

RESEARCH

Open Access

Time-varying bispectral analysis of visually evoked multi-channel EEG

Vinod Chandran

Abstract

Theoretical foundations of higher order spectral analysis are revisited to examine the use of time-varying bicoherence on non-stationary signals using a classical short-time Fourier approach. A methodology is developed to apply this to evoked EEG responses where a stimulus-locked time reference is available. Short-time windowed ensembles of the response at the same offset from the reference are considered as ergodic cyclostationary processes within a non-stationary random process. Bicoherence can be estimated reliably with known levels at which it is significantly different from zero and can be tracked as a function of offset from the stimulus. When this methodology is applied to multi-channel EEG, it is possible to obtain information about phase synchronization at different regions of the brain as the neural response develops. The methodology is applied to analyze evoked EEG response to flash visual stimuli to the left and right eye separately. The EEG electrode array is segmented based on bicoherence evolution with time using the mean absolute difference as a measure of dissimilarity. Segment maps confirm the importance of the occipital region in visual processing and demonstrate a link between the frontal and occipital regions during the response. Maps are constructed using bicoherence at bifrequencies that include the alpha band frequency of 8Hz as well as 4 and 20Hz. Differences are observed between responses from the left eye and the right eye, and also between subjects. The methodology shows potential as a neurological functional imaging technique that can be further developed for diagnosis and monitoring using scalp EEG which is less invasive and less expensive than magnetic resonance imaging.

Keywords: EEG, HOS, Human vision, Bicoherence, Time-varying

Introduction

Bispectral analysis of EEG data has been the subject of a number of studies. Some have used single channel data only. Others have used multi-channel EEG ensembles but few have investigated *multi-channel EEG* using *higher order spectral analysis* in a *time-varying* manner. This article revisits the theoretical foundations to justify such analysis and provides new results from the application of time-varying bispectral analysis to evoked EEG responses.

Background

Research on bispectral analysis of EEG signals dates back to the 1970s, not long after higher order spectral analysis emerged as a branch of study in the 1960s. Dumermuth et al. [1] demonstrated that there exists significant phase locking between alpha and beta components in intracranial EEG. Barnett et al. [2] used bispectral analysis to

examine waking and sleeping states and found significant quadratic phase coupling only in the EEG of wakeful subjects with high alpha activity. These early studies used steady state potentials. Bullock et al. [3] used bicoherence analysis of intracranial and subdural EEG in a time-varying framework in an attempt to classify the onset of epileptic seizures. They analyzed EEG from sleep, wakefulness and seizure states. Their results were not conclusive on the effectiveness of the bicoherence descriptor. They found the bicoherence to fluctuate abruptly within a few seconds. The fluctuations were not consistent across subjects during the seizure period although statistically significantly higher levels of bicoherence were observed. Muthuswamy et al. [4] modeled paroxysmal burst EEG as a non-linear time-invariant process and showed that the bicoherence in the delta-theta band of EEG bursts is significantly higher than baseline waveforms in animal subjects recovering from a brain trauma. It has been shown that the bispectrum of the EEG correlates with changes in consciousness level and the bispectral

Correspondence: v.chandran@qut.edu.au
School of Electrical Engineering and Computer Science Queensland University of Technology, Brisbane, Qld 4001 Australia

index (BIS) [5,6] derived from EEG bispectral parameters was developed as a clinical tool to monitor depth of anaesthesia during surgery. Tang and Norcia [7] used the bispectrum to study steady state visually evoked potentials. They called their method the coherent bispectrum. They used oscillatory visual stimuli and reported the presence of inter-modulation frequencies and evidence of nonlinear interactions. Shen et al. [8,9] investigated time-varying bispectral analysis of non-stationary EEG data considering piece-wise third order stationary segments and non-Gaussian autoregressive modelling. Minfen et al. [10,11] used higher-order spectral analysis of EEG for classification of brain functional states.

In this study, time-varying bispectral analysis is applied to transient EEG responses evoked by a stimulus or related to a sensory event. A classical Fourier approach is adopted and ergodicity is only assumed over short time intervals around windows that are at the same offset with respect to a stimulus-locked time reference. Many previous EEG studies such as [12] have used stimulus-locked time references and stimulus-locked time averaging but most of them use grand averaging in time and some use spectral analysis. They have not investigated time-varying bispectral analysis in the manner described in this work. Bicoherence changes are tracked in this study with millisecond resolution, better tracking resolution than in earlier studies such as [1]. Auto-bicoherence is mapped simultaneously for multiple channels to obtain a spatio-temporal view of the EEG response at selected locations in the bifrequency plane, providing enhanced processing and visualization capabilities compared to any previously reported work. Such analysis will be useful in understanding the neuronal activity involved in visual and auditory perception, motor planning and movements. It can provide new features for diagnosing neurological conditions and sensory impairment.

Methods

In this section, some equations defining higher order spectra [13-16], are revisited to provide a context and justification for the adoption of a classical short-time Fourier approach to time-varying bispectral analysis.

Random processes

Consider a real-valued random process, $x(t)$, that varies with time, as a signal model for any channel of EEG. An ensemble of many realizations of the random process can be used to define statistical averages or expected values that are deterministic quantities. At any given time instant t' , $x(t')$ is a random variable. For a first-order stationary random process, the probability density function $p[x(t')]$ is independent of the time t' . Descriptions of the random process that depend only on the statistics of one random variable such as the mean value are examples of first order

statistics. The mean or first order moment of the process is

$$E[x(t')] = \int x(t')p[x(t')] dx = m_1^x \quad (1)$$

A random process is not fully characterized by its first order statistics alone. The joint probability density function $p[x(t'), x(t'')]$ provides a second order description of the random process. For a second-order stationary process, this probability density is independent of the absolute value of the time instants and depends only on the time offset $\tau = t'' - t'$. The autocorrelation or second order moment of the random process is

$$m_2^{xx}(\tau) = E[x(t')x(t'')] = E[x(t')x(t' + \tau)] \quad (2)$$

Here E stands for the expectation operation over an ensemble of realizations of the process. Very often, only a single realization is available. If it is sufficiently long and the process is stationary to the second order, an estimate of the autocorrelation can be computed by averaging over time rather than over the ensemble. It is given by

$$\tilde{m}_2^{xx}(\tau) = E_t[x(t')x(t'')] = \int x(t')x(t' + \tau)dt' \quad (3)$$

For an ergodic process ensemble statistics (E) are equal to time statistics (E_t). Ergodicity also implies that time statistics do not change with time and an ergodic process is necessarily stationary. If these properties hold true up to n -th order statistics, the process is said to be n -th order ergodic. Ergodicity is not guaranteed for all processes. At best it is an assumption that holds fairly well in practice to allow reliable estimates of statistical parameters that characterize the process. If the process is ergodic, a single long realization may be divided into several shorter ones for statistical expectation computation. This division into blocks of time creates an ensemble of shorter realizations of the random process. As a trade-off, the range of possible time offset values (τ) is reduced. Assume that an estimate of the autocorrelation, $\tilde{m}_2^{xx}(\tau)$ has thus been obtained. In practice, this autocorrelation will usually tend towards zero for large offsets and the block size can be suitably chosen to be large enough for the autocorrelation to have decayed to nearly zero. If that is the case, the autocorrelation function will be absolutely integrable and its Fourier transform will exist. The Fourier transform of the autocorrelation function is the power spectral density referred to as the power spectrum of the process, $S_2^{xx}(f)$, where f represents frequency in cycles per second or Hertz when the independent variable time of the random process and the offset τ are measured in seconds. For deeper understanding of stationarity and ergodicity in random processes and stochastic calculus the reader is referred to [17]. The power spectrum reveals the harmonic structure or frequency components in the random

process. If the autocorrelation function is an impulse or Dirac delta function $\delta(\tau)$ (implying that values of the process separated even infinitesimally in time are uncorrelated) the power spectrum $S_2^{xx}(f) = 1$ is constant and the process is referred to as white. In this case, the power spectrum does not reveal much about the process. Noise processes like that produced by thermal fluctuations of electrons in a material tend to be broadband in their spectrum and nearly white. There are no selectively resonant structures that would concentrate power over particular frequencies or frequency bands. If the process is also a result of superposition of many uncorrelated processes, such as the potential difference produced across a resistor owing to the thermal fluctuations of millions of electrons, the Central limit theorem dictates that the probability density function will tend towards the Gaussian distribution. This is also true of joint probability density functions of all orders. The process is then referred to as a Gaussian process. A Gaussian probability density function in D dimensions is given by

$$p(X) = \frac{1}{(2\pi)^{D/2} [\det(C_X)]^{D/2}} \times \exp \left[-\frac{1}{2} (X - \mu_X) C_X^{-1} (X - \mu_X)^T \right] \quad (4)$$

where μ_X is the mean vector and C_X is the covariance matrix of the D -dimensional vector X . \det represents the determinant of a matrix. For D -th order statistics of a random process, these dimensions come from D time instants at which the process is considered, each a scalar random variable. Yet, only first order (mean value) and second order (pair-wise in the covariance matrix) suffice for a complete description of the Gaussian probability density. The white Gaussian random process model is often employed as a simplified model for noise. A Gaussian process is completely described by first and second order statistics because a Gaussian probability density function of any order can be represented using a mean vector and a covariance matrix alone. Deviations from Gaussian distributions can be detected or measured by using higher than second order moments. For example, skewness is computed from third moments and kurtosis from fourth order moments. Skewness is a measure of the asymmetry of the probability density and kurtosis is an indicator of the deviation of the tails of the distribution from those of a Gaussian. If a deviation from Gaussian is detected, it may be owing to particular frequency components in the process.

A Fourier representation for higher order moments was proposed by Shiryayev [13]. The third order spectrum or bispectrum was first applied to study nonlinear interactions in ocean waves [18]. The estimation of such higher

order spectra was placed on a firm mathematical and statistical foundation by Brillinger and Rosenblatt [14,15,19] in the same decade. The fast Fourier transform (FFT) algorithm and advances in computers facilitated the numerical estimation of higher order spectra and gave impetus to higher order spectral analysis research and its application in the latter half of the twentieth century. More review articles can be found in [16,20,21].

