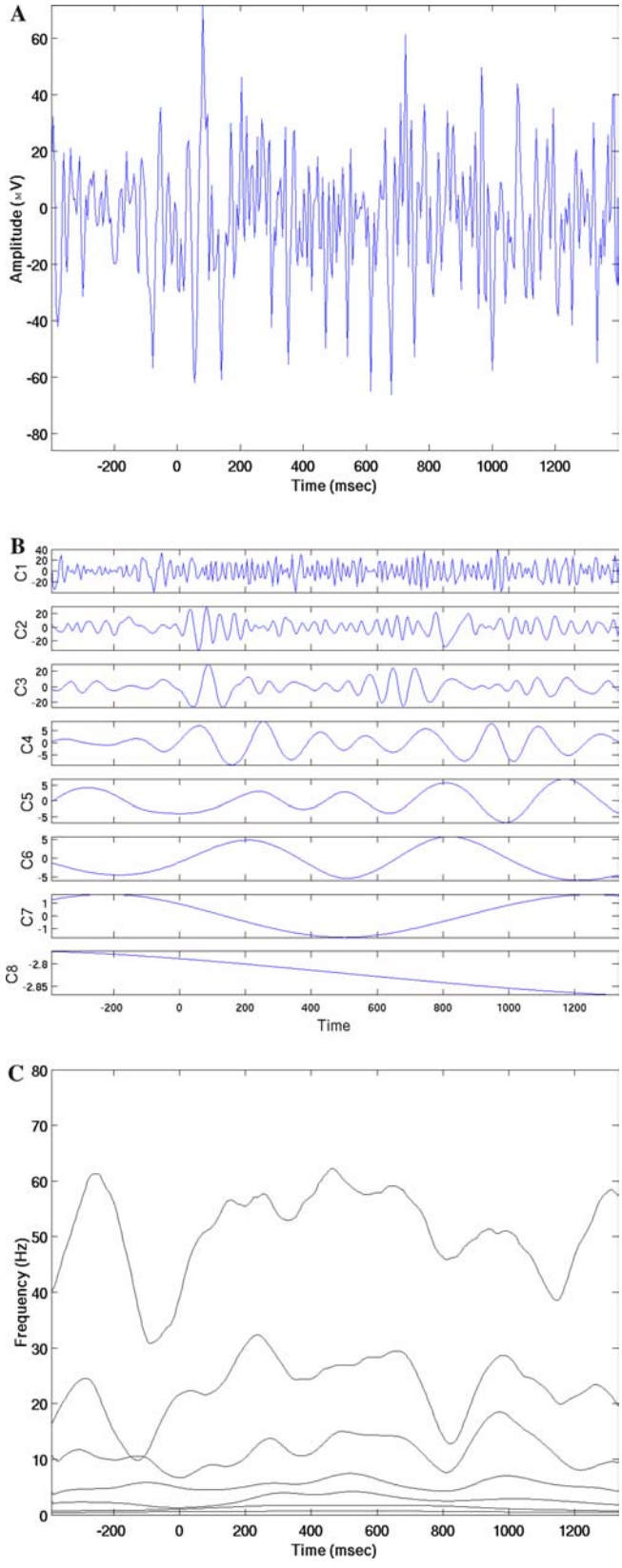

Fig. 5 A typical trial of field potential recording from area V4 **a**, its eight intrinsic mode functions IMFs, **b** and their smoothed instantaneous frequencies as functions of time **c**. Time 0 indicates the stimulus onset. Note that the C1 component is dominant in the gamma-frequency range. The large variation of instantaneous frequencies of IMF components indicates that the data are not stationary. The highest numbered components (*lowest frequency*) are equivalent to the trends in the data, suggesting that another benefit of EMD may be trend removal

