Journal of Neuroscience Methods 162 (2007) 49–63

JOURNAL OF NEUROSCIENCE METHODS

www.elsevier.com/locate/jneumeth

Comparison of spectral analysis methods for characterizing brain oscillations

Marieke K. van Vugt[a], Per B. Sederberg[b], Michael J. Kahana[c,*]

[a] *Department of Neuroscience, University of Pennsylvania, Philadelphia, PA, USA*
[b] *Department of Psychology, Princeton University, Princeton, NJ, USA*
[c] *Department of Psychology, University of Pennsylvania, 3401 Walnut Street, Rm. 302C, Philadelphia, PA 19104, USA*

Received 18 September 2006; received in revised form 13 December 2006; accepted 13 December 2006

Abstract

Spectral analysis methods are now routinely used in electrophysiological studies of human and animal cognition. Although a wide variety of spectral methods has been used, the ways in which these methods differ are not generally understood. Here we use simulation methods to characterize the similarities and differences between three spectral analysis methods: wavelets, multitapers and $P_{episode}$. $P_{episode}$ is a novel method that quantifies the fraction of time that oscillations exceed amplitude and duration thresholds. We show that wavelets and $P_{episode}$ used side-by-side helps to disentangle length and amplitude of a signal. $P_{episode}$ is especially sensitive to fluctuations around its thresholds, puts frequencies on a more equal footing, and is sensitive to long but low-amplitude signals. In contrast, multitaper methods are less sensitive to weak signals, but are very frequency-specific. If frequency specificity is not essential, then wavelets and $P_{episode}$ are recommended.

Keywords: Oscillations; EEG; Wavelets; Multitapers; $P_{episode}$

1. Introduction

Oscillations arise from an interaction between the intrinsic properties of neurons (excitability) and their interconnectivity, giving rise to synchronous activity (Buzsáki and Draguhn, 2004). Oscillations at various frequencies may be readily seen in electroencephalographic (EEG) recordings across species and are known to correlate with an animal's behavior and with the stimulating conditions present in the environment. Although early studies relied on visual inspection of the EEG signal to identify epochs of oscillatory activity and their behavioral correlates (Berger, 1929), the advent of modern computers now enables researchers to quantify the presence of oscillatory components in the EEG using spectral analysis methods. Spectral methods are widely used throughout the neurosciences and have yielded many new findings concerning the electrophysiology of both animal and human cognition (e.g., Klimesch et al., 1994; Kahana et al., 2001; Bastiaansen and Hagoort, 2003; Buzsáki and Draguhn, 2004; Kahana, 2006).

A myriad of spectral methods exist, which differ in the aspects of the data they highlight. However, exactly what aspects are highlighted by each method is often unclear. Our goal in this paper is to compare three methods used in the analysis of EEG oscillations. All three methods involve Fourier analysis. That is, they all seek to decompose the time series of EEG activity into sinusoidal functions whose amplitude and phase vary across frequency, but the shape of these functions differ for each method. In a traditional Fourier analysis, the function with which the signal is convolved[1] is a sinusoid of fixed length, and in order to improve temporal specificity, the analysis is performed on short windows ("windowing"). However, traditional Fourier analysis has a number of shortcomings: it has relatively poor time–frequency resolution (Bruns, 2004), the length of the window fixes the scale of the to-be-detected signal, and it is mainly designed for stationary[2] and regular signals (Mallat, 1998; Zhan et al., 2006). Because of the fixed window length, Fourier analysis is only useful in a limited frequency range that is optimized

* Corresponding author. Tel.: +1 215 746 3500; fax: +1 215 646 6848.
E-mail address: kahana@psych.upenn.edu (M.J. Kahana).

0165-0270/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2006.12.004

[1] A convolution measures the overlap between two functions by shifting them over one another and integrating over all shifts.

[2] 'Stationary' means that the signal has no significant change in its mean over time.

for your time window (Perrier et al., 1995). Moreover, none of the conventional oscillatory analysis methods differentiate oscillations from artifacts and evoked potentials, which can manifest as short events in Fourier space. Therefore, several alternatives have been proposed for analyzing brain oscillations, some of which will be discussed here.

We will examine three methods, namely wavelets (introduced for neural data in Kemerait and Childers (1972) and Schiff et al. (1994)), multitapers (introduced for the analysis of neural data in Mitra and Pesaran (1999)) and P_{episode} (introduced in Caplan et al. (2001)). Wavelets are functions that can come in many shapes, and the analyzed signal is decomposed into scaled and shifted versions of the oscillating waveform you are using. Because the wavelet functions are shifted and scaled versions of one another, the proportion between temporal width and frequency bandwidth remains the same for all frequencies. Therefore, a crucial difference from windowed Fourier analysis is that the size of the window depends on the frequency, which gives rise to more temporal precision for higher frequencies. Wavelets have a very good time–frequency resolution trade-off (Sinkkonen et al., 1995), making them quite useful for the analysis of non-stationary signals.

Multitapers are sets of functions that were designed to reduce bleeding between frequencies, rendering them well-suited for non-stationary processes with high dynamic ranges and/or rapid variations (Walden et al., 1998). An important distinction from wavelets is that the width of the function stays the same in absolute time across frequencies (similar to a Fourier transform). Finally, because *multi*tapers imply that the signal is convolved with multiple orthogonal tapers,[3] which are then averaged, the variance of this oscillation detection method across repeated measurements is reduced. In other words, the amplitude measurements taken with multitapers will have smaller error bars than, for example, wavelets.

Both wavelets and multitapers do not discriminate between short, high-amplitude power fluctuations and longer oscillations. The P_{episode} method addresses this issue because it was designed to detect "oscillatory episodes" and ignore transient voltage fluctuations (Caplan et al., 2001). This method characterizes whether oscillations at a given frequency are present or absent at a given time point in an ongoing EEG signal. It uses wavelets to determine the amplitude of oscillatory activity at a given frequency and time, and then applies an amplitude and duration threshold to characterize whether the signal is in an oscillatory state. Instead of measuring mean oscillatory power, Caplan et al.'s method measures the fraction of a time interval during which the signal exceeds the amplitude and duration threshold at a given frequency. This fraction is then termed the probability of being in an oscillatory episode at frequency *f*, or $P_{\text{episode}}(f)$. It is likely that the number of oscillatory cycles is more relevant than absolute length for information processing and computation (for an example see Jensen, 2006; Ward, 2003). Therefore, we will perform most analyses in this paper in units of oscillatory cycles at a given frequency as opposed to time in seconds, an alternative way of quantifying oscillatory power.

We will compare these methods by first analyzing simulated EEG data where the signal to be recovered is known. We then apply the three methods to empirical data. Comparing the results from simulations to effects in real data will allow us to highlight the differences between the three methods. We end the paper by offering recommendations to scientists interested in measuring oscillatory effects in EEG data.

2. Methods

2.1. Specifications of analysis methods

The first method we consider is wavelet analysis (Fig. 1a and d). Wavelets come in many shapes, each designed to capture different aspects of a time series. The Morlet wavelet is commonly used for the analysis of human EEG (Schiff et al., 1994) because its sinusoidal shape, which tapers at the ends, matches the signal we expect to extract from the EEG. This is crucial, because the success of wavelet analysis depends upon the suitability of the wavelet for detecting the desired signal (Özdemir et al., 2005). A small disadvantage of the wavelet is that it is non-orthogonal, hence computationally inefficient.[4] The Morlet wavelet is defined as follows (illustrated in Fig. 1g):

$$S^W = s(t) * \frac{1}{(\sigma_t\sqrt{\pi})^{1/2}}\, \mathrm{e}^{-(t^2/2\sigma_t^2)}\, \mathrm{e}^{2i\pi\, ft} \tag{1}$$

In this equation, S^W denotes the wavelet-transformed signal, $s(t)$ is the original signal, t and f represent time and frequency, respectively, and $*$ means convolution. The square root term causes the wavelet to be normalized to have an energy (squared integral) of 1. After the convolution, the absolute magnitude of the square of Eq. (1) will be taken. Wavelets have a length that scales inversely with frequency, such that the time–frequency product, or alternatively the number of cycles of oscillations within a wavelet, remains constant (the actual number of oscillations is set by the wavenumber k in $\sigma_t = k/\pi f$). It also means, however, that for higher frequencies the frequency resolution decreases and the temporal resolution increases (i.e., temporal and frequency resolution trade-off). In this paper, we use a wavenumber of 6, which is often used in human EEG analysis to strike a balance between temporal and frequency specificity (e.g., Sederberg et al., 2003, in press). In addition, the decomposition we use is a continuous wavelet transform, which has the advantage that we can investigate signals at arbitrary scales (as opposed to a decomposition at a fixed set of orthogonal frequencies–the discrete wavelet transform). However, it has the disadvantage that the wavelets used are not necessarily orthogonal (as is the case when the frequencies used are not logarithmically spaced) and, hence, the obtained power estimates are

[3] Tapers are functions that smooth the data by having a value of one in the middle and then slowly tapering off to zero at the edges.

[4] Non-orthogonal refers to overlap or correlation between the different wavelets, which essentially means that the overlapping part is convolved with your data twice. This is what is meant by 'computationally inefficient'.

redundant. The latter causes imprecision in the power estimates (Farge, 1992).