The bispectrum

The third order moment or triple correlation of a random process $x(t)$, assumed stationary, is defined as

$$m_3^{xxx}(\tau_1, \tau_2) = E[x(t)x(t + \tau_1)x(t + \tau_2)] \quad (5)$$

For many processes this function of two offset variables will decay with increasing delay or offset. If this function satisfies the Dirichlet conditions for existence of a Fourier spectrum, such as being absolutely integrable, its Fourier transform may be defined and is known as the bispectrum of the process,

$$B(f_1, f_2) = S_3^{xxx}(f_1, f_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} m_3^{xxx}(\tau_1, \tau_2) \times e^{-j2\pi(f_1\tau_1 + f_2\tau_2)} d\tau_1 d\tau_2 \quad (6)$$

The subscript 3 refers to a third order parameter. The superscripts xxx are a reminder that the spectrum is an auto-spectrum computed from the same random process rather than a cross-spectrum computed from different random processes. The bispectrum of a one dimensional random process is a function of two frequencies as shown above. The pair of frequencies is referred to as a bifrequency. The bispectrum of a Gaussian random process is zero because the third moment is zero. Moments of higher order are not all zero for a Gaussian process. But cumulants [16,21] of order three and greater are zero and it is conventional to define higher order spectra in terms of cumulants of a random process rather than moments. This work is restricted to the third order and third order moment and cumulant spectra are identical, both referring to the same bispectrum. This definition of the bispectrum may be extended to include moment functions from harmonic random processes that are not absolutely integrable but have finite power, or finite energy within a period, similar to such extension for the power spectrum. The spectrum will then theoretically comprise of impulses or Dirac delta functions rather than being continuous.

Estimation of the bispectrum

It has been shown that the bispectrum can be estimated in the Fourier domain [19]. Strictly speaking, a stationary random process does not have a Fourier transform

because existence conditions are not satisfied. Any single realization of the process will have finite power but infinite energy. A Fourier transform is therefore defined using the Fourier–Stieltjes integral and Cramer spectral representations as

$$x(t) = \int_{-\infty}^{\infty} e^{j2\pi ft} dX(f) \quad (7)$$

where

$$E[dX(f)] = 0 \quad (8)$$

It can be shown [16] that for real-valued $x(t)$

$$m_2^{xx}(\tau) = \begin{cases} \int_{-\infty}^{\infty} e^{j2\pi(f\tau)} E[dX(f)dX^*(f)] & f_1 = f_2 = f \\ 0 & f_1 \neq f_2 \end{cases} \quad (9)$$

where $*$ represents the complex conjugate operator. The power spectrum of a random process can thus be expressed as an expected value in the frequency domain.

$$S_2^{xx}(f)df = E[dX(f)dX^*(f)] \quad (10)$$

The power over a small interval in the frequency f is the expected value of a product of Fourier–Stieltjes spectral representations that are complex conjugates. The power spectrum is real-valued and does not have any phase information. The zero value for $f_1 \neq f_2$ can be subjected to some scrutiny. If it were not the case, the second moment, m_2^{xx} would turn out to be a function of t which is a contradiction to the assumption of a stationary random process, $x(t)$. However, it holds because the infinitesimal bandwidth processes, $dX(f_1)$ and $dX(f_2)$ are complex-valued and in a phasor representation such as Rice's representation [17] they are not “phase-locked” when $f_1 \neq f_2$. The phase of the product will be uniformly distributed in the interval $[0, 2\pi)$ and the expected value of the product will be zero in this case. When $f_1 = f_2 = f$, the product becomes real-valued and can have a non-zero expected value. A different but formal proof is given in [22] where it is shown that $dX(f)$ is an orthogonal process. The proof starts by making a single realization of the random process periodic outside a finite interval. In the same manner it can be shown [16] that the bispectrum is a triple product of such representations.

$$m_3^{xxx}(\tau_1, \tau_2) = \begin{cases} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{j2\pi(f_1\tau_1 + f_2\tau_2)} E[dX(f_1)dX(f_2) \\ \times dX^*(f_1 + f_2)] & f_1 + f_2 = f_3 \\ 0 & f_1 + f_2 \neq f_3 \end{cases} \quad (11)$$

and by comparing with Equation (6) and the inverse Fourier relationship,

$$S_3^{xxx}(f_1, f_2)df_1df_2 = E[dX(f_1)dX(f_2)dX^*(f_1 + f_2)] \quad (12)$$

The reason for a non-zero expected value when the resonance condition, $f_1 + f_2 = f_3$ is satisfied is similar to that for the power spectrum as explained above except that the triple product here can become “phase-locked” and does not have to be real-valued. The term “phase-locked” is used with inverted commas because the phase is not necessarily constant in every realization. The bispectrum is a complex-valued function of two frequency variables and retains relative phase information between Fourier components.

A difficulty in the use of the frequency domain for estimation of the power spectrum or the bispectrum of a random process lies in the fact that differentiability of the spectral representation cannot be guaranteed. However, estimates of higher order statistical or spectral parameters have to be computed numerically for practical purposes.

Sampling of random processes

Sampling of stationary random processes is discussed in [17] where Rice's representation for stationary random processes in terms of in-phase and quadrature components is utilized in a theoretically rigorous treatment. A discussion is also found in [23]. For the sake of simplicity a descriptive justification is provided here. Assume that the random process has a finite bandwidth or has been filtered using a linear phase filter that does not disturb the higher order spectral parameters of interest and is thus made band limited. The process can now be sampled at the Nyquist rate of twice the bandwidth without any loss of information. It is now a discrete-time random process whose power spectrum and higher order spectra of interest are still the same within the relevant band of frequencies. Criteria to prevent aliasing in higher order frequency domains are not exactly identical as that for the power spectrum because of the multiple frequency dimensions, and the interested reader is referred to [24–26]. Here, attention is focussed only in the regions in higher order frequency space where the conditions imposed by the Nyquist criterion hold and suffice. Sampling is assumed to be without any aliasing in the power spectrum or bispectrum in the range of frequencies of interest. Sampling also imposes a limit on the resolution with which statistical parameters can be estimated in the time domain because the offset variables cannot be smaller than the sampling interval. Assume further that higher order statistical parameters of interest, such as the third moment, can be estimated to the desired accuracy using expected values, in practice, averaging over finite length epochs. Theoretically, these parameters are defined in the limit of the epoch length approaching infinity. In practice, a long enough epoch length will ensure that the estimates converge to a desired degree for processes whose moments converge asymptotically. Harmonic random processes have ensemble mean and

autocorrelation functions that are periodic in time and are a class of cyclostationary random processes [17]. For cyclostationary processes higher order spectral parameters do not asymptotically converge as the time goes to infinity. Correlations for such processes are cyclic and do not decay to zero. Theoretically rigorous treatment of the calculus of various classes of random processes can be found in [17]. If a cyclostationary process were divided into sub-processes using blocks that are multiples of a period, the random-phase assumption for individual harmonic components will not hold. This can be addressed by phase randomization procedures, for example by using blocks with random shifts. A phase-randomized [17] version of the cyclostationary process is a stationary process with the same moments as that computed for the cyclostationary process over one period. There are pitfalls in the phase randomization approach pointed out in [27] and ergodicity can sometimes be destroyed. For estimation of the bispectrum, it is assumed here that an ensemble of realizations is available from a number of trials and phase randomization is not necessary.

Let there be N epochs. Let each epoch be processed using blocks of M samples, appropriately large such that estimates in time are averaged over a large number of periods and the inaccuracy from averaging over a non-integral number of periods is small. A finite length block of samples satisfies the conditions for existence of a discrete-time Fourier transform (DTFT)

$$X(f) = \sum_{k=0}^{M-1} x[kT] e^{-j2\pi kfT} = \sum_{k=-\infty}^{\infty} x[kT] \Pi[k/M] e^{-j2\pi kfT} \quad (13)$$

where the rectangular window in time is given by

$$\Pi[k/M] = \begin{cases} 1 & 0 \leq k < M \\ 0 & k \geq M \end{cases} \quad (14)$$

and T is the sampling interval. The DTFT can be computed at any desired frequency and uniquely between 0 and $1/T$. With higher order spectral analysis, it is not merely a question of reducing spectral leakage from windowing but understanding the effect that leakage will have on estimates of relevant parameters. Spectral leakage and windowing in the estimation of the bispectrum are discussed in [28,29]. In general, spectral leakage from statistically independent or random phase components will have a similar effect on higher order spectra as the addition of white random noise. It will lower the fraction of power that is phase-coupled. Tapered windows will introduce a modulation effect and spurious phase-locked low frequency components unless the mean value is removed from each block before application of the window and computation of the discrete Fourier transform

(DFT). Assume that each realization of the discrete-time random process repeats itself outside the M samples. It is then converted into a hypothetical cyclostationary process which has nearly the same higher order moments of a given order provided M is sufficiently large. This assumption is different from the harmonizability of deterministic moment functions [21]. The FFT algorithm to compute the DFT can now be employed to get samples of the spectrum as the expected value.

$$X(p) = X(f) \Big|_{k=\frac{p}{T}} = \sum_{k=0}^{M-1} x[p] e^{-j2\pi kp} \quad (15)$$

Bispectrum estimate in the Fourier domain

The bispectrum of the process can then be obtained in the Fourier domain as the expected value of a triple product of Fourier coefficients using the DTFT or the DFT. It can be shown [16] that

$$E[X(f_1)X(f_2)X^*(f_1+f_2)] = MS_3^{xxx}(f_1, f_2) \quad (16)$$

Periodicity outside the M -length interval and third order stationarity are utilized in the simplification. The bispectrum can be therefore estimated as an average of a triple product of Fourier coefficients over an ensemble of realizations. It can be seen from Equation (12) that the bispectral density is actually a triple product divided by the term $df_1 df_2 = (\frac{1}{MT})^2$ but since this is a constant for a given sampling interval and block length it may be ignored when bispectral values are compared and M and T remain constant. It can also be removed when the bispectrum is normalized with power spectrum values at the frequency components involved as it will appear both in the numerator and the denominator.