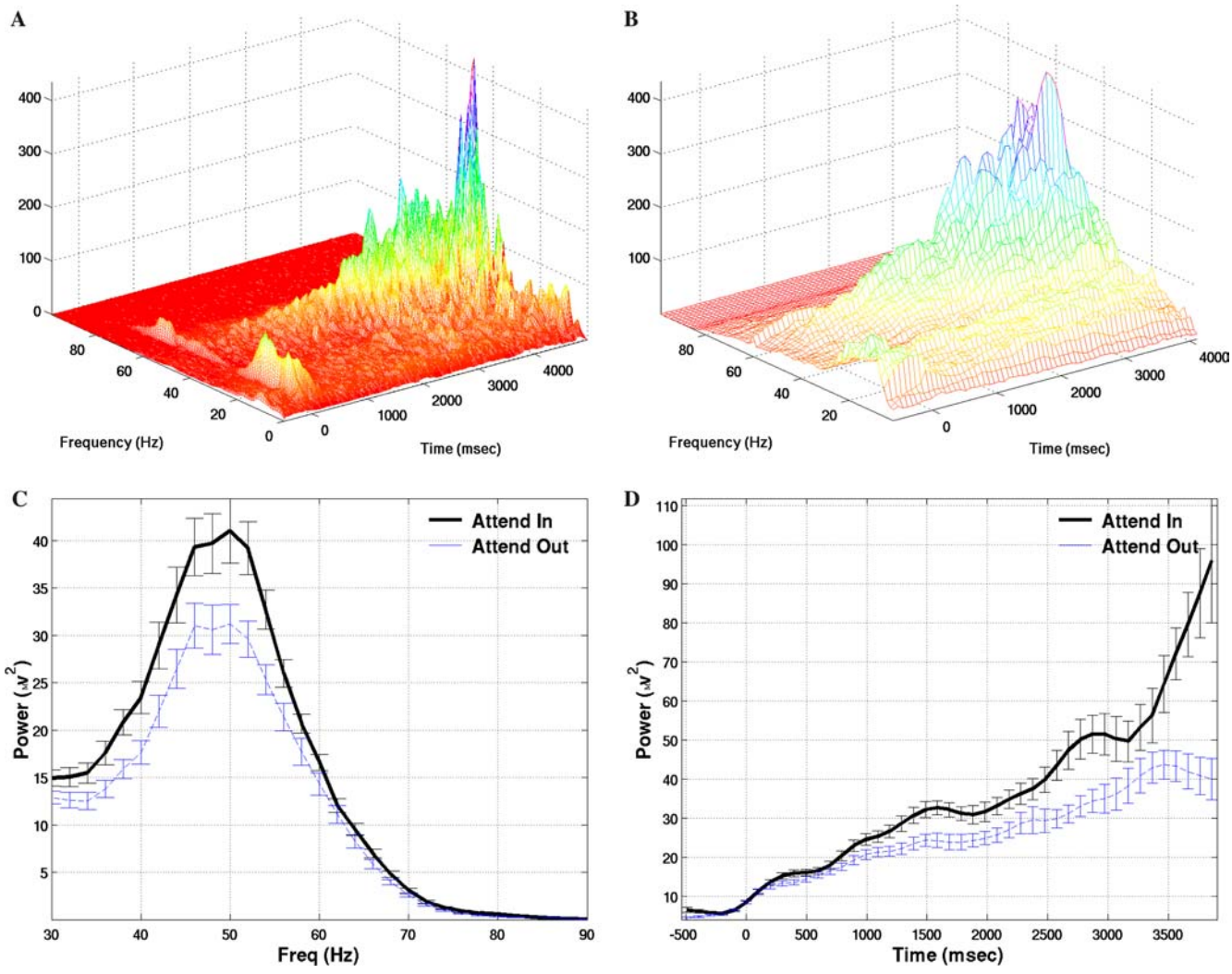

Fig. 6 a, b Time–frequency Hilbert spectrum **a** and the short-time Fourier spectrum **b** for V4 recording site 1 for attention inside the receptive field. **c** Hilbert spectra of site 1 for the two conditions, attention inside the receptive field (*fat lines*) and attention outside the receptive field (*thin lines*) at 2500 ms after stimulus onset. Note that both conditions have peaks about 50 Hz, but the peak is larger for the "inside" condition. **d** Hilbert spectra at 50 Hz as a function of time for the two conditions. Two prominent patterns are evident: (1) elevated gamma power when the monkey directed attention within the RF compared to attention outside the RF, a result further confirming and extending previous findings (Fries et al. 2001) by revealing the detailed time course of task-related changes in gamma power; and (2) "climbing" gamma activity in the temporal profile of which the slope is surprisingly linear. Error bars were obtained by the bootstrap resampling method