Researchers have recently begun to adopt multitapers (Thomson, 1982; Percival and Walden, 1993; Mitra and Pesaran, 1999) for EEG analysis (Raghavachari et al., 2001; Hoogenboom et al., 2006) (also see Fig. 1b and e). In the multitaper method, the data are multiplied with special windows before the frequency decomposition. The windows used are often Slepian windows or Discrete Prolate Spheroidal Sequences (DPSS), examples of which are shown in Fig. 1h. These window functions are designed to prevent bleeding of power to neighboring frequencies, as often tends to happen for wavelets. After this operation, a Fourier transform is performed, and the absolute square is taken of the resulting signal (an alternative procedure is to convolve the data with the DPSS window directly).[5] The convolution is repeated with a number of K different (orthogonal) windows (where the number K depends on the window size (T) and bandwidth (W), $K = 2TW - 1$). Each repetition reduces the variance in the estimate by $\sqrt{K}$ (Raghavachari et al., 2001). In fact, each window gives an independent estimate of the signal because the windows are orthogonal, thereby making the power estimate more reliable for noisy data. The complete estimate of the oscillatory power through this method then becomes:

$$S^M = \frac{1}{K}\sum_{k=1}^{K} \Delta t \left| \sum_{t=1}^{N} \mathrm{DPSS}_{t,k} s(t)\, \mathrm{e}^{-i2\pi f t\, \Delta t} \right|^2 \quad (2)$$

Here again t and f represent the time and frequency, and $s(t)$ denotes the signal. $\mathrm{DPSS}_{t,k}$ is the kth taper function at time point t. The DPSS functions at frequency f are defined as the Fourier transforms of the solutions to the following integral equation, and can be obtained easily from programs like MATLAB™:

$$\int_{-W}^{W} \frac{\sin T\pi(f - f')}{\sin \pi(f - f')} U_k(T, W; f')\, \mathrm{d}f' = \lambda_k(T, W) U_k(T, W; f') \quad (3)$$

In the simulations presented below, we use a bandwidth of 1 for the DPSS functions, giving us a frequency resolution of ±1 Hz, and a window size of 0.3 s (size of the taper), unless otherwise noted (note that this gives rise to only a single taper). The effect of using different parameters is considered in Section 4.

The advantage of multitapers is that they are designed to detect non-stationary signals with large-amplitude transients and also have good anti-frequency leakage properties. An important difference from wavelets is that, whereas the wavelet width changes with frequency, the width of the DPSS function (the envelope) remains constant, while periodicity varies with taper number.

Finally, the P_{episode} method (Caplan et al., 2001, 2003; Ekstrom et al., 2005) was designed to measure sustained oscillations, and is used after oscillatory power has been computed by another method (typically wavelets; Fig. 1c and f). This method was designed to detect sustained oscillatory processes, whereas conventional methods will detect any process that exhibits a partial oscillation cycle, such as artifacts (e.g., spikes) or evoked potentials. For example, P_{episode} could make the distinction between the sustained hippocampal slow oscillation and large-amplitude irregular activity that has been reported in rats (Wolansky et al., 2006). Even though an autocorrelation analysis can also detect differences between oscillatory activity of longer and shorter duration, it does not do so very well on a trial-by-trial basis. P_{episode} not only detects periodicity, but also quantifies the fraction of time spent in an oscillatory episode, and can thereby detect graded differences in the distributions of signal lengths between different conditions. In addition, when one does not use a pre-whitening or other filtering procedure to compensate for the $1/f^{\alpha}$ fall-off of the spectrum, autocorrelation analyses will detect high-frequency oscillations poorly because high-amplitude oscillations at lower frequencies dominate the correlation spectrum.

To compute P_{episode}, one first estimates the "background" spectrum by taking the mean wavelet power over all experimental time. We typically include the entire experiment in the background estimation because we are interested in oscillations that exceed a baseline during certain periods of the task. This method of background estimation was introduced in Schiff et al. (1994), but alternative background epochs could be used (e.g., time spent looking at a fixation cross before the task starts). This background is then fit with a linear function in log–log space, in accordance with the theoretical $1/f^{\alpha}$ power spectrum that EEG is thought to have (Freeman, 2006). This procedure will remove the frequency bias (higher power at lower frequencies).

We then process the entire EEG signal and mark, for every frequency, the time intervals that exceed the power threshold for a period exceeding the duration threshold (D_T) in cycles (usually set to 3). The power threshold is defined as the P_T^{th}, usually 95th percentile of the fit to the background power spectrum, meaning that on average 95% of the background signal is eliminated. P_{episode} is then defined as the fraction of the time interval of interest that exceeds both thresholds. The duration and power thresholds were chosen based on experimentation with different parameters but are quite robust to the exact choice of parameters (see e.g., Caplan et al., 2001; Fig. 6). The P_{episode} computation is illustrated in Fig. 1i.

The advantage of the P_{episode} approach is that it puts all frequencies on equal footing, whereas both wavelets and multitapers lose information at higher frequencies due to the $1/f^{\alpha}$ fall-off of the power spectrum (Linkenkaer-Hansen et al., 2001). In addition, P_{episode} only quantifies oscillations that are truly periodic because they must be sustained for a certain number of oscillatory cycles.

2.2. Simulations

In order to examine the characteristics of the different oscillatory analysis methods, we used simulated EEG in which the "signal" was known. Here, the signal s is defined as a simple sinusoid of known frequency f, phase ϕ, time course $[t_1, t_2]$ and

[5] This alternative approach is identical because the convolution of two functions is identical to the multiplication of their Fourier transforms.

amplitude a:

$$s(t) = \begin{cases} a \sin(2\pi f t + \phi), & t_1 < t < t_2 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

This signal was added to a background EEG, which has a $1/f^\alpha$ frequency spectrum, in agreement with human EEG. The background EEG was constructed as described in Yeung et al. (2004). Briefly, 500 sinusoids of the desired duration, with random phase and frequency where summed. The amplitude of each sinusoid was adjusted to create a $1/f^\alpha$ fall-off of the power spectrum. Sampling rate in this simulated EEG was set to 256 Hz, in agreement with most of our experimental data.

Trials were then simulated as 2-s sequences of EEG, over which power was averaged (the signal length of 2 s was chosen to facilitate comparison with empirical data). To avoid edge effects, buffers with one more second of data were appended to these sequences at each side and were discarded after wavelet/multitaper convolution. A non-parametric Wilcoxon rank sum test was conducted between samples of 100 trials with the added oscillatory signal and samples of 100 trials comprising only background EEG. A signal is "detected" when the p-value of this rank sum test is lower than 0.05. Each signal in such a test has a fixed length, frequency and amplitude; only the background varies.

The criteria we consider for comparing the three methods of interest are detection threshold, temporal characteristics, and frequency specificity. Detection threshold is a measure of the sensitivity of the method, i.e., how small can the signal be and still be detected? To determine the temporal characteristics of the methods, we investigate the effect of increases in amplitude versus signal length on the detected power. Finally, to assess frequency specificity we quantify the frequency spread of the detected signal.

2.3. Empirical data

In addition to the simulated data, we applied these analysis methods to a large dataset of human EEG described more extensively in Sederberg et al. (in press). Intracranial EEG (iEEG) was collected from 35 subjects who underwent long-term invasive monitoring to determine seizure focus in cases of pharmacologically intractable epilepsy. These patients had arrays of subdural and/or depth electrodes implanted for a period of 1–2 weeks to determine the focus of the epilepsy. The placement of the electrodes was determined by the patients' medical needs. The signal was amplified and sampled at 200, 256, 500, 512 or 1024 Hz by means of a Bio-Logic, XLTek, Neurofile or Nicolet EEG system (depending on the site). Signals were band-pass sfiltered between 0.3 and 70 Hz or 0.1 and 100 Hz (depending on the amplifier). Data were also notch-filtered at 50 or 60 Hz (depending on the country), to eliminate electrical line and equipment noise. A kurtosis threshold of five was used to identify and discard events with artifacts (Delorme et al., 2001).

While the iEEG was being recorded, subjects performed a delayed free recall task, which involved learning lists of 15 or 20 words, then solving simple arithmetic problems for ~ 20 s. After the arithmetic task, subjects recalled the words from the just-presented list in any order. Electrophysiological signals were synchronized with behavioral events to a precision of < 4 ms.

Sederberg et al. (in press) were primarily interested in the *subsequent memory effect* (SME) (Karis et al., 1984; Fell et al., 2001; Paller and Wagner, 2002)—the electrophysiological activity measured during the item presentations that distinguishes words that will subsequently be recalled from those that will not. This was investigated by comparing 2-s sequences of EEG (over which the power was averaged) for recalled versus not-recalled words. As with the simulated data, they avoided edge artifacts by including a 1-s buffer on either side of the 2-s epoch, which was discarded after wavelet/multitaper convolution.

3. Results

3.1. Basic characteristics of the three methods

To visualize the basic spectral properties of the three methods, spectrograms for simulated data are shown in Fig. 1. In this figure, we added a signal to background EEG activity between 200 and 800 ms, with a frequency of 10 Hz and an amplitude of 5% of the background activity at that frequency. Note that both the strength with which the signal is distinguished from the background activity and the structure detected in the background activity itself differ between the three methods. In particular, multitapers and P_{episode} tend to detect more structure in the random background EEG than wavelets. Also notice that P_{episode} does not fall off as a function of frequency, as both wavelets and multitapers do (however, this could be prevented by pre-whitening the data, i.e., removing autocorrelations such that the power spectrum becomes flat).

In order to compare the performance of these three methods, we investigated sensitivity of detection in amplitude, time, and frequency domains. In the amplitude domain, we determined the smallest signal that can be detected by each method, i.e., the detection threshold. In the time domain, we examined how detection of signal length and amplitude trade-off. Finally, an analysis of frequency specificity determined how much frequency bleeding occurs.

3.2. Detection threshold

First, we examined how wavelets, multitapers, and P_{episode} could detect signals as a function of their amplitude. Fig. 2 shows how the three analysis methods differ in terms of the minimum amplitude (in units of the amplitude of the background signal at that frequency) required to detect a signal (i.e., to find a significant difference between background EEG and signal using a rank sum test). If the method is more sensitive, then the minimum amplitude required to detect a signal will be lower. The variability of the methods is determined by repeating the procedure $n = 200$ times with different noise backgrounds.