For the rest of the article, $B(f_1, f_2)$ is used instead of $S_3^{xxx}(f_1, f_2)$ to refer the the auto-bispectrum of $x(t)$ and $P(f)$ is used instead of $S_2^{xx}(f)$ to refer to the auto power spectrum.

The bispectrum of a real-valued process satisfies a number of symmetry properties in the bi-frequency domain. Assume that time is expressed in sampling intervals and frequencies are normalized by the sampling frequency. The principal domain or non-redundant region [16,30] of computation of the bispectrum for such a process is given by the triangular region (shown in Figure 1) in bi-frequency defined by

$$f_1 \geq 0; f_1 = f_2; f_1 + f_2 \leq 1 \quad (17)$$

Bispectrum magnitude and bicoherence

Expressing the bispectrum estimate in polar form

$$B(f_1, f_2) = E[|X(f_1)X(f_2)X^*(f_1+f_2)|e^{j\phi(f_1)+\phi(f_2)-\phi(f_1+f_2)}] \quad (18)$$

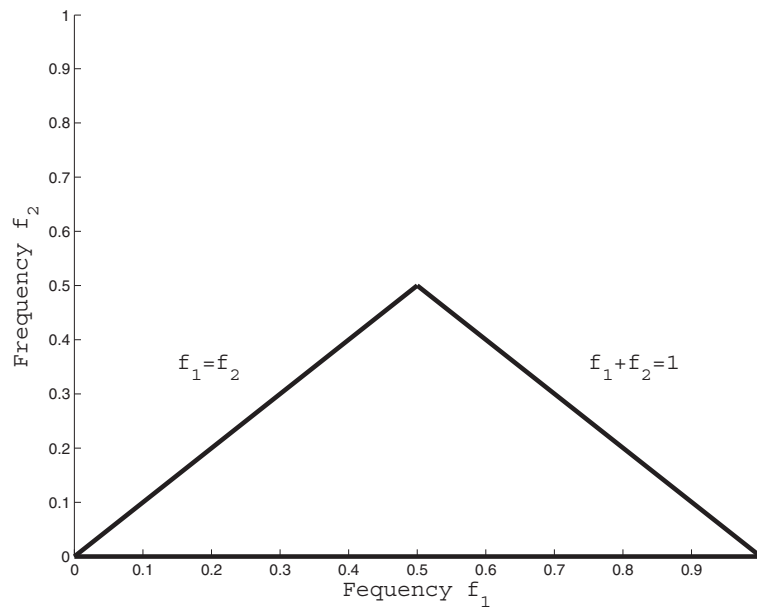


Figure 1 The principal domain of the auto-bispectrum of a sampled signal. Frequencies are normalized by one-half of the sampling frequency.

If the phases, $\phi(f_1)$, $\phi(f_2)$, $\phi(f_1 + f_2)$ are independent and uniform random in the interval $[0, 2\pi)$ the bispectrum will be zero. If there is perfect phase coupling between the Fourier components and $\phi(f_1) + \phi(f_2) - \phi(f_1 + f_2)$ is zero for every realization, the bispectrum will be non-zero. The bispectrum has been used to study non-linear wave coupling and in this context it has been normalized to assume values between 0 and 1 similar to coherence in second order statistics. The squared magnitude of the normalized bispectrum [31] is referred to as bicoherence by

$$b^2(f_1, f_2) = \frac{|B(f_1, f_2)|^2}{P(f_1)P(f_2)P(f_1 + f_2)} \quad (19)$$

Although this bicoherence spectrum measures the degree of phase coupling at the particular bi-frequency, it is not guaranteed to be between 0 and 1 when numerically estimated from a finite number of realizations. An alternative normalization, that satisfies this property better is given in [32] by

$$b^2(f_1, f_2) = \frac{|B(f_1, f_2)|^2}{E[|X(f_1)X(f_2)|^2] E[|X^*(f_1 + f_2)|^2]} \quad (20)$$

The bicoherence is a measure of the fraction of power at the component $f_1 + f_2$ that is owing to phase-coupled Fourier components at f_1 and f_2 as opposed to arising from random-phase components, or from random additive broadband noise that has non-zero power spectral density at these frequencies. If there is a random phase component at any of the frequencies f_1 , f_2 or $f_1 + f_2$ the bicoherence will reduce because at least one of the denominator terms in Equation (20) will increase. Because

the bicoherence is a ratio it is sensitive to small values of the denominator that are close to zero. When the denominator is zero, the ratio should actually be 0/0 but can be a large value subject to precision in the numerical computations. This will rarely occur with broad band random processes but can occur with harmonic random processes that have line spectra. One means of avoiding this is to add white Gaussian random noise of small amplitude to the input prior to processing. This will have the undesirable effect of lowering all bicoherence values from their true value. If the power of the noise is small compared to the power of sinusoidal components at the frequencies of interest its negative bias on the true value of bicoherence will be small and negligible.

Estimation of bicoherence

The procedure to estimate bicoherence using the Equation (20) is as follows.

- (1) Collect N epochs in an ensemble. Let the epochs be processed as M sample blocks. The sampling frequency, f_s , should be above twice the highest frequency of interest. An anti-aliasing filter may be applied provided it does not disturb phase relationships in the range of frequencies of interest. The frequency resolution is set by the length of each block, M , and given by $\Delta f = \frac{1}{MT}$ where $T = \frac{1}{f_s}$ is the sampling interval.
- (2) Estimate the signal power and add low amplitude white Gaussian noise keeping the signal to noise ratio maintained high.

- (3) For each block, remove the mean value.
- (4) Obtain the Fourier coefficients using the FFT.
- (5) Form the products of Fourier coefficients required, $X(f_1)X(f_2)X^*(f_1 + f_2)$ and $X(f_1)X(f_2)$, and retain $X^*(f_1 + f_2)$. Do this for every bi-frequency in the non-redundant region of computation.
- (6) Average the intermediate products above over all N realizations.
- (7) Divide as shown in the Equation (20) to obtain an estimate of the bicoherence.

Statistical reliability of bicoherence estimates

In general, the estimation of higher order spectra becomes progressively unreliable as the order increases. Increasingly larger numbers of epochs are required to achieve estimates with similar variance. An asymptotic theory of these estimates is discussed in [19]. The bias and variance of any higher order spectrum will actually depend on the true value and exact expressions are not known for arbitrary true values [33].

The estimate of the bicoherence from a finite number of realizations has a bias and variance and it must be determined whether the value is statistically significantly different from zero. For a finite value of N , it has been shown that the bicoherence of a Gaussian process is Chi-squared distributed [16]. 95% of the bicoherence values are expected to be below $\frac{3}{N}$. Statistics of the higher order coherence are discussed in [33,34].

The development of higher order spectral analysis above assumes that the process is ergodic. This is not satisfied by the EEG signal. Statistics may be stationary over a short window in time but ergodicity may still not be satisfied. A framework for using short-time higher order spectral analysis is developed next.

Non-stationary random processes and time-varying spectra

Extensions of the theory of sampling to non-stationary random processes have been made by Gardner [27,35], Garcia et al. [36] and others. Higher order spectra have been analyzed in time-varying frameworks using Wigner-Ville distributions and polynomial phase modelling in [37-42]. Essentially, Wigner-Ville distributions transform the second order correlation with respect to the time offset variable τ but leave the time variable t' in Equation (2) to take the non-stationarity of the signal into account. A symmetric form $x(t' - \tau/2)x(t' + \tau/2)$ is used in the definition of the autocorrelation. Higher order forms have been defined with multiple time offset variables. In order for ergodicity to hold, it is necessary that the time t' refer to the same time instant in every realization of the non-stationary process. This is only possible if the process is time-locked to the reference $t' = 0$ such that there is correspondence of this time in every realization of the

process with regard to the non-stationarity. This is not true of arbitrary non-stationary random processes but it can be imposed on data acquisition as done in this work. Phase is meaningless in the Wigner-Ville approach unless the cross Wigner-Ville distribution is used to access the phase (see [43]). Better resolution in the time-frequency plane is possible with the Wigner-Ville approach than with classical short-time Fourier analysis.

Wavelet transforms have also been used in defining transformed representations for higher order cumulants [44]. It may be noted here that in general the wavelet transform is quite well suited to capture transient information in the time-varying scalogram but not all mother wavelets lend themselves to representations that involve phase or utilize correlations. A real-valued mother wavelet is usually defined to be symmetric and does not have phase. Morlet wavelets (discussed in [45]) are complex-valued and are in fact complex exponentials at a central frequency modulated by a Gaussian window in time. Retaining both the real and imaginary parts of the transform, a phase can be defined as a function of the central frequency and time. It is in this respect similar to a short-time Fourier transform (STFT) with a Gaussian window. However, the approach is not well suited for phase coupling investigations because expectation by averaging over short time windows is not the same as ensemble averaging and triads of frequencies that satisfy resonance conditions for sum or difference interactions are not as easily identified in scale space as in a linear frequency space. A wavelet transform based bicoherence is developed in [45]. This work reported that lack of ensemble averaging resulted in spurious coherence. Filtering and application of thresholds based on global maxima are used to remove unwanted noise and extract evolutionary characteristics from the bicoherence.

Another approach to time-varying bispectral analysis of non-stationary EEG signals is presented in [9]. A non-stationary signal is modelled using a non-Gaussian autoregressive model with time-varying parameters under the assumption of piece-wise stationarity. This approach is suitable for analysis of different states but not particularly for tracking evolutionary characteristics with time.

A classical STFT approach with a sliding window is presented here. This methodology is preferable here over other methods because it provides a direct phase coupling interpretation of the results and permits relatively easy means of establishing a confidence interval to indicate whether the bicoherence observed is significantly different from zero or not. Further, the methodology developed here examines EEG frequency components in the delta, alpha and beta bands considering narrow band frequencies of interest such as 4, 8 or 20Hz. The STFT is an appropriate tool for such analysis unlike broadband analysis.