identified as a significant change point, partial reconstruction with IMFs from K+1 up to the residual reveals the slow varying trend in the data. The procedure works efficiently for data containing a slow varying trend. Otherwise, an ad hoc procedure can be simply adopted by selecting the last few low-frequency components. An example of this procedure is given in Fig. 7a, where a single-trial field potential recording in Fig. 5a is considered. We see that $K = 3$ is the change point where the mean is significantly ($p < 0.01$ at the 99% confidence level) different from the zero-mean null hypothesis. The low-frequency components are, therefore, estimated from IMF = 4 up to the residual (the last component). Adding up these low-frequency components gives the slow varying trend of the data, which is shown in Fig. 7b (thick line), also superimposed on the raw single-trial field potential recording (thin line).

Following the above procedure, we selected low-frequency components for individual trials. By averaging these low-frequency components over all trials, we obtained the AVEP, as shown in Fig. 7c. Computing the AVEP by EMD (Fig. 7c, thick lines) offers a striking contrast to that obtained by directly averaging the single-trial data (Fig. 7c, thin lines). The difference between these two approaches becomes less pronounced as more trials are averaged. By verifying that a realistic average can be computed from single-trial EMD components, it is suggested that the analysis of these single-trial components may also prove useful in some applications.

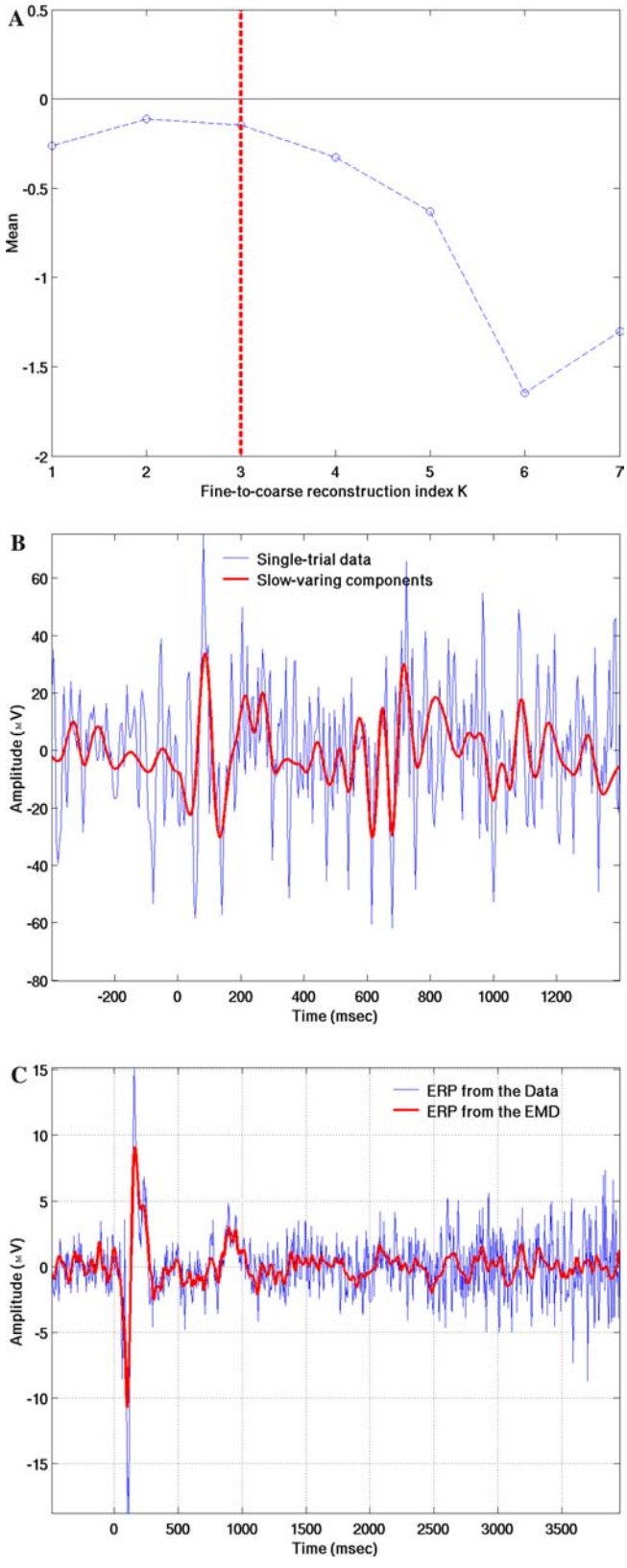

Fig. 7 **a** The mean of the fine-to-coarse EMD reconstruction as a function of index K. The *vertical dash-line* at $K = 3$ indicating that the mean departs significantly from zero ($p < 0.01$ at the 99% confidence level). **b** A single-trial field potential recording (*thin line*) and the estimated slow varying trend partially reconstructed from IMFs 4 to 7 (*fat line*). **c** AVEPs obtained from