We see that the detection threshold decreases with signal length, such that longer signals are easier to detect. This is to be expected because detection is dependent on mean power in a fixed time interval, and if the signal is longer, mean

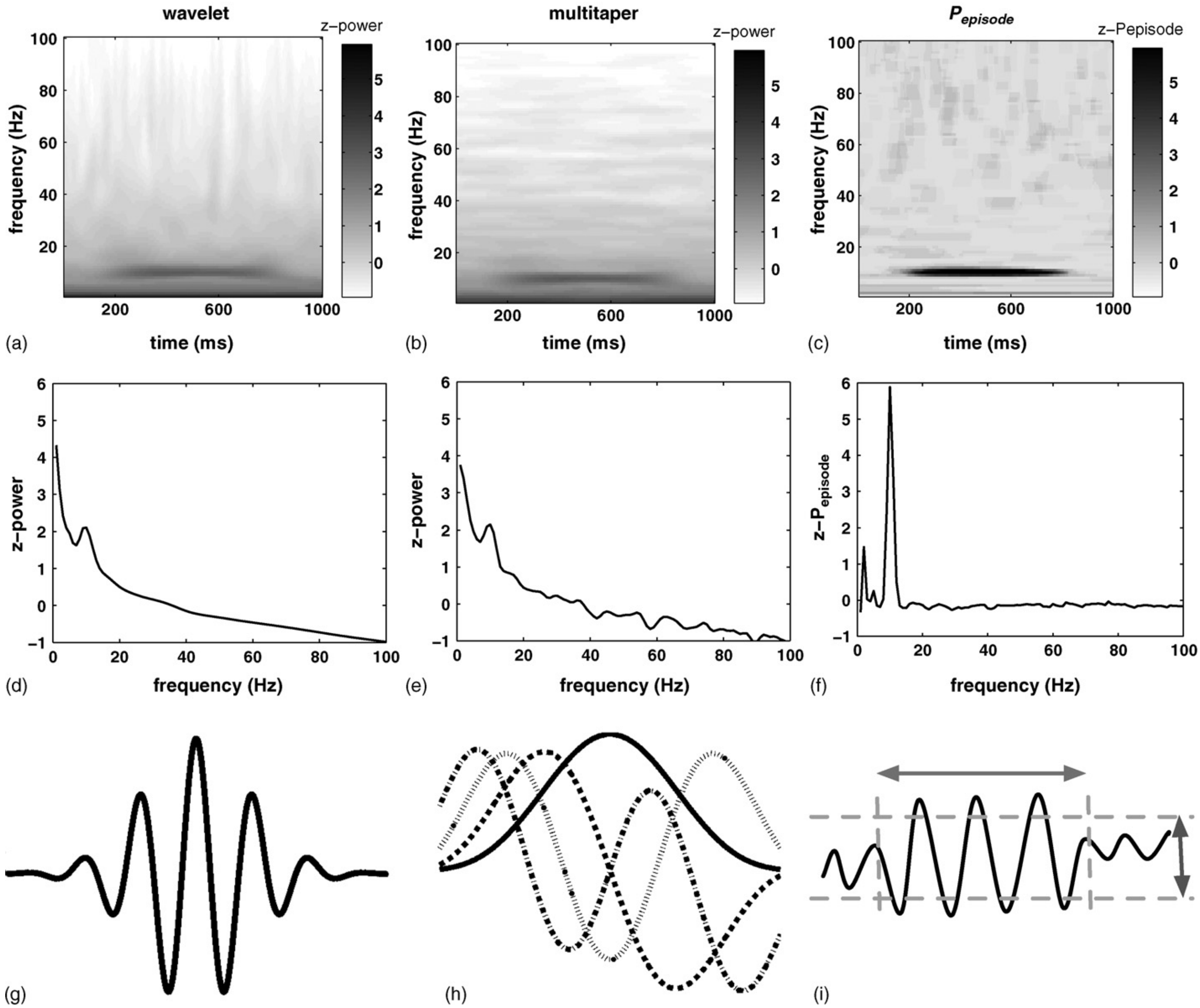

Fig. 1. Basic characteristics of the three oscillatory detection methods. Spectrograms (a–c) and spectra (d–f) of a 10 Hz signal of simulated data for wavelets (left), multitapers (middle) and P_{episode} (right). In all cases, a signal from 200 to 800 ms was added to a noisy background, with a frequency of 10 Hz and an amplitude of 5% of the background activity. These spectrograms/spectra show a mean over 100 trials. Technical details of each of the methods are shown in (g–i): a sample wavelet (g), sample multitapers (h) and illustration of the P_{episode} method (i).

power increases. Also note that the detection thresholds increase as a function of frequency (the slope of detection threshold versus frequency is significantly different from 0 for signals longer than one cycle by linear regression ($p < 0.0381$)). Just as the detection threshold decreases as a function of the signal length, the detection threshold increases as a function of frequency because the duration of the signal is much shorter at higher frequencies when the number of cycles is held constant. Consequently, it is more difficult to detect signals at higher frequencies.

We also see that at low frequencies (e.g., 5 Hz), wavelets are able to detect smaller and shorter signals than P_{episode}. This is related to the three-cycle duration threshold of P_{episode}, which prevents very short quasi-periodic activity from being detected. The reason that P_{episode} still sometimes detects signals smaller than three cycles is that high amplitudes can give rise to temporal spreading of the signal when measured by wavelets, causing them to occasionally exceed the P_{episode} duration threshold even when the signals themselves are below this duration threshold. In general, multitapers fall between wavelets and P_{episode} in their sensitivity, or are comparable to wavelets. At higher frequencies, however, multitapers become less sensitive, whereas P_{episode} increases in sensitivity.

3.3. Trade-off between time and amplitude

Increasing either the amplitude or duration of an oscillatory signal will make it more easily detectable via the wavelet and multitaper methods. In contrast, the P_{episode} method is primarily sensitive to signal length. Thus, when P_{episode} is used in combination with either wavelets or multitapers, one can, in principle, disambiguate the length and amplitude of signals.

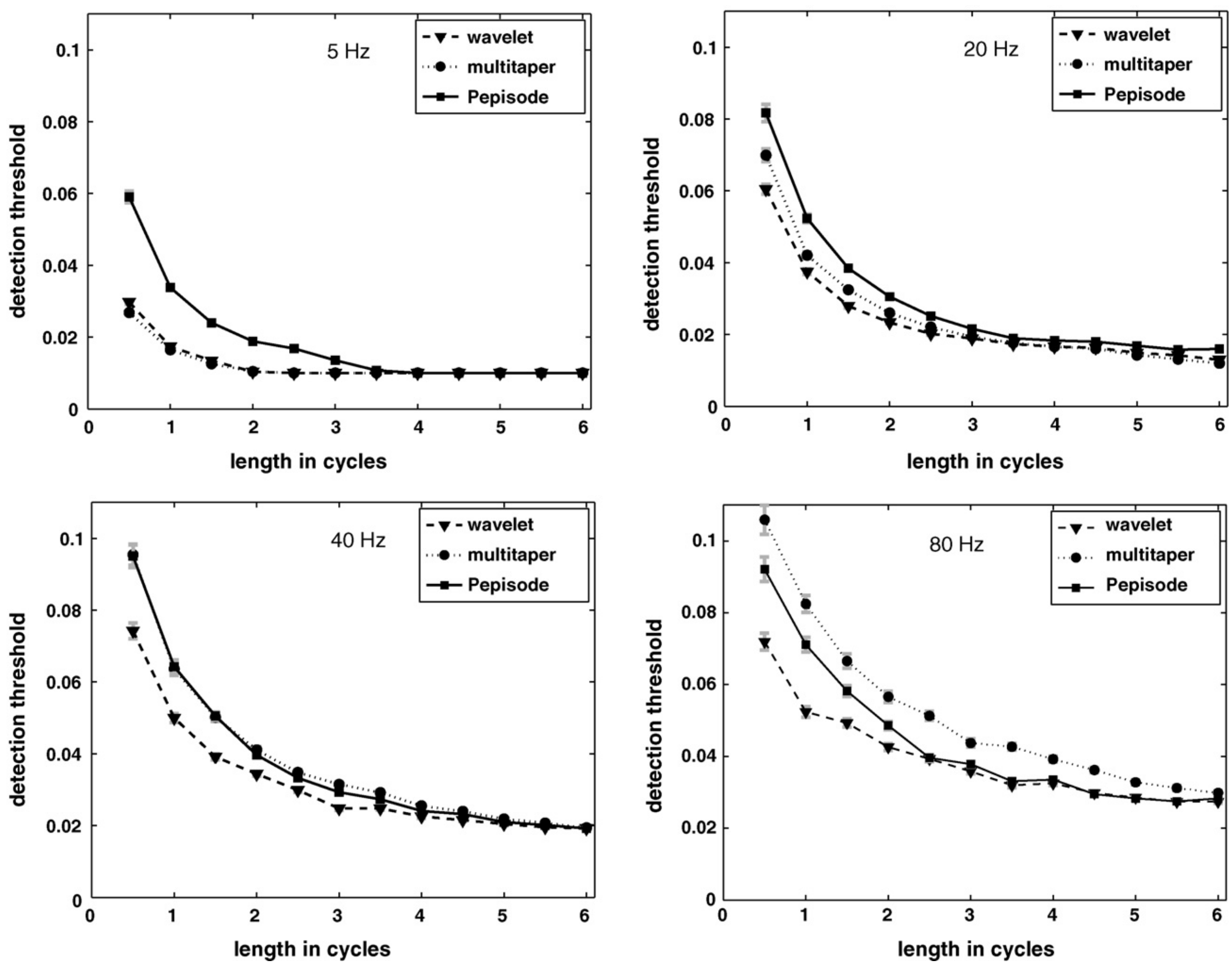

Fig. 2. Detection thresholds. Detection threshold of a signal (minimum signal amplitude that can be detected) as a function of signal length in cycles (one cycle has a length of 1/frequency seconds). The amplitude of the signal is expressed as fraction of the amplitude of the background EEG at that frequency (error bars are standard error of the mean).

We examined the responses of wavelets, multitapers and P_{episode} to increases in the to-be-detected signal length, amplitude, or their combination. We quantified the responses by determining the increase in measured oscillatory activity (either normalized power or P_{episode}) due to an increase in signal amplitude (all amplitudes used were supra-threshold for P_{episode}) as a function of signal length in cycles. Each response measurement was normalized by dividing it by the maximum response for that method, thereby placing all three methods on the same scale. This measure therefore shows what effect an increase in amplitude has on the detection of the signal by each method. When the index is near zero, it indicates that the detection method is not sensitive to amplitude, whereas large values indicate that the method responds strongly to changes in the amplitude of the signal.