Short time Fourier analysis

The Fourier transform defined as an infinite integral for continuous signals or an infinite summation (DTFT) for discrete-time signals is not a useful tool for the analysis of non-stationary signals. Spectral representations of non-stationary signals change with time. A short-time Fourier spectrum may be computed using a window in time centred at a given time instant. As a function of time, this spectrum provides information about the signal in a time-varying manner. The squared modulus of the STFT is known as the spectrogram. The main disadvantage of the STFT is that the frequency representation is uniformly sampled. The frequency bins of the same size regardless of whether high frequencies or low frequencies are being analysed. A logarithmic spacing of frequency, with the frequency step size proportional to the frequency, would permit better analysis of most real-world signals. The time-frequency bins of the STFT are rectangular. The main advantage is that it can be computed fast using the FFT algorithm. In order to keep the processing simple and to facilitate the establishment of values that are significantly different from zero, DFT using a sliding window is adopted as the method of analysis for non-stationary random processes here. It is assumed that the statistics of the process are stationary up to the third order within a window and change slowly as the window is translated. The process is also assumed to be ergodic provided a correspondence between windows is established across multiple epochs. The correspondence across realizations can be achieved by using a reference time that is constant for all epochs with regard to the physical processes and state changes that are taking place. For evoked EEG this reference is the time of application of the stimulus. Segments of each epoch that are equally offset in time from the reference can be assumed to have similar statistics and constitute an ensemble of epochs.

Time-varying bicoherence

Assume that N epochs of a nonstationary random process are obtained along with a separate signal that provides a stimulus-lock time reference. Assume that the epochs are sampled and the sampling interval is T . The sample at the reference time is indexed 0. The epochs extend from $-M_1$ to $+M_2$ samples with respect to the reference. The objective is to examine the bicoherence of the process as a function of time around the reference. This can provide useful information about the process prior to, during and after the stimulus indicating the degree of phase-coupling between harmonic components in the process. A sliding window of $W = M$ samples is used and the window is centred at t_w . If the window is short enough compared to the transient changes in the underlying process, an assumption of stationarity within the window will hold. The ensemble of epochs for any given t_w can be assumed

ergodic if the process can be assumed to behave similarly in each epoch at a given offset t_w from the stimulus-locked reference. A time-varying version of the bispectrum can then be defined using Fourier transform coefficients with the STFT centred at time t_w ,

$$B(f_1, f_2, t_w) = E[X(f_1, t_w)X(f_2, t_w)X^*(f_1 + f_2, t_w)] \quad (21)$$

and a time-varying bicoherence can be defined using

$$b^2(f_1, f_2, t_w) = \frac{|B(f_1, f_2, t_w)|^2}{E[|X(f_1, t_w)X(f_2, t_w)|^2]E[|X^*(f_1 + f_2, t_w)|^2]} \quad (22)$$

Estimation of time-varying bicoherence

The procedure to estimate time-varying bicoherence is as follows.

- (1) Vary the position of the centre of the window from $-M'_1 + W/2$ to $M'_2 - W/2$ skipping S samples between windows. For each position, consider the ensemble of windowed segments as a separate random process.
- (2) Estimate the bicoherence for this process.

Bicoherence can be visualized in the bi-frequency plane as an image. Time varying bicoherence, therefore, will need to be visualized as a montage of images or as an animated set of frames. A contour map with contours starting from the 95% significance level value is useful. Such a map is shown in Figure 2. For a multi-channel signal such as an EEG montage, topographic maps that show the parameter of interest simultaneously at all channels is often employed. They provide useful information about relative values and the spatial distribution. Such a map is shown in Figure 3.

Only one value of bicoherence at each channel can be visualized in this manner at a time from one plot. Visualization of the time evolution requires the display of a sequence of two-dimensional plots.

Application of time-varying HOS to evoked EEG response

Higher order spectra, as spectral representations of cumulants or moments, are strictly valid only for ergodic random processes. These requirements are not satisfied by evoked EEG potentials. However, if a sufficiently small observation window in time is employed, the process can be assumed to be stationary within the window. The time of application of the stimulus serves as a reference. Windows at a given offset from this reference can be assumed to yield consistent averages over an ensemble if the speed of response does not vary. For spontaneous responses such as reaction to an auditory or visual stimulus this can be expected. For learned responses such as memory recall it may not hold true.

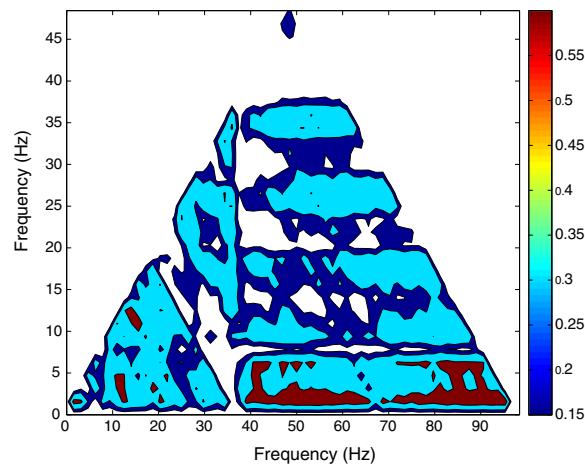


Figure 2 Bicoherence of visual evoked EEG. Bicoherence at channel O2 plotted over the principal domain in bifrequency space computed using a 128 point window around a point in time 10ms after the presentation of a flash visual stimulus to the right eye. The sampling frequency is 200Hz. Only Bicoherence values above the 95% significance level of 0.15 are coloured and the colour bar is shown on the right. It can be noted that bicoherence is significantly high over many frequency combinations and the values appearing as brown are particularly high around 0.6. This picture can change with time.

Let $x_{c,W,t_W}(t)$ be an evoked EEG response signal from channel c over a window W centred on time t_W at time t relative to the stimulus. W is short enough for the process to be considered stationary. An array of EEG channels is processed simultaneously. Bicoherence values are obtained for each channel as a function of time t_W . For the sake of simplicity of notation, bicoherence will be referred to as $B(f_1, f_2)$ and with the channel names mentioned on the plots or tables rather than included as a subscript. Time will implicitly refer to t_W and the subscript W is omitted as the window does not change except for sliding across in time.

Assume that each epoch of this process extends periodically outside the window with the same statistics up to third order. Short time higher order spectra are the computed using ensemble averaging. If N epochs are used in the ensemble average, bicoherence values for white Gaussian noise are Chi-squared distributed and the 95% confidence level is $\frac{3}{N}$ [33,34].

Response to a given stimulus is recorded several times in different epochs. In practice, the assumption of ergodicity over a window will only hold if the response is at the same speed. This can be expected for most involuntary responses such as those to a visual stimulus or an auditory stimulus from the same individual or a set of individuals who are similar in this regard. For responses such as memory-recall, the time delays can vary considerably and the stimulus-locked ergodicity assumption may not hold even for the same individual. For a voluntary movement such as a hand squeeze, the assumption will hold provided each subject complies with the request to perform the movement as instructed. The time of application

of the stimulus is recorded synchronously with the EEG channels using an auxiliary channel. The length of the window, MT , should be long enough to get the desired frequency resolution $\Delta f = \frac{1}{MT}$ and yet short enough for stationarity up to order 3 for bispectral analysis to hold true. This can be a challenge. Exactly how quickly, phase synchronization and desynchronization occurs between frequency components in the response of neural populations to stimuli is not fully known. Whether such phenomena last tens of cycles or hundreds of cycles of the component frequencies and whether they are consistent and correlated with other known markers of such responses is not yet clearly understood. Experimental analysis with real EEG data is a means to provide answers to such questions.

A phase coupling interpretation of bicoherence will strictly hold only for rhythmic components of the EEG such as alpha and beta waves. Transient or paroxysmal events such as bursts will result in high bicoherence if they occur at nearly the same offset from the stimulus onset and retain their waveform characteristics across an ensemble. In this case, they will always fall within some processing window(s) and they will be phase synchronized to the envelope signal that defines the burst.

Simulations

In order to demonstrate the usefulness of the time-varying bicoherence analysis with phase synchronized waves and bursts in the context of EEG signal processing, three simulations are performed. In each case, 20 realizations of a signal are generated with sampling frequency 200Hz and length 128 samples extending from $t = 0$ to $t =$

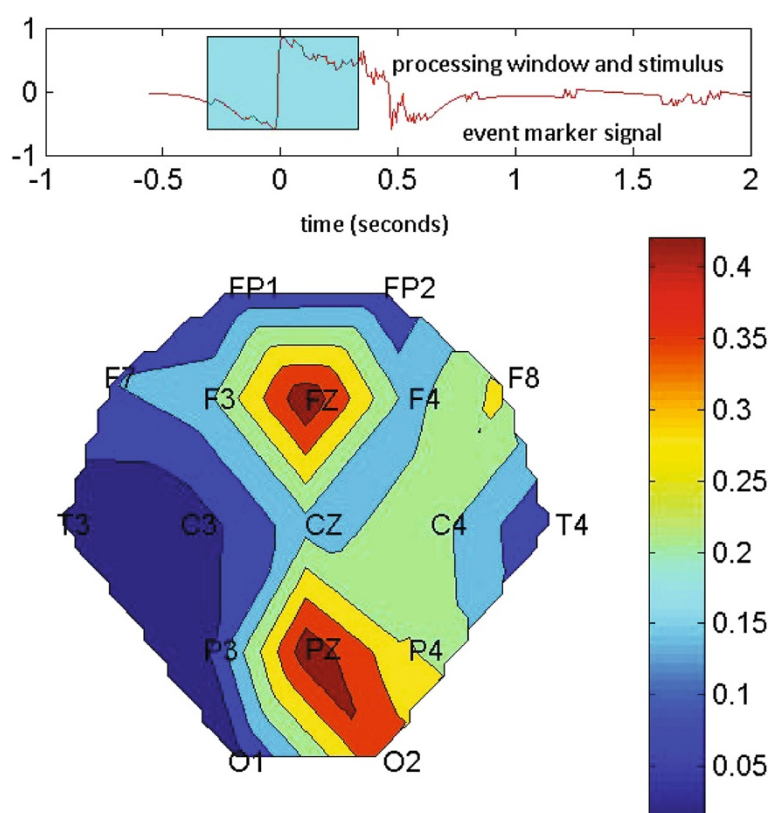


Figure 3 Bicoherence multiple channel map. Bicoherence at (8, 8Hz) at each of 17 EEG electrodes as a spatial distribution on the scalp. A colour bar is shown on the right and values above 0.15 are statistically significantly above zero bicoherence. Bicoherence is computed using a 128 point window which extends in time over the blue region as shown on the 1D plot on top and the window is centred 10ms after the presentation of a flash visual stimulus to the right eye. The time of presentation of the stimulus is indicated by the transition in the 1D plot and this is used as the zero time reference in each epoch processed.