the EMD method (*fat line*) by selecting the low-frequency IMFs and that obtained directly from data by direct ensemble averaging (*thin line*). We see that there is a close match between these two approaches. The single-trial low-frequency components that went into the average might also prove useful for applications requiring single-trial analysis

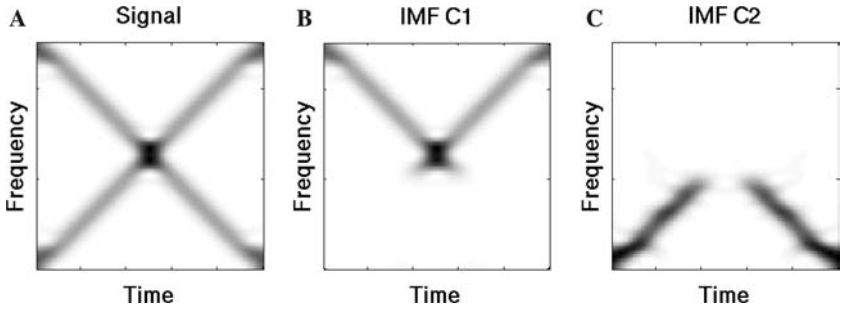

Fig. 8 An example of EMD of a two-component signal. **a** The spectrogram of the composite signal that consists of two chirp components with opposite chirp rates and their time–frequency representations cross each other. **b, c** The spectrograms of the first two IMFs derived from the EMD, indicating the IMFs do not retain the original signal components. This example stresses that care must be taken in the physical interpretation of the decomposed IMF components

4 Discussion

In this paper, application of the HHT method to the analysis of neurobiological time-series has been presented. Extensive computer simulations have been conducted to validate the performance of the HHT method. Application of the HHT to field potentials from cortical area V4 showed that field potentials were resolved into sets of intrinsic components having different degrees of oscillatory content. The high-frequency components were identified as gamma-band (30–90 Hz) oscillations, whereas the low-frequency components were the major contributions to the AVEP. We also showed that the magnitude of time-varying gamma activity was enhanced when the monkey attended to a visual stimulus as compared to when it was not attending to the same stimulus. Comparison with Fourier analysis showed that the HHT may offer better temporal and frequency resolution. These results support the idea that the magnitude of gamma activity reflects the modulation of V4 neurons by visual spatial attention (Desimone and Duncan 1995). The EMD, coupled with instantaneous frequency analysis, may prove to be a valuable technique for the analysis of neural data.