Fig. 3 shows how P_{episode} is mostly sensitive to increases in length of the signal, not its amplitude, whereas multitapers and wavelets are sensitive to both (i.e., the estimates of power they provide increase by a large amount when either the amplitude or the length of the signal to be detected increases). The data in Fig. 3 are for a signal at 40 Hz; results at other frequencies are similar, except that lower frequencies tend to be contaminated more by duration threshold issues of P_{episode}.[6]

By comparing the results from P_{episode} and the other methods, we can disambiguate length and amplitude of signals because when *only* P_{episode} detects a signal it is likely that the underlying oscillation has a fairly long duration, but does not have a high amplitude compared to the no-signal condition. Conversely, P_{episode} will be less likely to detect signals that are very short in number of cycles (as shown in Fig. 2), even if the amplitude is high.

3.4. Frequency specificity

Although multitapers are not very sensitive to small-amplitude signals, they were designed to have good frequency specificity (due to their anti-leakage properties). This feature is

[6] For lower frequencies, we would need a very long (in time) signal for the cycle range we study here, which does not fit in a 2-s interval. At higher amplitudes, power will bleed across time, and the detected signal may pass the duration threshold, thereby leading to an apparent increase of P_{episode} as a function of amplitude.

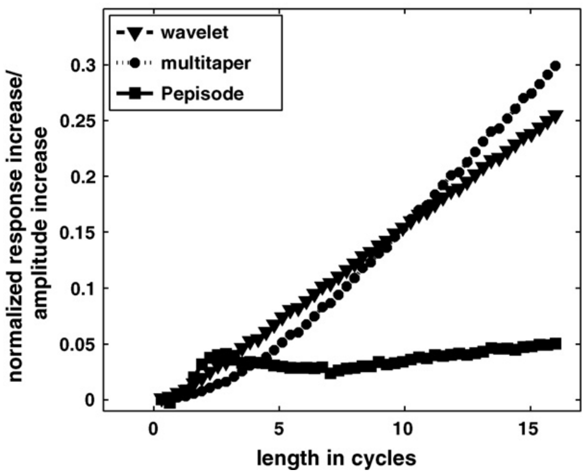

Fig. 3. Trade-off between amplitude and length of signals. Increase in normalized response of the oscillation detection method divided by the increase in amplitude. The signal is at 40 Hz and added at three different amplitudes. When this index is near zero, the detection method is not sensitive to amplitude, whereas large values indicate a strong response to amplitude.

shown in Fig. 4 in which we examine how the different methods respond to a signal at frequencies surrounding the frequency of interest. For frequencies ranging from 1 to 100 Hz we used a rank sum test to assess the reliability of the difference between samples of artificial EEG data with and without an embedded oscillatory signal. Fig. 4 illustrates the *p*-value resulting from the rank sum test at each frequency and for each of the three spectral analysis methods. When the *p*-value is higher than the dotted line (corresponding to $p = 0.05$), the signal is considered to be detected. The spread around the signal frequency (width of the peak) is an indication of frequency specificity. We added two types of signals: a signal with a fixed duration of 500 ms (left column of Fig. 4), and a signal with a duration of four cycles of the oscillation (right column of Fig. 4). The signal amplitudes were 1 and 4%, respectively, of the background signal at each frequency of interest (we used different amplitudes in order to to remain within the dynamic range of the different detection methods).

The height of the peaks shown in Fig. 4 determines how well the signals are detected. When comparing the peak heights across different repetitions of the simulation with a rank sum test, almost all heights of the peaks are different between methods. The only case where this is *not* true is when comparing multitapers and wavelets at 5 Hz: $p = 1$[7] and $p = 0.370$ for a four-cycle and a 500-ms signal, respectively. The width of the peak is a measure of the frequency specificity. We quantify the width of the peak by comparing distributions of the number of points above the significance threshold, for every repetition of the simulation. The widths are different between the three spectral methods when using a rank sum test, with the only exceptions being the comparison of multitapers and wavelets at 20 and 80 Hz for a 500-ms signal ($p = 0.136$ and $p = 0.440$); and between wavelets and P_{episode} for a four-cycle and a 500-ms signal at 80 Hz ($p = 0.070$ and $p = 0.42$, respectively).

Another interesting feature of Fig. 4 is the comparison of absolute length versus time in cycles across frequencies. For lower frequencies, the frequency specificity of the different methods is virtually identical for both the fixed and variable-length cases. However, as the frequency of the to-be-detected signal increases to 40 Hz and higher, P_{episode} and wavelets show more frequency bleeding for the fixed-length signal (P_{episode} to a lesser extent), whereas multitapers remain compact. Also note the general fall-off in significance levels (i.e., the height of the peaks) as frequencies increase, especially for the signal with a fixed number of cycles (right column of Fig. 4). These differences in measurements between the three oscillatory methods demonstrate the effect of considering a signal in terms of a fixed cycle length versus a fixed duration in seconds, where P_{episode} has an advantage for when cycle length is kept constant and multitapers have an advantage when signal duration is kept constant.

3.5. Signals near-threshold

Our previous analyses showed how P_{episode} is not very sensitive to signals below its duration threshold of three cycles, whereas wavelets are, especially at low frequencies. At higher frequencies however, P_{episode} tended to become more sensitive than wavelets for short but supra-threshold signals. These results are valid when a signal is compared to a noise background. However, from its construction, P_{episode} is *especially* sensitive to differences between conditions that are right around its amplitude and duration thresholds. In order to test this, we generated populations of signals of different amplitudes and lengths, and compared them to each other, to find the smallest distance between populations that still allowed them to be distinguished.

In Fig. 5, we show how detection decreases when the signals in the two comparison conditions become more similar in amplitude (left column) or length (right column). In each of these simulations, a sample with the maximum amplitude/length is compared to a sequence of samples with smaller amplitudes/lengths. As expected, the significance decreases for higher amplitudes/lengths because the difference between the two populations decreases (note the logarithmic scale for the *p*-values, which shows smaller *p*-values as higher data points on the ordinate). The error bars (standard error of the mean) are created by repeating the simulation 100 times with a 1% variation in amplitude (when the length is the predictor variable) or length (when the amplitude is the predictor variable), in order to mimic differences between trials within a certain condition (as in actual experiments).

For all frequencies larger than 5 Hz, we see the predicted result of a regime where P_{episode} detects a difference between very similar conditions (in terms of amplitude or duration of the signal) better than multitapers or wavelets. At 5 Hz, P_{episode} does not have an advantage in the amplitude regime, likely due to ceiling effects. The conditions shown in this figure are likely to appear in empirical EEG data, where two conditions are likely to differ in gradation between two signals than in the total presence or absence of a signal (as we explored in the previous simula-

[7] In this case, both wavelets and multitapers are always detected and reach the maximum possible detection statistic, and hence the two vectors are identical.

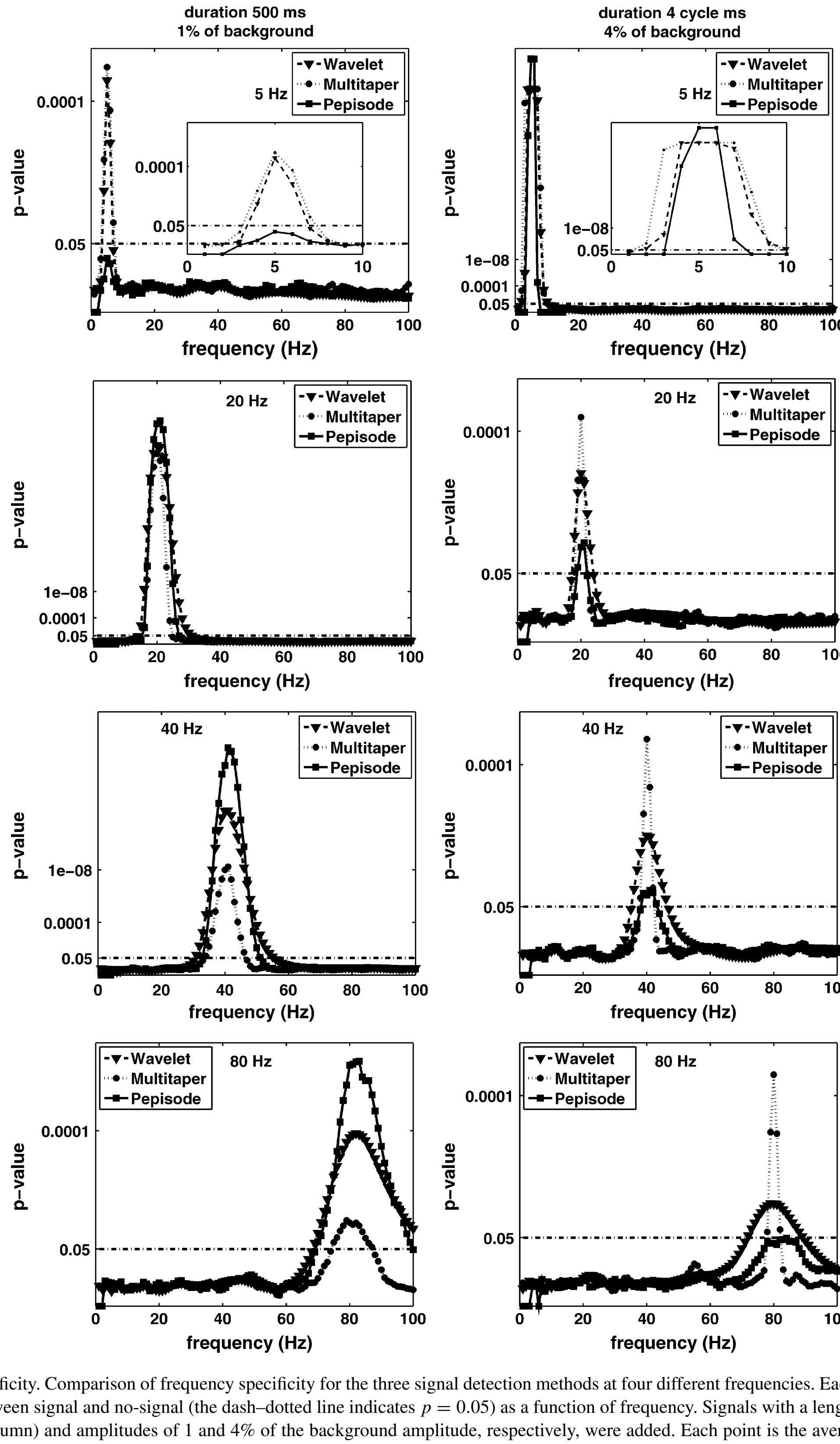

Fig. 4. Frequency specificity. Comparison of frequency specificity for the three signal detection methods at four different frequencies. Each panel shows the p-value for the comparison between signal and no-signal (the dash–dotted line indicates $p = 0.05$) as a function of frequency. Signals with a length of 500 ms (left column) or four cycles (right column) and amplitudes of 1 and 4% of the background amplitude, respectively, were added. Each point is the average p-value for 50 runs of the analysis. Insets show a close-up for the frequencies 0–10 Hz.

tions). Also note that P_{episode}'s advantage is bigger for length variations (right column) than for amplitude variations (left column), which is in agreement with Fig. 3, reaffirming that P_{episode} is more sensitive to differences in signal length than amplitude.