0.635s. Each realization is a rectangular windowed block centred at time $t_w = 0.315$ s of a random signal in the time domain. Three types of input signal are simulated: (a) phase-coupled and random phase sinusoids, (b) alpha burst with jitter and (c) beta burst with jitter.

Phase-coupled and random phase sinusoids

The input signal comprised of sinusoids at 8, 16, 20, and 40Hz. The sinusoid at 16Hz had a phase which was the sum of the phases of two sinusoids at 8Hz in each realization. The sinusoids at 8Hz had a random individual phase in each realization. A similar triad of two components at 20 and 40Hz, however, was random phase and had individual phases that were random in each realization and there was no relationship between their three phases. The amplitude of each sinusoid was the same and constant, equal to 1. The time domain waveform for a single realization of this signal is shown in Figure 4a. Gaussian random noise was added to each realization with 10dB SNR. Twenty realizations were generated. A power spectrum estimate for the signal using an averaged periodogram is shown in Figure 4b. As expected there are

peaks in the power spectrum at 8, 16, 20, and 40Hz, and the peaks at 8 and 20 Hz are approximately twice the other because there are two random phase components at each of these frequencies. The bicoherence for this signal is shown in Figure 4c. Note the high value of bicoherence at (8Hz, 8Hz) because of the phase-coupled triad of components (8, 8, and 16Hz). The bicoherence at (20, 20) is however low because the triad of Fourier components (20, 20, and 40Hz) is not phase-coupled. How do we know whether a value is low or high? If 20 realizations are averaged in the estimate, the 95% significance level for bicoherence (see Table 1) is 0.15. This implies that if the input were white Gaussian noise, for which theoretically the bicoherence is 0, the distribution of bicoherence [34] is such that around 95% of the values will be less than 0.15. Therefore, a value above 0.15 can be considered non-zero with 95% confidence. Why is the bicoherence not equal to 1 at a perfectly phase-coupled triad? The bicoherence is lowered by the presence of random phase components at each of the frequencies involved arising from additive random noise. Why would a triad of Fourier components in a real world signal be phase-coupled? When an input

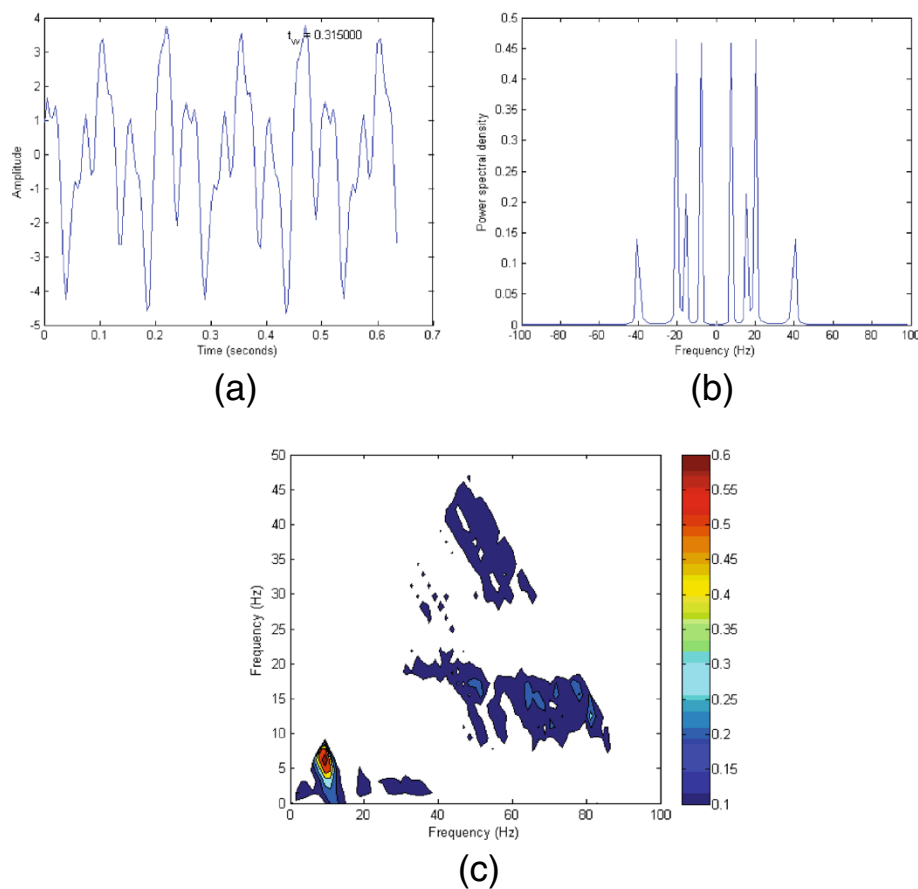


Figure 4 Phase-coupled and random phase sinusoids. (a) time domain signal. This corresponds to a window of length 128 samples centred around 0.315s, (b) power spectrum, (c) bicoherence. Note that the triad (8, 8, 16Hz) is phase-coupled and shows up with high bicoherence around (8, 8) while the triad (20, 20, 40Hz) is random phase and the bicoherence at (20, 20) is not significantly high. The power spectrum is unable to reveal this difference.

signal passes through a non-linear system, any quadratic non-linearity will generate sum and difference frequencies of input sinusoidal components which will then be phase coupled. Synchronized firing of pulses can also generate phase relationships.

Alpha burst with jitter

The input signal comprised of a short burst of an alpha frequency component at 8Hz. The envelope of the burst

Table 1 Significance levels of bicoherence

Number of epochs	b_{95}^2
10	0.3
20	0.15
100	0.03

The expected value of bicoherence at any bifrequency (f_1, f_2) is 0 if the harmonic components are random-phase. This also holds for a Gaussian random process. However, when a finite number, N , of epochs is used in estimation, bicoherence values will be non-zero. 95% of the values will lie below b_{95}^2 .

was the convolution of a short, unit-amplitude rectangular pulse of four samples width with itself four times. The width of the burst was thus 13 samples or 65ms and its waveform was a coarse approximation to a Gaussian. The width of the burst is only long enough to about one-half of a cycle of the alpha component and therefore the signal looks similar to a spike. Jitter was introduced to this burst and its position varied uniformly randomly by four samples or 20ms on either side from realization to realization. The alpha component was synchronized to the envelope of the burst such that it maintained its phase relative to the start of the burst. The amplitude of the alpha component was constant, equal to 1. Again, 20 realizations were generated and Gaussian noise was added to each realization keeping SNR at -15 dB. A typical time domain waveform of this input signal is shown in Figure 5a. The power spectrum is shown in Figure 5b. Note that there is no sharp peak at 8Hz and the power is spread in the frequency domain as expected from the confinement in time owing to the burst. The

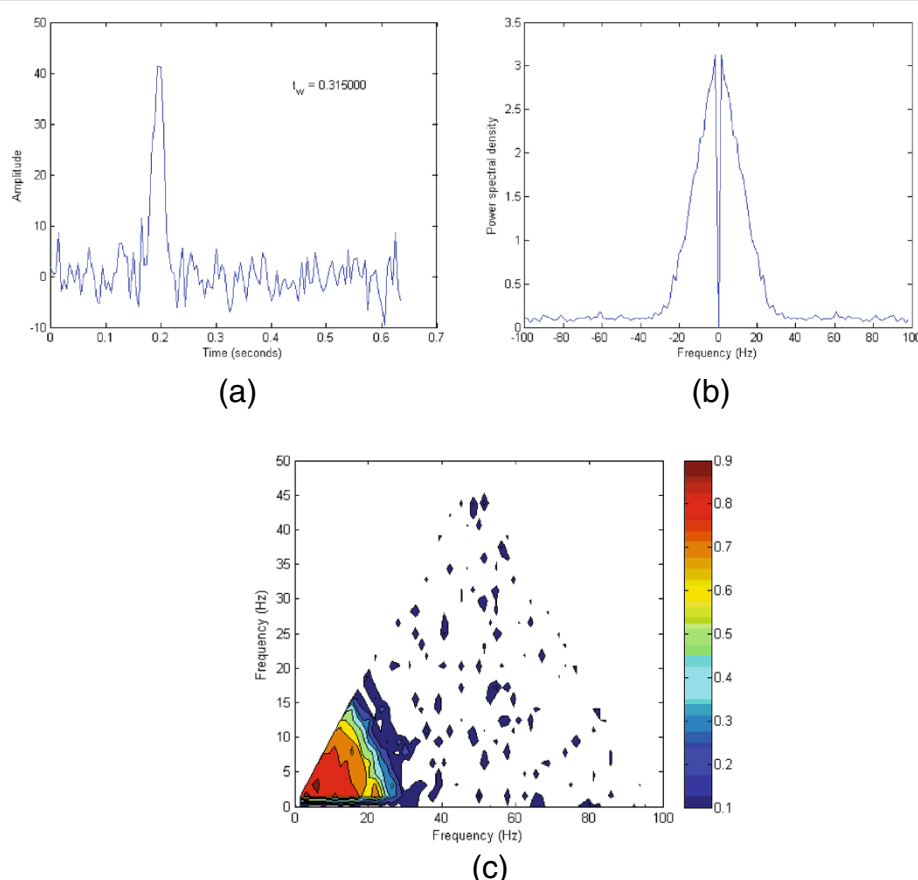


Figure 5 Alpha burst with jitter. (a) time domain signal, **(b)** power spectrum, **(c)** bicoherence.