A potential difficulty in the application of HHT to the analysis of neurobiological time-series lies in how to characterize the spectra for an entire ensemble of trials collected under the same experimental condition. In this contribution, we have developed strategies to deal with multiple-trial recordings. In the selection of the IMF component with the frequency of interest from individual trials, the Fourier transform is essentially used to facilitate the identification. As an alternative, the Hilbert transform can also be used, where we first derive the instantaneous frequency from each IMF, and then identify the main frequency by computing the mode frequency of each IMF (i.e. the value of instantaneous frequency occurring most frequently in the IMF). In either case, it is required to have a priori information related to a particular hypothesis that is to be confirmed (e.g. activity is in the gamma-frequency range). As an initial data-screening technique, it is suggested to restrict the Hilbert spectra from individual trials to have the same number of bins in time-frequency space and then derive the mean of the Hilbert spectra. As a precaution, it is useful to compute a standard time-frequency representation, for example, STFT, to check the fidelity of the results.

Many signal processing techniques, such as the wavelet transform, are available for time-series decomposition. The uniqueness of the EMD method is that the decomposition is based on the local characteristic time scales of the data, and the basis functions (or IMFs) used to represent a given time-series are nonlinear functions that are directly extracted from the data. Therefore, the time scale is defined by the data per se, rather than by a pre-determined value. Fourier analysis, whose basis functions are limited to sinusoidal functions, cannot separate these IMFs without using pre-assigned cut-off frequencies. This is the crucial difference between EMD and Fourier-based filtering. Comparison with Fourier analysis has shown that EMD offers much better temporal and frequency resolution.

The WT is also an iterative decomposition process aimed at progressively scrutinizing finer and finer scales in a time-series. Compared to the EMD method, there are two major differences. First, linear time-invariant filters are essentially used in the WT decomposition, which precludes the possibility of adapting to local variations in oscillatory components. Second, the basis function (mother wavelet) of the WT is pre-determined, rather than directly extracted from the data as in EMD. Furthermore, a range of potential basis functions, from the simplest Haar wavelet to the more complicated higher-order Daubechies wavelet, are available for use. This is a potentially serious problem since there is no firm guidance for selection of the mother wavelet, and inappropriate selection may adversely influence the result of the analysis, particularly if the wavelet shape does not match that of the considered time-series in every time instant.

The outcome of selecting low-frequency components in individual trials is similar to the simple low-pass filtering approach, but the procedures are radically different. Low-pass filtering based on Fourier methods is a linear operation and an *a priori* determined cut-off frequency must be specified, whereas EMD-based filtering is nonlinear and its time scale is defined by the data itself. Unlike low-pass Fourier filtering, EMD-based filtering does not remove components

above an arbitrary cut-off frequency some of which might be useful in certain application such as constructing the AVEP.

While the HHT is a potentially powerful analysis tool for nonlinear and non-stationary data, great care must be taken in the physical interpretation of the decomposed IMF components. The decomposition does not guarantee a well-defined physical meaning, although the IMF components, in most cases, carry physically meaningful interpretations of the underlying dynamic processes. An example is shown in Fig. 8, where the analyzed signal (its spectrogram shown in Fig. 8a) consists of two chirp components with opposite chirp rates and their time–frequency representations cross each other. Figure 8b, c shows the spectrograms of the first two IMFs derived from the EMD. Clearly, there is a problem with the physical interpretation of these two IMFs since they are not consistent with physical intuition for the phenomena under investigation. This observation underscores the need for all of the relevant IMF components to be interpreted together if the data being investigated do not possess a clear, physically meaningful separation of scales.

The HHT is a potentially useful addition to our repertoire of non-stationary and nonlinear signal processing tools. Although conceptually quite simple, the technique still needs to be better understood. Due to the empirical nature of the EMD method, it lacks an analytical definition. The only way of better understanding the technique, so far, has been to resort to extensive numerical simulations in well-controlled situations (Flandrin et al. 2004; Wu and Huang 2004). Therefore, the lack of theoretical foundation of the method clearly calls for further theoretical analysis. In addition, implementation of EMD requires some further improvements, such as in management of the end points for cubic splines interpolation in the EMD process and selection of the stopping criteria for the sifting procedure. The influence of the end points is one problem encountered while implementing the EMD method; the problem becomes more dominant when the signal is very short. The decomposition relies on envelope calculations derived from a cubic spline interpolation between local extrema. The splines, however, are notoriously sensitive to the end points. It is thus important to make sure that end effects do not propagate into the interior and corrupt the data. We solve this problem by extending both the beginning and the end of the data by the addition of characteristic waves which are defined by the two consecutive extrema for both their frequency and amplitude. The extension takes place every iteration so that the additional waves are continuously changing in frequency and amplitude.