3.6. Differences in detection of very short, half-cycle signals

The detection-threshold analysis showed that P_{episode} rarely detects very short signals (with fewer cycles than its duration

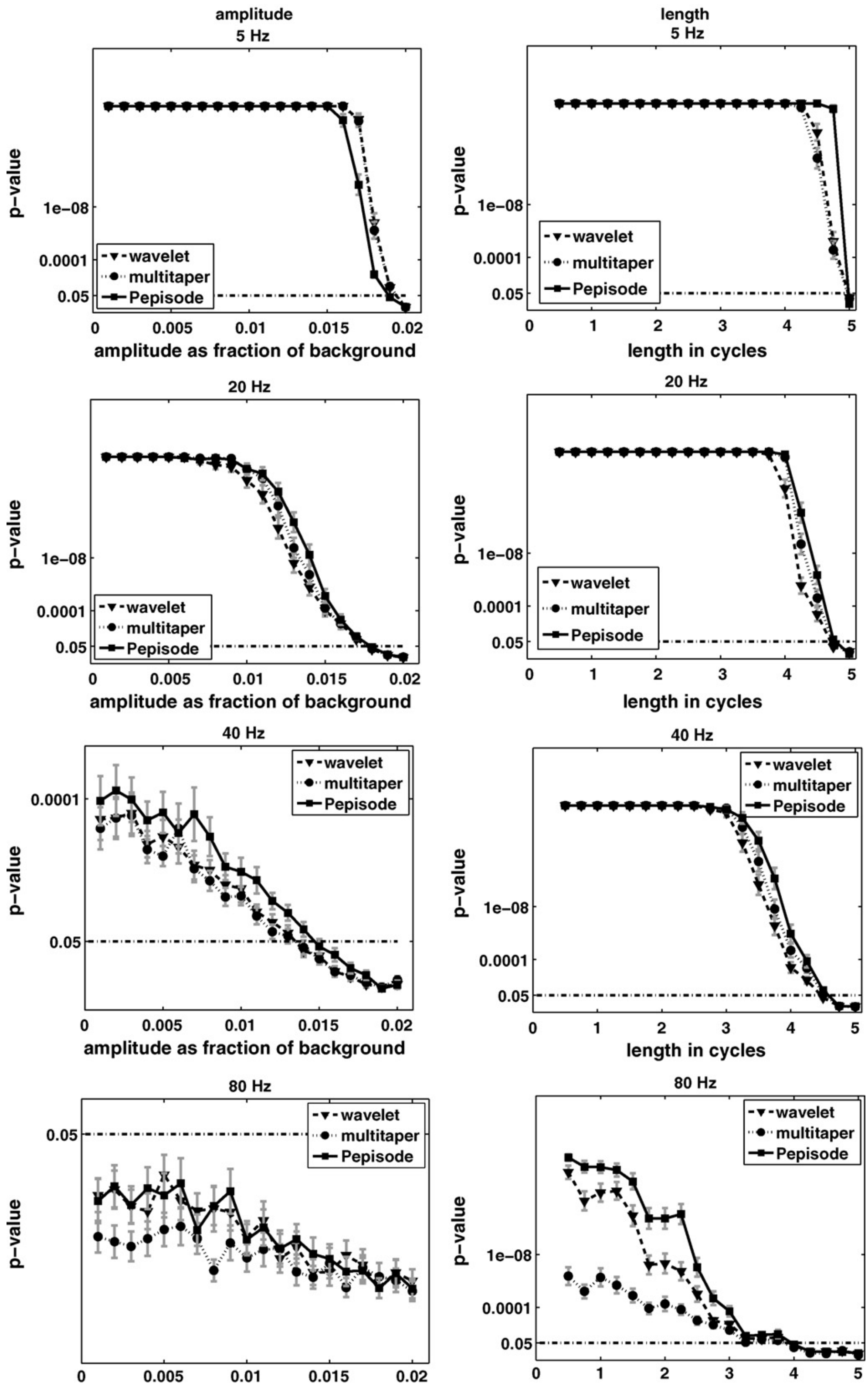

Fig. 5. The effect of near-threshold populations. To investigate near-threshold signal detection in the three methods, a population of signals with the highest amplitude (left column) and length (right column), respectively, is compared to populations with varying lengths and amplitudes. Lower *p*-values indicate better detection. P_{episode} shows a regime with better detection than the other two methods when the populations get closer to one another (i.e., the graphed amplitude/length comes closer to the maximum amplitude/length). Error bars are standard error of the mean.

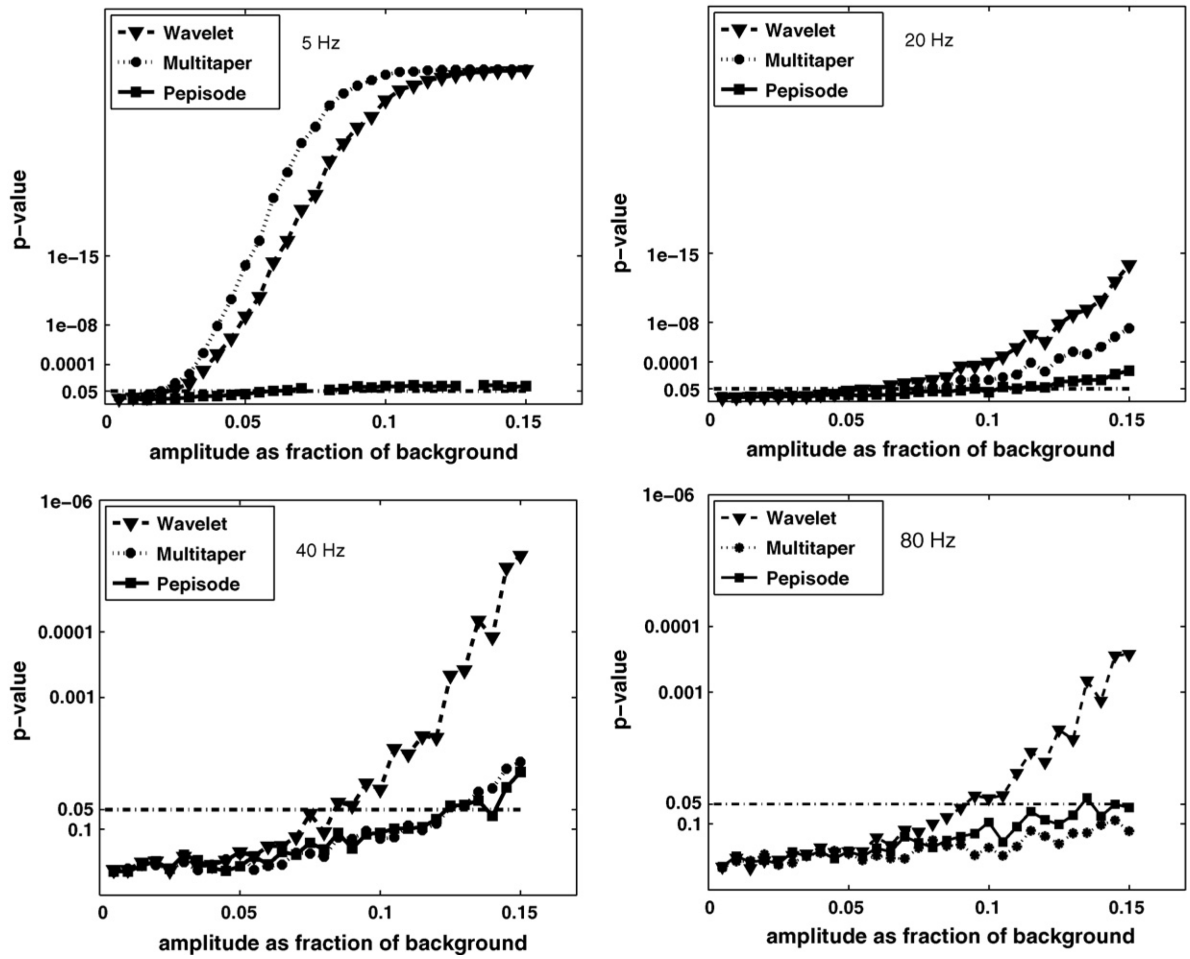

Fig. 6. The effect of amplitude on the detection of a very short signal. A signal of half a cycle is added to the background EEG with varying amplitudes (abscissa). Detection is measured as the *p*-value of the rank sum test comparing populations of samples with and without signals.

threshold). We decided to demonstrate this feature empirically, as shown in Fig. 6. When we added an oscillatory signal lasting only half a cycle to the background EEG data, both wavelets and multitapers detected these oscillatory signals, whereas the P_{episode} method did not. This is because P_{episode} requires oscillatory power to exceed an amplitude threshold for a minimum of three cycles in order to be considered an oscillatory episode. Thus, P_{episode} will be restricted to *sustained* oscillatory phenomena.

Further, this analysis also shows that for higher amplitudes, multitapers exhibit a more significant difference between the signal and no-signal populations than for lower amplitudes (they have the most significant *p*-values, and the difference in *p*-values is largest at the higher amplitudes). This demonstrates an advantage for the detection of high-amplitude signals by multitapers.