power spectrum, however, would be the same regardless of any synchronization between the burst envelope and the alpha frequency sinusoid. The bicoherence, shown in Figure 5c shows significant high values around (8, 8). If there was no phase relationship between the envelope and the sinusoid, some spikes would be positive and others negative of varying amplitudes, depending on which part of a cycle of the alpha wave coincides with the burst envelope, and the bicoherence would tend to zero as the Fourier phases are random relative to each other. But that is not the case here. The spread of bicoherence values around (8,8) is similar to what would arise from leakage when a window is applied to each realization except that here the window is much shorter. This simulation example shows that information from transient events can be captured in the time-varying bicoherence. Application of time-domain windows of short duration as may occur with wavelet transforms at different scales can have a significant impact on the bispectrum. Although a wavelet transform based approach might appear to be suited to detection of such transient waveforms, it must be pointed out that in the presence of jitter, ensemble averaging of time-domain signal realizations can destroy any signature

of the event. Without such averaging, a wavelet transform based feature can capture the event in each realization separately at appropriate scales. However, when the signal to noise ratio falls very low, as is the case with EEG signals, it will not be possible to obtain a signature of the event from any single realization and ensemble averaging will not help because of the jitter. Bicoherence, however, as computed here, will be significantly high provided enough realizations are averaged in the estimate and it is tolerant of any jitter as long as the event is within the processing window in all the realizations. This is owing to the translation invariance property of the bispectrum; linear frequency dependent phase shifts arising from translation cancel out in the triple product of Fourier coefficients in Equation 16 or 18.

Beta burst with jitter

The input signal here comprised of a short burst of beta frequency at 20Hz. The rest of the details are the same as above for the alpha frequency burst except that the width of the burst can hold more than one cycle and therefore the typical time domain waveform shown in Figure 6a is no longer just a single spike. The power spectrum

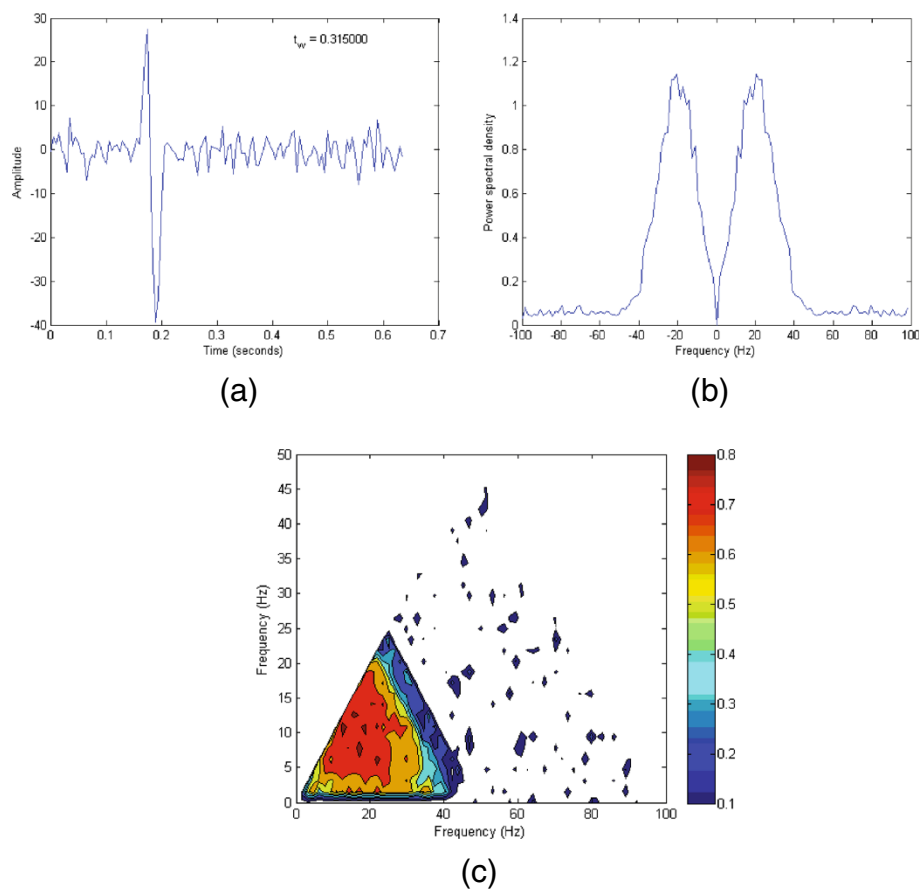


Figure 6 Beta burst with jitter. (a) time domain signal, **(b)** power spectrum, **(c)** bicoherence.

shown in Figure 6b is again spread out in the frequency domain but has its peak around 20Hz. The bicoherence shown in Figure 6c shows significantly high bicoherence at (20, 20) and in the region around it, especially towards lower frequencies. This simulation example shows that bicoherence can detect transient events that show phase synchronization characteristics and also reveal differences between them.

In practice, when time-varying bicoherence is computed, the processing window will slide across into the event, around it and then away from it. The event is not tracked with high resolution in bifrequency space with the processing parameters used above, but that is not always necessary. The methodology developed here is intended to use information from one or more such events—wave or transient—that occur in the EEG signal, and allow channels to be grouped based on the similarity of their activity in response to a stimulus.

It is difficult to validate the methodology with real EEG data because there does not exist a signal model for evoked response that would predict significant bicoherence or quadratic coupling at particular bifrequencies

which can then be compared to experimental findings. Confidence in the usefulness of the methodology can be improved if the spatial and temporal patterns of bicoherence can be correlated with simultaneously acquired data from a complementary methodology like functional magnetic resonance imaging. Such data are not available for this work.

In this study, tests are reported that examine whether the bicoherence evolution patterns can be used to establish relationships between different regions of the brain. EEG electrodes from the array are segmented into groups based on their bicoherence evolution in the evoked response. Signals from two electrodes that have the same bicoherence as a function of time following the stimulus will have the same fraction of power that is quadratically phase-coupled. Whether this suggests involvement in the particular response or joint inactivity is not clearly established by this procedure alone. The segment maps of electrode groups thus produced is analyzed together with knowledge about different lobes and cortical regions involved in the particular response. For the visual response, the primary visual cortex is in the occipital

region and other cortical regions and information pathways have been identified by neuroscience research [46]. This is a degree of validation of the methodology. Such segmentation can be done using various bifrequencies.

Application to real data

EEG data were collected using 19 EEG channels. 17 of these channels are shown in Figure 3. The standard 10–20 system of EEG electrode placement (explained in [47]) was used. Auxiliary electrodes were placed above and below one eye and to the right and left of the eyes to capture eye movements in the horizontal and vertical directions. These signals were used in rejecting epochs where eye blinking occurred because the interest is in visual processing and the EEG signal is desired to be free of motor activity related potentials. An LED light flash was used as the stimulus. A stimulus-locked signal was obtained synchronously with the EEG. This signal provides a sharp transition at the time of application of the stimulus. It is used to segment a continuously recorded EEG ensemble into epochs.

Left hemisphere EEG electrodes were referenced to an electrode behind the left ear and right electrodes similarly to a reference behind the right ear. These reference electrodes are not expected to vary significantly as a result of the stimulus. Medial electrodes were referenced to ground. A 50Hz power supply frequency rejection filter is provided with the system. Ag–AgCl electrodes with conductive brine Gel were used with the KT88-2400 Contec 24 channel EEG system for data acquisition. Electrode contact impedances were checked to be low enough using the LED indicator on the system.

Data were captured with each subject blindfolded in the left eye for one session and then in the right eye in the following session. Samples were continuously collected at 200Hz and segmented to yield 512 sample epochs. The epoching was done offline. Epochs were screened and selected free of artifacts such as eye blinks. The initial screening process was software driven and it was followed by manual examination of the output to finalize the selection process. All epochs processed were ensured to be free of noise from eye blinking. Each epoch represented 2.56s starting 0.56s prior to the application of the stimulus. Data were collected from five subjects. A variable number of usable epochs were obtained from each session following the screening process. More than 20 usable epochs from each session were available from three out of five subjects, labelled as E01, E02 and E05. These are processed in the work reported here.

Processing

Processing decisions must be made with regard to the type of window, the length of the window and the time duration over which the window centre will be

moved. Although windows other than rectangular, such as Hanning, Hamming or Blackman, are routinely employed in spectral analysis and provide better suppression of leakage, they are not preferred in this context. These windows are essentially multiplying the data samples in time and will result in a convolution of the true spectrum with the spectrum of the window. This introduces additional complexity in analysis and interpretation and is less desirable than some spectral leakage in this context. A rectangular window is therefore adopted. Each epoch was segmented to contain data for 2s following a visual stimulus. This is enough time for the response to a flash visual stimulus to be registered. Visual evoked potential markers identified from grand averages lie within about one second from the stimulus. Further, data are to be clean from eye blinks and larger time intervals make this more difficult. Since data are sampled at 200Hz, this is about 400 samples. If time-varying bispectra are to be computed over an interval of at least one second following the response, the upper limit on the window length becomes one-half this interval or 200 samples. A lower limit is set by the frequency resolution desired. For a frequency resolution better than 2Hz, the window length must be greater than 0.5s or 100 samples. A window length of 128 samples makes it convenient to use an FFT algorithm. If the window length is increased, the frequency resolution will be better but if the underlying phenomenon is transient and short-lived, the trade-off in time resolution will be a problem. If the window is made shorter, the frequency resolution is too poor for EEG frequencies of interest such as the alpha at around 8Hz. A rectangular sliding window of 128 samples was therefore used. This yields a frequency resolution of 1.5625Hz in the output of each short-time FFT. In order to move this window with its centre starting from the zero time reference derived from the stimulus, each epoch must contain at least 64 samples before the reference. There can be a small offset between the time reference and the actual stimulus event. In order to account for this and to facilitate some processing before the onset of the stimulus additional samples prior to the zero time reference were included in each epoch. This number was selected to be 112 such that each epoch was 512 samples, a multiple of 2. This resulted in epochs between -0.56 and 2s with respect to the time reference. For the results presented here, the time duration prior to the onset of the stimulus is not important. The centre of the rectangular window is moved from -0.07 to 1.2s, covering the region of time over which the response is likely to manifest itself significantly and allowing for any small negative offset. It may be noted that even when the window is centred at -0.07 a number of data samples come from post-stimulus measurements. Twenty epochs are processed here providing ensemble averaging. Bicoherence above 0.15 implies that there is

95% confidence that the true value is not zero as shown in Table 1.