To summarize, we have introduced here a new method for analyzing neurobiological field potential data. The HHT method offers an alternative to, and advantages over, Fourier-based methods. We are enthusiastic that this new technique will prove itself of general value in the field of neural data analysis.

Acknowledgements Supported by the Whitehall Foundation, NIMH-IRP, NIMH 67776, NIMH 64204 and NSF 0090717.

References

Balocchi R, Menicucci D, Santarcangelo E, Sebastiani L, Gemignani A, Ghelarducci B, Varanini M (2004) Deriving the respiratory sinus arrhythmia from the heartbeat time-series using empirical mode decomposition. Chaos Solitons Fractals 20:171–177

Bendat JS, Piersol AG (1986) Random data: analysis and measurement procedures. Wiley, New York

Boashash B (1992) Estimating and interpreting the instantaneous frequency of a signal – Part I: fundamentals. Proc IEEE 80:520–538

Daubechies I (1992) Ten lectures on wavelets. SIAM: Philadelphia

Desimone R, Duncan J (1995) Neural mechanisms of selective visual-attention. Annu Rev Neurosci 18:193–222

Efron B, Tibshirani RJ (1993) An introduction to the boostrap. Chapman & Hall/CRC, London/Boca Raton

Flandrin P, Rilling G, Goncalves P (2004) Empirical mode decomposition as a filter bank. IEEE Sig Proc Lett 11(2):112–114

Freeman WJ (2004a) Origin, structure, and role of background EEG activity. Part 1. Analytic amplitude. Clin Neurophysiol 115(9):2077–2088

Freeman WJ (2004b) Origin, structure, and role of background EEG activity. Part 2. Analytic phase. Clin Neurophysiol 115(9):2089–2107

Fries P, Reynolds JH, Rorie AE, Desimone R (2001) Modulation of oscillatory neuronal synchronization by selective visual attention. Science 291:1560–1563

Gabor D (1946) Theory of communication. IEEE J Comm Eng 93:429–457

Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen N-C, Tung CC, Liu HH (1998a) The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time-series analysis. Proc R Soc Lond A 454:903–995

Huang W, Shen Z, Huang NE, Fung YC (1998b) Engineering analysis of biological variables: An example of blood pressure over 1 day. Proc Nat Acad Sci USA 95:4816–4821

Jenkins GM, Watts DG (1968) Spectral analysis and its applications. Holden-Day, San Francisco

Liang HL, Bressler SL, Desimone R, Fries P (2005a) Empirical mode decomposition: a method for analyzing neural data. Neurocomputing 65–66:801–807

Liang HL, Lin QH, Chen JDZ (2005b) Application of the empirical mode decomposition to the analysis of esophageal manometric data in gastroesophageal reflux disease. IEEE Trans Biomed Eng (in press)

Liang HL, Lin Z, McCallum RW (2000) Artifact reduction in electrogastrogram based on the empirical mode decomposition method. Med Biol Eng Comput 38:35–41

Mallat S (1998) A wavelet tour of signal processing. Academic Press, New York

Oppenheim AV, Schafer RW (1989) Digital signal processing. Prentice Hall, Englewood Cliffs

Percival DB, Walden AT (1993) Spectral analysis for physical applications. Cambridge University Press, New York

Potamianos A, Maragos P (1994) A comparison of the energy operator and Hilbert transform approach to signal and speech demodulation. Sig Processing 37:95 –120

Rilling G, Flandrin P, Goncalves P (2003) On empirical mode decomposition and its algorithms. In: IEEE-EURASIP workshop on nonlinear signal and image processing, Grado (I)

Wu Z, Huang NE (2004) A study of the characteristics of white noise using the empirical mode decomposition method. Proc Roy Soc Lond A 460:1597–1611