3.7. Application to intracranial EEG data

Whereas the previous analyses were limited to simulated EEG data, we also applied our spectral analysis methods to human intracranial EEG data. For this purpose, we selected the large dataset reported by Sederberg et al. (in press). In that study, subjects studied lists of common words for a subsequent recall task. Using wavelet analysis methods, Sederberg et al. (in press) found that specific patterns of oscillatory power during the period when a word was being studied predicted the word's subsequent recall.

In order to examine how the three methods discussed here agree in this dataset, we first computed the correlation between mean power or P_{episode} for each 2-s word presentation event for each electrode. Fig. 7, which contains histograms of these correlations (where each data point is the correlation at one electrode), gives us some insights in the similarity of the signals detected by these methods.

Wavelets and multitapers correlate more highly with one another than with P_{episode} ($F = 4506$ ($p \ll 10^{-8}$), two-way ANOVA), as we would expect based on, in particular, Fig. 3. This correlation decreases on average for higher frequencies ($F = 1844$ ($p \ll 10^{-8}$), two-way ANOVA), mostly because at higher frequencies multitapers will not detect signals of only a few cycles at higher frequencies, whereas P_{episode} still can if the number of cycles is above the duration threshold. Therefore, the three methods detect slightly different populations of signals, which will differ along one or more of the dimensions analyzed in the previous sections.

To illustrate more precisely the similarities and differences between the three methods in an analysis of empirical data, we calculated the significance spectra for two sample electrodes

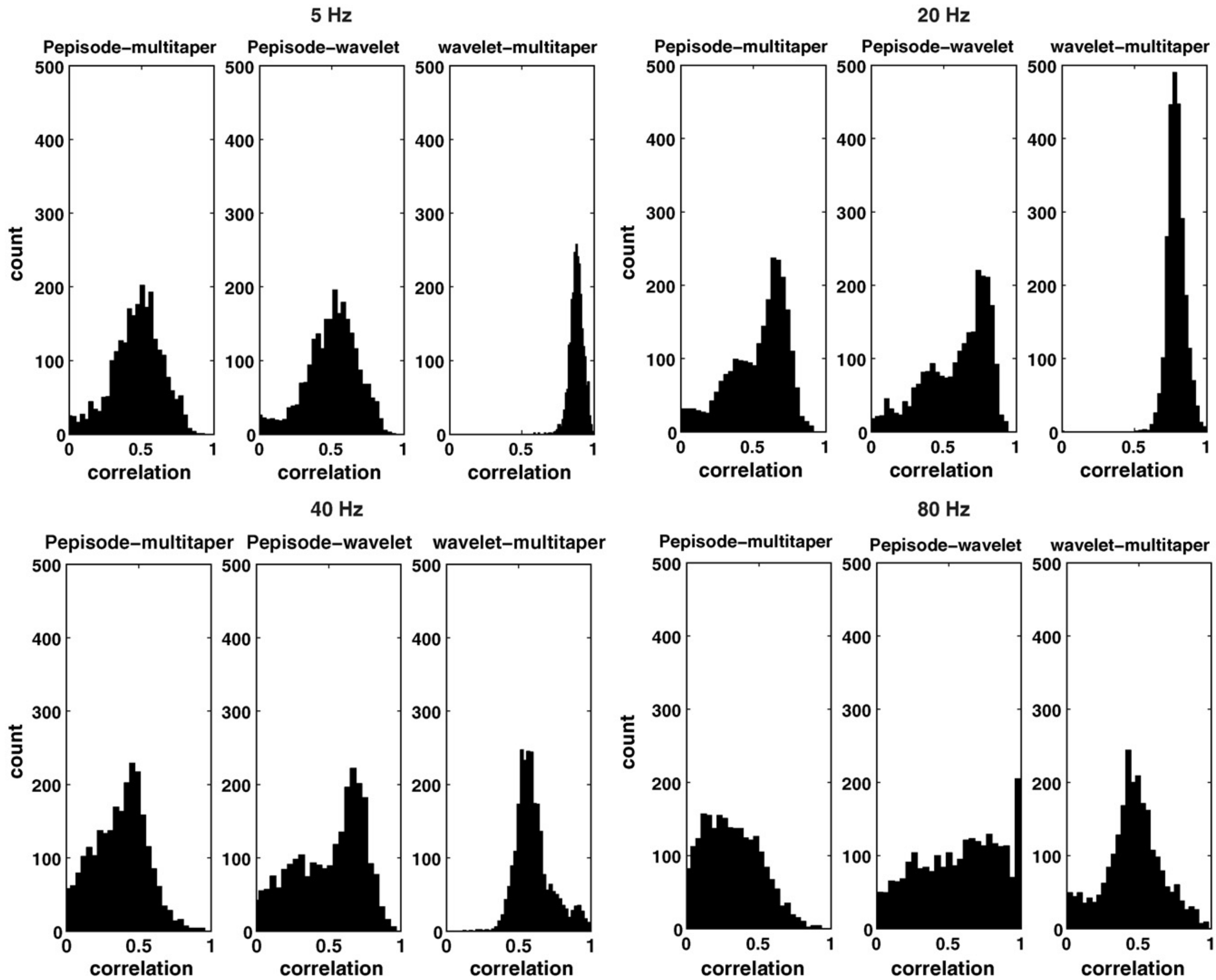

Fig. 7. Correlations across 2-s trials (for different electrodes) between mean wavelet/multitaper power and $P_{episode}$ at different frequencies.

published in Fig. 1 of Sederberg et al. (in press). Fig. 8 shows the significance of the difference between the power during recalled and not-recalled words for the same two electrodes using wavelets, multitapers and $P_{episode}$, respectively. In the top electrode (located in the hippocampus), $P_{episode}$ shows a peak at 32 Hz that the other methods do not show, and, conversely, wavelets and multitapers detect a negative peak at 6–12 Hz and a positive peak at 64 Hz that $P_{episode}$ does not. In the second electrode (located in the left inferior pre-frontal cortex (LIPC)), $P_{episode}$ shows a low-frequency peak that wavelets and multitapers do not show. The significance spectra for wavelets and multitapers are quite similar. As discussed earlier, the peaks uniquely detected by $P_{episode}$ are likely due to a relatively sustained signal that is not high-amplitude. The peaks that $P_{episode}$ does not detect, but wavelets and multitapers do, are likely due to short, high-amplitude signals. Also note that for the second electrode, multitapers distinguish two peaks at 32 and 60 Hz, whereas those are blurred into a single peak by wavelets. $P_{episode}$ shows a hint of the separation of the two peaks. These results are to be expected based on our frequency specificity analyses (Fig. 4).

4. Discussion

We reported analyses of three spectral methods: wavelets, multitapers, and $P_{episode}$. Wavelets perform a frequency decomposition by convolving the data with a localized sinusoid that scales with frequency. Multitapers are based on a windowed Fourier transform, where the windows are shaped to minimize power leakage. Instead of measuring oscillatory power, as the other two methods do, $P_{episode}$ measures the length of time during which an oscillation is present. It has amplitude and duration thresholds that allow for comparisons across frequencies.

To compare these three methods, we generated artificial EEG data that consisted of a $1/f^{\alpha}$ spectral background plus a to-be-detected oscillatory signal. Varying the target signal's properties allowed us to compare each method's ability to detect oscillations under different conditions. We compared the three methods according to the minimum amplitude they were able to detect, their trade-off between time and amplitude in signal detection, and their frequency specificity. By performing many of these analyses in units of cycles of an oscillation as well as normal time, we were able to observe differences between the novel

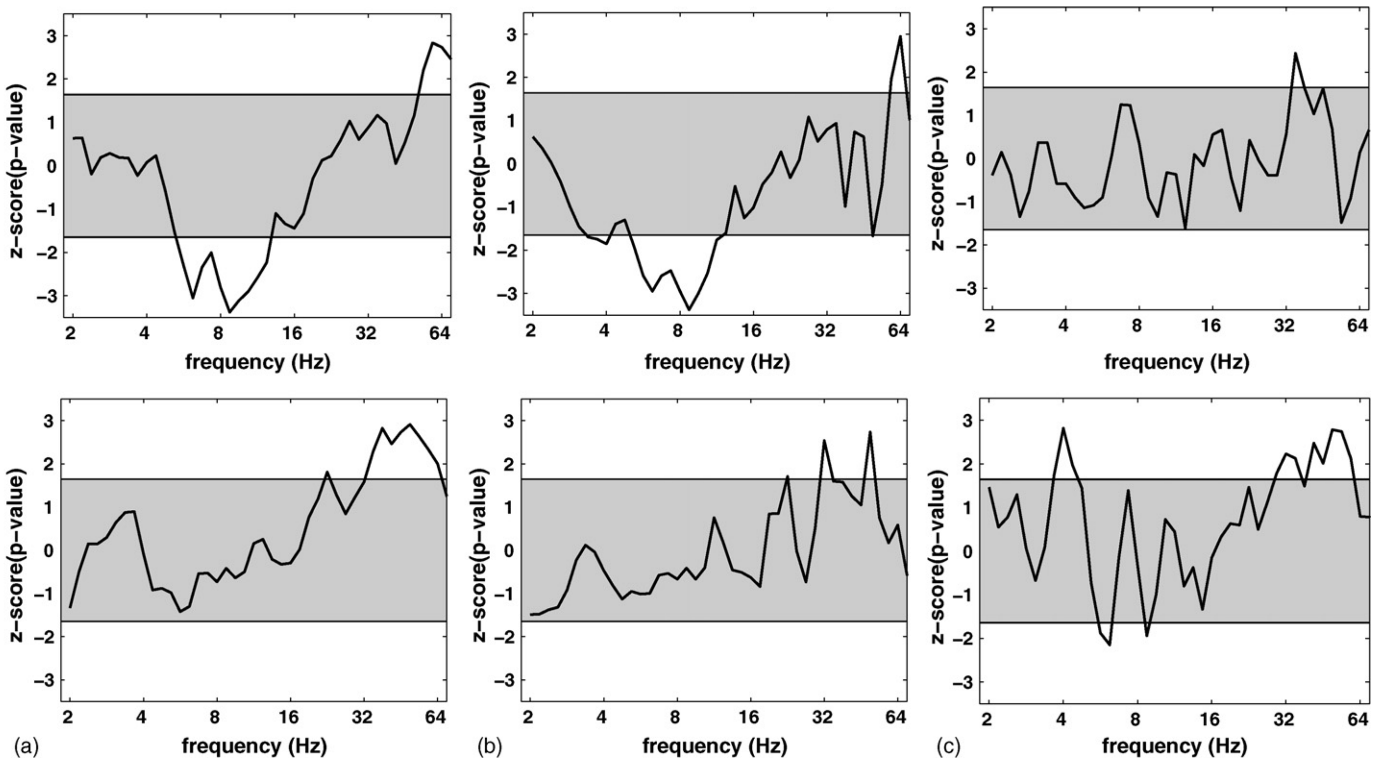

Fig. 8. Sample electrodes exhibiting oscillatory subsequent memory effects (first row: hippocampus; second row: LIPC (BA47)). Each panel shows the *Z*-transformed significance value of the difference in power between recalled and not-recalled words in the time bin 1000–2000 ms after word onset (rank sum test). The sign indicates the direction of the effect, and the gray area indicates the $p > 0.05$ significance threshold. The wavelet spectra in the left column are replications of Sederberg et al. (in press), Fig. 1(a and c); middle and right-hand columns show multitaper and P_{episode} spectra, respectively, for the same electrode. Notice how the different methods can detect different types of signal.