A large number of plots can be generated using the methodology and bicoherence can be visualized as a function of a number of different variables—over the bifrequency plane, with time and with channels. In this study, the alpha wave and its harmonics are of particular interest. The alpha frequency can vary over a few Hertz with age and from individual to individual. It is typically around 8Hz for an adult. A number of adjacent frequencies were examined in the tests but 8, 4, and 20Hz and their bifrequency combinations are selected for figures in this work. Only auto-bicoherence is reported in this study. A frequency of 8Hz actually implies the frequency closest to it when the 128 point DFT is computed. This is the 5-th frequency bin at 7.8125Hz. Similarly 9Hz refers to the closest DFT bin at 9.375Hz. For the sake of clarity, frequencies are kept as integers in the discussion here, with the effect of finite resolution implicit.

Bicoherence is tracked with time for all EEG channels. It is computed with two sample skips or every 10ms. Bicoherence patterns are compared across pairs of electrodes. Time-varying bicoherence plots are shown in Figure 7. A mean of absolute difference of bicoherence (BMAD) is calculated for each pair of patterns. Each bicoherence lies between 0 and 1. The BMAD value also lies between 0 and 1. If two bicoherence patterns are exactly identical it will be zero.

Electrodes are separated into segments based on using the following procedure:

- (1) Group all electrodes, $x = 1 \dots K$, into one group labelled $i = 1$.
- (2) Take the first electrode, say x , at level i and examine $\text{BMAD}(x, y)$ for all $y > x$.
- (3) Apply a threshold in the range, $(\text{BMAD}_{\max}, \text{BMAD}_{\min})$, to each where the range of threshold is between the maximum and minimum values.
- (4) If $\text{BMAD}(x, y) > \text{Th}$, then change the label of y to $i + 1$.
- (5) Replace i with $i + 1$ and go to step 2 until all electrodes are processed.
- (6) If the number of segments is greater than 5 terminate else lower the threshold and proceed to step 1.
- (7) Repeat the above procedure until the lowest possible threshold is reached.

The above segmentation results in greater than 5 segments in most but not all cases. It imposes the constraint that segment separation be achieved using the same threshold for separating every group from its parent. Sometimes there is no single threshold that will necessarily produce five or more segments. The attempt is to divide

the electrodes into a small number of regions that might have a correspondence with known cortical regions and it will be difficult to interpret what is happening if there are too many segments.

Electrode groups are plotted using different colours based only on the order in which they get separated. All electrodes in a given segment have the same colour. No value is attached to the colour and it is not a measure. An indication of the BMAD between two electrodes within a segment is provided by drawing a line between the two and making the thickness of the line logarithmically proportional to the closeness of their BMAD value to zero. This is done from an arbitrarily chosen (depending on the first electrode to separate into that segment) electrode for that segment to all other electrodes in the segment. In these plots, electrodes are represented by three concentric circles of a given colour corresponding to the segment. Such maps are shown in Figure 8.

Results

Time-varying bicoherence obtained on the data is analyzed at different frequencies and over different channels and as a function of time.

Bicoherence distribution over frequencies

Significant bicoherence was observed over large regions in the bifrequency plane and this pattern changed with time. A typical plot of bicoherence over the entire principal domain in bifrequency space is shown in Figure 2. This plot corresponds to the EEG channel O2, computed with a window around time $t = 0.10$ s. This plot is for subject E01. Significantly high bicoherence (above 0.15) extends over the several EEG frequency bands. Particularly high values above 0.5 are observed in the delta band (around 2–4Hz), alpha band (around 8–12Hz), beta band (around 20Hz) and extending to much higher harmonics.

It is quite possible that actual phase-coupling occurs at discrete frequencies and their harmonics and what is observed over larger frequency bands is owing to spectral leakage arising from the finite length DFT. If an attempt were being made to hypothesize or validate a signal model this would be of consequence. If it is desired to use the bicoherence values without any assumption of a parametric signal model for the purpose of analyzing spatial-temporal functioning of the brain it does not matter whether there is leakage as long as the processing remains consistent across any comparisons. The latter is the case here. Specific bifrequencies are selected and the bicoherence fluctuation with time is computed for all channels.

Bicoherence distribution over channels

Figure 3 shows the distribution of bicoherence at (8, 8) for E01 computed with a window around time $t = 0.10$

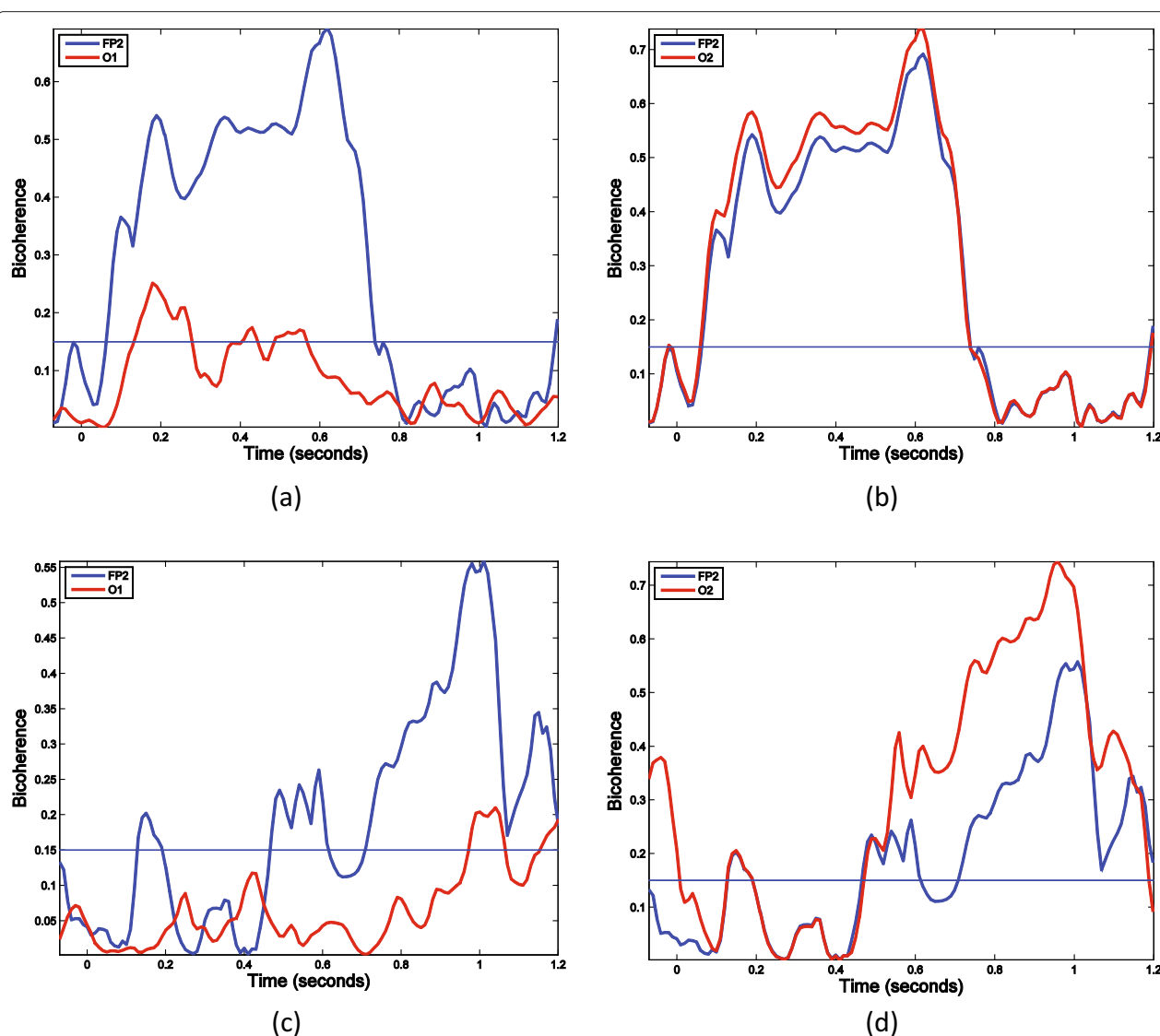


Figure 7 Bicoherence evolutions at (4, 4). Bicoherence evolutions at (4, 4) as functions of time for subject E01: **(a)** over Fp2 and O1 for a flash stimulus viewed through the left eye, **(b)** over Fp2 and O2 and the left eye, **(c)** over Fp2 and O1 and the right eye, **(d)** over Fp2 and O2 and the right eye. There is high correlation in **(b)** and the BMAD values will be small. In **(d)**, the correlation immediately following the stimulus with a delay of about 100ms is nearly perfect for the next 400ms.

seconds. These are the same conditions as those in Figure 2 except that only one bifrequency is selected. Bicoherence is plotted here for 17 channels and is interpolated in between to yield a contour-filled colour plot. It can be observed that some electrodes such as Fz, Pz and O2 show very high bicoherence. These plots can be hard to interpret because high bicoherence values can emerge and fade away if they were viewed as a sequence with time, and there is no knowledge of the involvement of different regions of the brain at different times following a stimulus to validate this. There are also a number of such plots possible, one for each bifrequency, and there is no signal model or precise understanding of the role of quadratic non-linearity in the firing of neural

populations to guide any selection based on bifrequency. The problem can be simplified to some extent by making pair-wise channel comparisons, examining how bicoherence evolves with time for two electrodes. It can then be determined whether there is strong correlation or similarity between the bicoherence variations with time at the two channels.