P_{episode} method and the more commonly used wavelets and multitapers.

As expected, P_{episode} did not detect very short signals in units of oscillatory cycles (Fig. 2). Also, its detection capability was much less frequency dependent than either wavelets or multitapers (Fig. 4). We demonstrated how wavelets are, in general, quite sensitive to low-amplitude signals (Fig. 2), whereas multitapers can best detect low-frequency high-amplitude signals (Fig. 6) and have a high degree of frequency specificity (Fig. 4). In addition, we showed how P_{episode} is not sensitive to the amplitude of the signal (once it exceeds the amplitude threshold), whereas the other two methods are (Fig. 3). These effects were further demonstrated in empirical data from an intracranial EEG study of human episodic memory (Figs. 7 and 8).

In contrast to wavelets and multitapers, P_{episode} was specifically designed to detect oscillatory episodes. Numerous theories propose that many of the processes that underlie cognition and awareness are sub-served by oscillatory synchrony between brain areas, where, for example, oscillations have to be sustained to carry information (see e.g., Engel et al., 2001; Varela et al., 2001; Tononi and Edelman, 1998). Because P_{episode} subtracts background activity, it can be much more sensitive than either wavelets or multitapers when conditions differ only slightly from one another, but have distributions that are near the duration or amplitude thresholds of P_{episode} (cf. Fig. 5). We performed many of the analyses in this paper in units of cycles of oscillations and fractions of background activity, both to highlight these subtle detection capabilities of P_{episode} in comparison to the other oscillation detection methods and to introduce this new way of thinking about oscillations.

Because P_{episode} will only detect oscillations that exceed a duration threshold in cycles, one could use P_{episode} as a first-pass filter to find real oscillations, and then use wavelets or multitapers to detect their amplitudes. In fact, using P_{episode} in concert with multitapers and wavelets, one could disentangle length and amplitude of the detected signal, as shown in Fig. 8. Another consequence of this approach is that oscillations of the same number of cycles are detected roughly equally at different frequencies by P_{episode}, whereas the ability of multitapers and wavelets to detect signals with the same number of cycles decreases at higher frequencies because they are shorter in time (e.g., Fig. 2).

Nonetheless, a number of questions remain. How crucial are the particular methods and analyses we used for obtaining these results? How could changing the parameters that these methods use affect the results? How do the assumptions about the $1/f^{\alpha}$ background spectrum affect P_{episode}? Can these results be extended to other types of analysis, for example instantaneous measures of power as opposed to averaged power? Finally, how good are the simulation methods? These questions will be discussed in the next sections.

4.1. Differences between methods due to parameters

It has been observed (Bruns, 2004) that differences between oscillation detection methods (in that case Hilbert transform, wavelet transform and Fourier transform) could be largely eliminated when the equation parameters were adjusted. In this manuscript, the parameters that have the greatest effect on signal detection are the wavenumber for wavelets, window size and bandwidth for multitapers, and the amplitude and duration threshold for P_{episode}.

The frequency specificity for wavelets could potentially be improved by using longer wavelets (which then decreases temporal resolution). For example, at higher frequencies (80 Hz), an increase to a wavenumber of 8 will narrow the signal detection peak to about 2/3 that of a wavenumber of 5, but at 5 Hz, there is a much smaller effect ($\sim 9/10$). Given that P_{episode} is based on the wavelet measure, its frequency specificity will follow that of wavelets, but will always be slightly more narrow (see Fig. 4). Conversely, increasing the bandwidth and/or decreasing the window size would not only decrease the frequency specificity of multitapers (or increase their leakage across frequencies (Bronez, 1992)), but also decrease their variance. Increasing the bandwidth from 1 to 5 will widen the peak (see Fig. 4) by roughly a factor 5 at 80 Hz, but only a factor 2 at 5 Hz. Increasing the window size from 0.1 to 0.5 will narrow the frequency specificity peak by a factor of 3 at both 5 Hz and 80 Hz. Furthermore, increasing the bandwidth and decreasing the window length will also slightly decrease the amplitude sensitivity, especially in the beta–gamma band (20–80 Hz), with roughly a factor 0.9. In conclusion, wavelets and multitapers can be made to behave more similarly when adjusting their parameters. P_{episode} however, because it measures a proportion of time as opposed to oscillatory power, will always behave distinctly.

For both wavelets and multitapers, changing their basis function will also affect both amplitude and frequency sensitivity, but those effects are harder to predict. The crucial difference between wavelets and multitapers, however, is that in wavelets the signal is *convolved* with the basis function, whereas for multitapers the signal is *windowed* with the basis function and then subjected to a FFT procedure. The multitaper parameters in the current simulations were set such that the multitaper spectrograms looked visually like those of the wavelets. A crucial difference however is that for wavelets, the temporal and frequency specificity scales with frequency, whereas for multitapers it does not. One interesting approach, developed by Lilly and Park (1995), is to merge the characteristics of wavelets and multitapers into a multiwavelet. [8] In a multiple wavelet analysis the functions with which the signal is convolved are Slepian-like (i.e., similar to the DPSS functions used in multitaper spectral analysis, which reduce frequency leakage compared to ordinary wavelets) and scale with frequency. Multiple orthogonal Slepian-like functions are then averaged (note that in contrast to wavelets, the Slepian-like functions are not scaled versions of one another, even though they are self-similar). This multiwavelet approach has yielded better time and frequency resolution than multitapers, especially for events with high variability (Zanandrea et al., 2004).

Finally, by decreasing the duration threshold, P_{episode} will act more like wavelets (Caplan et al., 2001). Note that P_{episode} will still be qualitatively different because it measures time in an oscillatory episode as opposed to the oscillatory power of that episode. In addition, P_{episode} can be made more frequency-sensitive at higher frequencies by using multitapers as power estimates, at the cost of being unable to detect very small signals.

4.2. Assumptions concerning the background spectrum for P_{episode}

One concern is whether P_{episode} depends on the assumption of a $1/f^{\alpha}$ shape of the power spectrum. In real data, the spectrum may often have “bumps” at particular frequencies, such as theta, or may diverge from $1/f^{\alpha}$ at low or high frequencies due to high- or low-pass filtering. We have repeated some of our analyses (e.g., frequency specificity) with spectra that were not linear in log–log space, and found that the qualitative results did not change very much (for frequency specificity we still found the best specificity for multitapers when the signals were relatively large amplitude and for P_{episode} when the signals were relatively short). Appropriate fits, even for non-linear spectra, can be further ensured by using a robust fit algorithm as opposed to a least-squares fit, which is less affected by outliers. Other alternatives would be fitting a non-linear function to the background, or fitting only the part of the frequency spectrum that is linear and removing the other frequencies from consideration. In order to determine the correct fit, it is important to visually inspect the fits to the background spectrum for the dataset of interest before interpreting P_{episode} results.

Also, it should be mentioned that, in principle, any fit to the background spectrum could be used, depending on the research question one is interested in. For example, when one is interested in how many oscillations in the delta range are present in one sleep stage compared to another, one would fit the peaked spectrum in the first sleep stage and use that as the background distribution for the other. It is important to keep in mind that the resulting data can only be interpreted with reference to the baseline one uses.

4.3. Interval power versus instantaneous power

In the simulations described above, we have presented the results for mean power over a certain interval in time. An alternative analysis would be of instantaneous power at a particular moment in time. In fact, in order to detect instantaneous power and phase information, better methods than those discussed here are available, such as the Hilbert transform (used by e.g., Freeman and Rogers, 2002) and its cousin, empirical mode decomposition (Huang et al., 1998). P_{episode}, however, is by definition applied to intervals, not isolated time points. Nevertheless, it would be possible develop a form of P_{episode} that could be used in the instantaneous power case, but with a different meaning: one would count the fraction of trials containing an oscillation that exceeds the duration and amplitude thresholds for every event-related point in time. Thus, this analysis would quantify the probability that the system is in an oscillatory episode at any given point in

[8] Note that the term ‘multiwavelet’ is used for two different methods. The method discussed here deals with wavelets that have the form of Slepian-like functions (Lilly and Park, 1995); another meaning of the term refers to wavelet bases in multiple dimensions (Goodman and Lee, 1993).

time across events. The properties of this proposed analysis are, however, beyond the scope of this paper.

4.4. Validity of simulation methods

The simulations in this paper were performed with background EEG created from a sum of sinusoids, similar to Yeung et al. (2004) and Zhan et al. (2006). Then, the to-be-detected signal, as a pure sinusoid, was added into the background. One may reasonably question the biological validity of this approach, since real EEG is likely to be created from a more stochastic process. An alternative approach would be to simulate the EEG from either an Ornstein-Uhlenbeck process (Steyn-Ross et al., 1999) or an autoregressive process (Kaipio and Karjalainen, 1997), and then add a signal from an additional stochastic process. However, this reduces our control over the exact amplitude, length, and frequency content of the signal. This could be remedied by filtering the signal to have $1/f^{\alpha}$ characteristics. We decided to use the Yeung et al. method because (a) it had already been used for EEG simulations before and (b) it made the signal (a sinusoid of known frequency, length and amplitude) quite congruent with the background activity (i.e., a collection of other sinusoids).

5. Conclusions

Based on the results presented here, we propose a few guidelines for the use of these oscillation detection methods.

(1) Using *both* P_{episode} and either multitapers or wavelets gives more information about the underlying signal than any method alone (Fig. 3). P_{episode} will be more sensitive to differences in length (see also Fig. 5), whereas the other two methods will confound length and amplitude of the signal. Among these two, wavelets will provide more sensitive detection (i.e., detecting lower amplitude signals), whereas multitapers will give better frequency specificity.

(2) If one is interested in detecting oscillations that exceed certain number of cycles, then P_{episode} is the method of choice. This method makes it harder to detect oscillations at lower frequencies. Short-duration, low-frequency oscillations are very difficult to distinguish from evoked activity of a non-oscillatory nature, and therefore P_{episode} will exclude those signals.

(3) The method of choice depends on the frequency one is interested in. For higher frequencies, P_{episode} is the more sensitive choice for shorter signals, whereas longer signals will be better detected by multitapers (Fig. 4). Also, multitapers offer better frequency localization than the other methods, provided the signal is strong enough.

(4) As with many data analysis problems, it is often useful to apply more than one technique to one's data. Our analyses suggest that when multiple oscillation detection methods are used in concert one can distinguish different types of oscillatory signals. For example, if oscillations are detected by wavelets, but not multitapers, at low frequencies, this would suggest that short-lived high-amplitude oscillations are present in the time series. Multitapers will show less diffuse peaks in the power spectrum than wavelets because of less frequency bleeding. Finally, P_{episode} can detect more subtle differences between two conditions, especially when these differences occur in the time domain. In addition, P_{episode} is more sensitive to short-duration high-frequency signals than are wavelets and multitapers, so long as the oscillatory signals exceed the duration threshold.

Acknowledgements

We would like to thank Jeremy Caplan and Joshua Jacobs for helpful discussions the manuscript. We also acknowledge Sridhar Raghavachari for help with the multitaper analyses. Finally, we thank two anonymous reviewers for their thoughtful comments. This work was supported by the CELEST grant SBE 0354378, Conte center grant P50 MH062196 and NIH grants RO1 MH061975, RO1 MH055687 to MJK and predoctoral NRSA MH72138 to PBS.

References

Bastiaansen M, Hagoort P. Event-induced theta responses as a window on the dynamics of memory. Cortex 2003;39:967–992.

Berger H. Übert das Elektroenkephalogramm des menschen (On the human electroencephalogram). Archiv Psychiatry Nervenkrankheiten 1929;87: 527–70.

Bronez TP. On the performance advantage of multitaper spectral analysis. IEEE Trans Signal Process 1992;40(12):2941–46.

Bruns A. Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches? J Neurosci Methods 2004;137:321–32.

Buzsáki G, Draguhn A. Neuronal oscillations in cortical networks. Science 2004;304(5679):1926–29.

Caplan JB, Madsen JR, Raghavachari S, Kahana MJ. Distinct patterns of brain oscillations underlie two basic parameters of human maze learning. J Neurophysiol 2001;86:368–80.

Caplan JB, Madsen JR, Schulze-Bonhage A, Aschenbrenner-Scheibe R, Newman EL, Kahana MJ. Human theta oscillations related to sensorimotor integration and spatial learning. J Neurosci 2003;23:4726–36.

Delorme A, Makeig S, Sejnowski T. Automatic artifact rejection for EEG data using high-order statistics and independent component analysis. Proceedings of the third international ICA conference. San Diego, December; 2001 [http://cloe.ucsd.edu/BSS-validation.html].

Ekstrom AD, Caplan J, Ho E, Shattuck K, Fried I, Kahana M. Human hippocampal theta activity during virtual navigation. Hippocampus 2005;15:881–89.

Engel A, Fries P, Singer W. Dynamic predictions: oscillations and synchrony in top-down processing. Nat Rev Neurosci 2001;2(10):704–16.

Farge M. Wavelet transforms and their applications to turbulence. Annu Rev Fluid Mech 1992;24:395–457.

Fell J, Klaver P, Lehnertz K, Grunwald T, Schaller C, Elger CE, et al. Human memory formation is accompanied by rhinal-hippocampal coupling and decoupling. Nat Neurosci 2001;4(12):1259–64.

Freeman WJ. Origin, structure, and role of background EEG activity. Part 4. Neural frame simulation. Clin Neurophysiol 2006;117(3):572–89.

Freeman WJ, Rogers LJ. Fine temporal resolution of analytic phase reveals episodic synchronization by state transitions in gamma EEGs. J Neurophysiol 2002;87(2):937–45.

Goodman TNT, Lee SL. Wavelets in wandering subspaces. Trans Am Math Soc 1993;338(2):639–754.

Hoogenboom N, Schoffelen JM, Oostenveld R, Parkes LM, Fries P. Localizing human visual gamma-band activity in frequency, time and space. NeuroImage 2006;29(3):764–73.

Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, et al. The empirical mode decomposition and the Hilbert spectrum from nonlinear and non-stationary time series analysis. Proc R Soc Lond A 1998;454:903–95.

Jensen O. Maintenance of multiple working memory items by temporal segmentation. Neuroscience 2006;139(1):237–49.

Kahana MJ. The cognitive correlates of human brain oscillations. J Neurosci 2006;26(6):1669–72.

Kahana MJ, Seelig D, Madsen JR. Theta returns. Curr Opin Neurobiol 2001;11:739–44.

Kaipio JP, Karjalainen PA. Simulation of nonstationary EEG. Biol Cybern 1997;76:349–56.

Karis D, Fabiani M, Donchin E. P300 and memory: individual differences in the Von Restorff effect. Cogn Psychol 1984;16:177–216.

Kemerait RC, Childers DG. Signal detection and extraction by Cepstrum techniques. IEEE Trans Inform Theory 1972;18(6):745–59.

Klimesch W, Schimke H, Schwaiger J. Episodic and semantic memory: an analysis in the EEG theta band. Electroencephalogr Clin Neurophysiol 1994;91:428–41.

Lilly JM, Park J. Multiwavelet spectral and polarization analyses of seismic records. Geophys J Int 1995;122(3):1001–21.

Linkenkaer-Hansen K, Nikouline VV, Palva JM, Ilmoniemi RJ. Long-range temporal correlations and scaling behavior in human brain oscillations. J Neurosci 2001;21(4):1370–77.

Mallat S. A wavelet tour of signal processing. San Diego: Academic Press; 1998.

Mitra PP, Pesaran B. Analysis of dynamic brain imaging data. Biophys J 1999;76:691–708.

Özdemir AK, Karakaş S, Çakmak ED, Tüfekçi DI, Arikan O. Time–frequency component analyser and its application to brain oscillatory activity. J Neurosci Methods 2005;145:107–25.

Paller KA, Wagner AD. Observing the transformation of experience into memory. Trends Cogn Sci 2002;6(2):93–102.

Percival DB, Walden AT. Spectral Analysis for Physical Applications. Cambridge, UK: Cambridge University Press; 1993.

Perrier V, Philipovitch T, Basdevant C. Wavelet spectra compared to Fourier spectra. J Math Phys 1995;36(3):1506–19.

Raghavachari S, Rizzuto DS, Caplan JB, Kirschen MP, Bourgeois B, Madsen JR, et al. Gating of human theta oscillations by a working memory task. J Neurosci 2001;21:3175–83.

Schiff SJ, Aldroubi A, Unser M, Sato S. Fast wavelet transformation of EEG. Electroencephalogr Clin Neurophysiol 1994;91:442–55.

Sederberg PB, Kahana MJ, Howard MW, Donner EJ, Madsen JR. Theta and gamma oscillations during encoding predict subsequent recall. J Neurosci 2003;23(34):10809–14.

Sederberg PB, Schulze-Bonhage A, Madsen JR, Bromfield EB, McCarthy DC, Brandt A, Tully MS, et al. Hippocampal and neocortical gamma oscillations predict memory formation in humans. Cerebral Cortex, doi:10.1093/cercor/bhl030, in press.

Sinkkonen J, Tiitinen H, Näätänen R. Gabor filters: an informative way for analysing event-related activity. J Neurosci Methods 1995;56:99–104.

Steyn-Ross ML, Steyn-Ross DA, Sleigh JW, Liley DTJ. Theoretical electroencephalogram stationary spectrum for a white-noise-driven cortex: evidence for a general anesthetic-induced phase transition. Phys Rev E 1999;60 (6):7299–311.

Thomson DJ. Spectrum estimation and harmonic analysis. Proc IEEE 1982;70:1055–96.

Tononi G, Edelman GM. Consciousness and complexity. Science 1998; 282(5395):1846–51.

Varela F, Lachaux JP, Rodriguez E, Martinerie J. The brainweb: phase synchronization and large-scale integration. Nat Rev Neurosci 2001;2(April(4):229–39.

Walden AT, Percival DB, McCoy EJ. Spectrum estimation by wavelet thresholding of multitaper estimates. IEEE Trans Signal Process 1998;46:3153–65.

Ward L. Synchronous neural oscillations and cognitive processes. Trends Cogn Sci 2003;7(12):553–9.

Wolansky T, Clement EA, Peters SR, Palczak MA, Dickson CT. Hippocampal slow oscillation: a novel EEG state and its coordination with ongoing neocortical activity. J Neurosci 2006;26(23):6213–29.

Yeung N, Bogacz R, Holroyd CB, Cohen JD. Detection of synchronized oscillations in the encephalogram: an evaluation of methods. Psychophysiology 2004;41(6):822–32.

Zanandrea A, Da Costa JM, Dutra SLG, Rosa RR, Saotome O. Spectral and polarization analysis of geomagnetic pulsations data using a multitaper method. Comput Geosci 2004;30(8):797–808.

Zhan Y, Halliday D, Jiang P, Liu X, Feng J. Detecting time-dependent coherence between non-stationary electrophysiological signals–a combined statistical and time–frequency approach. J Neurosci Methods 2006;156(1–2):322–32.