Bicoherence evolution with time

Figure 7a through Figure 7d plot bicoherence at (8, 8) for subject E01, at two electrodes each, as a function of time. Figure 7a shows the bicoherence at Fp2 and O1 for a flash stimulus to the left eye. Fp2 is in the frontal region and close to the eyes whereas O2 (and O1) is close to the

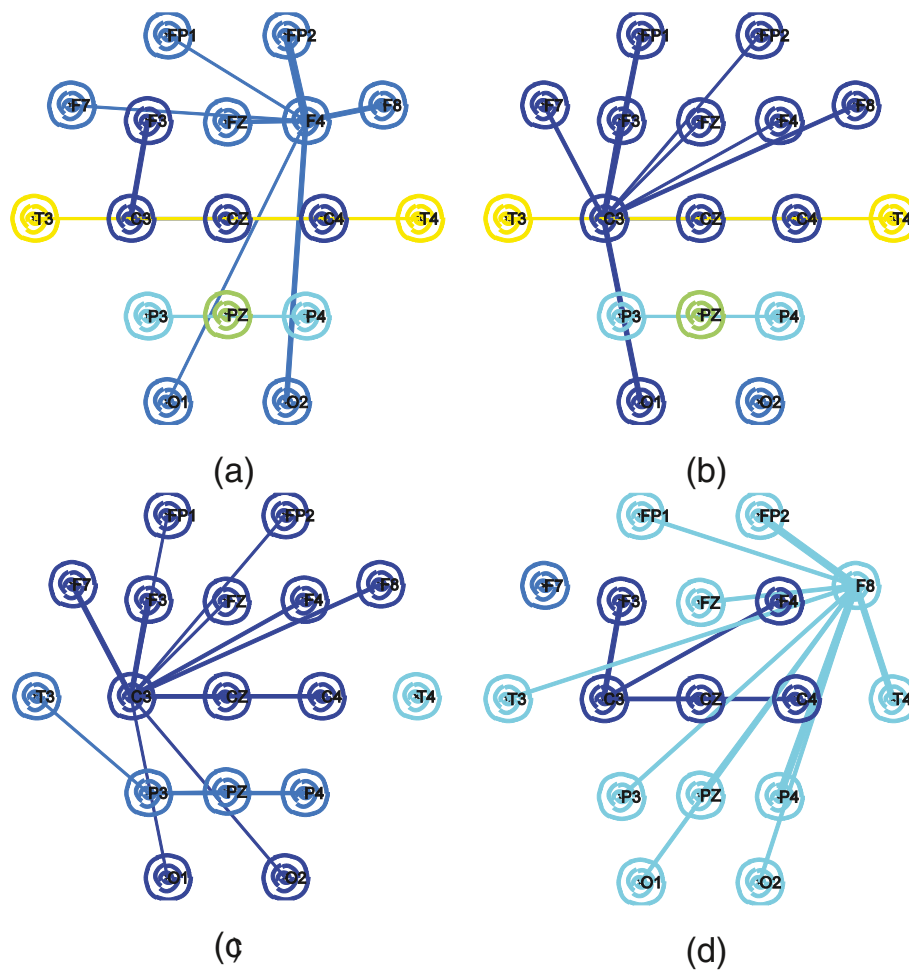


Figure 8 Segmentation maps (4, 4). Segmentation of electrodes based on bicoherence at (4, 4): **(a)** for E01 and the left eye, **(b)** for E01 and the right eye, **(c)** for E02 and the left eye, **(d)** for E02 and the right eye. In each region, electrodes are linked to a representative electrode using lines whose thickness is indicative of the inverse of the BMAD value on a logarithmic scale. The thicker the line, the closer the two electrodes are in terms of their bicoherence fluctuation.

occipital region, known for visual processing activity in the human brain. Figure 7b shows bicoherence at electrodes Fp2 and O2 following a left stimulus. Figure 7c,d are similar bicoherence plots for a flash stimulus to the right eye.

It can be observed that the two bicoherence variations in Figure 7b (left eye) are very highly correlated. The mean absolute difference of bicoherence (BMAD) between Fp2 and O2 in this plot is very small. In Figure 7d (right eye) there is very high correlation between Fp2 and O2 only during the 100 to 500ms duration following the stimulus. After that the two functions separate although the trends are still quite positively correlated as judged visually. The BMAD taken over the entire 1.27s will be higher than that for Figure 7b.

Figure 7a,c show that O1 is not highly correlated with Fp2 for either left or right eye responses and the BMAD

values will be much higher. Based on these plots, Fp2 will be grouped with O2 for the left eye response even with a low threshold on BMAD. They will be grouped in the same segment for right eye response only if the threshold of separation is large.

In order to tell left from right, markers may need to be computed over significant intervals in time such as the 100 to 500ms duration. In this study, attention is focussed on holistic differences over the entire time response from -0.07 to 1.2s. These are used to segment the electrode array into groups. Segment maps are examined for comparison with existing knowledge about visual signal processing in the human brain.

Segment maps

Figures 8, 9 and 10 show segment maps that divide the electrode array into groups according to the temporal

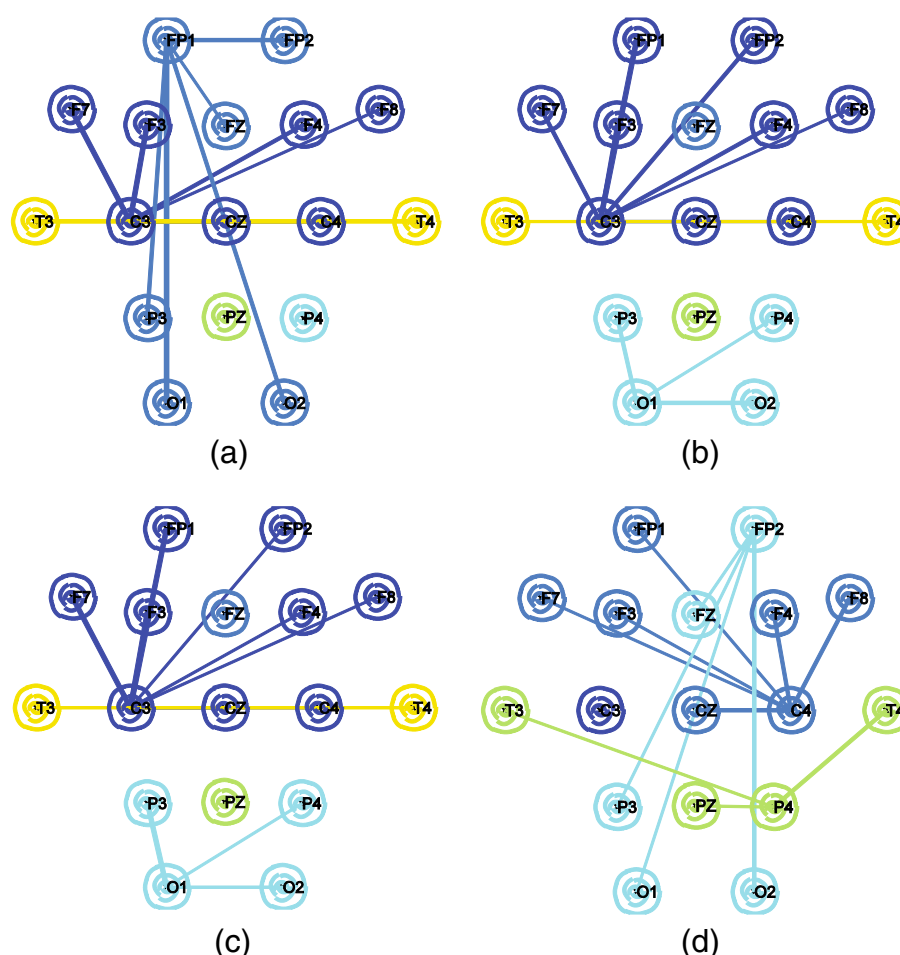


Figure 9 Segmentation maps (8, 8). Segmentation of electrodes based on bicoherence at (8, 8): **(a)** for E01 and the left eye, **(b)** for E01 and the right eye, **(c)** for E05 and the left eye, **(d)** for E05 and the right eye.

evolution of bicoherence using the algorithm described above. The objective is to divide the electrodes into groups and examine (a) whether they correspond to lobes—frontal, central, parietal and occipital, and (b) the nature of the connection between frontal regions close to the sensors (eyes) and the occipital region known to be involved in visual processing.

Figure 8a through Figure 8d use bicoherence at (4, 4) and show maps for subjects E01 and E02 viewing a flash stimulus through the left eye and the right eye. All the figures indicate a relationship between the frontal and occipital regions as being grouped in the same segment. Figure 8c groups parietal lobes and a temporal lobe in the same segment. Figure 8d groups parietal and temporal lobes also with the frontal and occipital but leaves the central electrodes separate. This requires further investigation.

Figure 9a through Figure 9d show a set of segment maps for subjects E01 and E05 viewing a flash stimulus

through the left and right eyes, using bicoherence at (8, 8). Figure 9a,d show frontal and occipital electrodes in the same segment but Figure 9b,c do not. It is likely that because the frequencies 8 and 16Hz, are also involved in motor activity, the latter is influencing bicoherence evolution phenomena. In Figure 9b,c, the BMAD values between Fp1, Fp2 and O1, O2 are relatively small although they appear in different segments. In Figure 9b, $BMAD(Fp2, O2)$ is 0.060, $BMAD(Fp1, O1)$ is 0.028 and $BMAD(O1, O2)$ is 0.099. In Figure 9c $BMAD(Fp2, O2)$ is 0.110, $BMAD(Fp1, O1)$ is 0.064, $BMAD(O1, O2)$ is 0.127.

Figure 10a through Figure 10d show similar segment maps for subjects E01 and E05 using bicoherence at the much higher frequencies of (20, 20). In Figure 10a,b,d frontal and occipital electrodes are connected. FZ is also included in the same segment. In Figure 10c, they form two separate segments. This is partly owing to the segmentation algorithm. Electrodes separate from parent groups based on one threshold distance from a

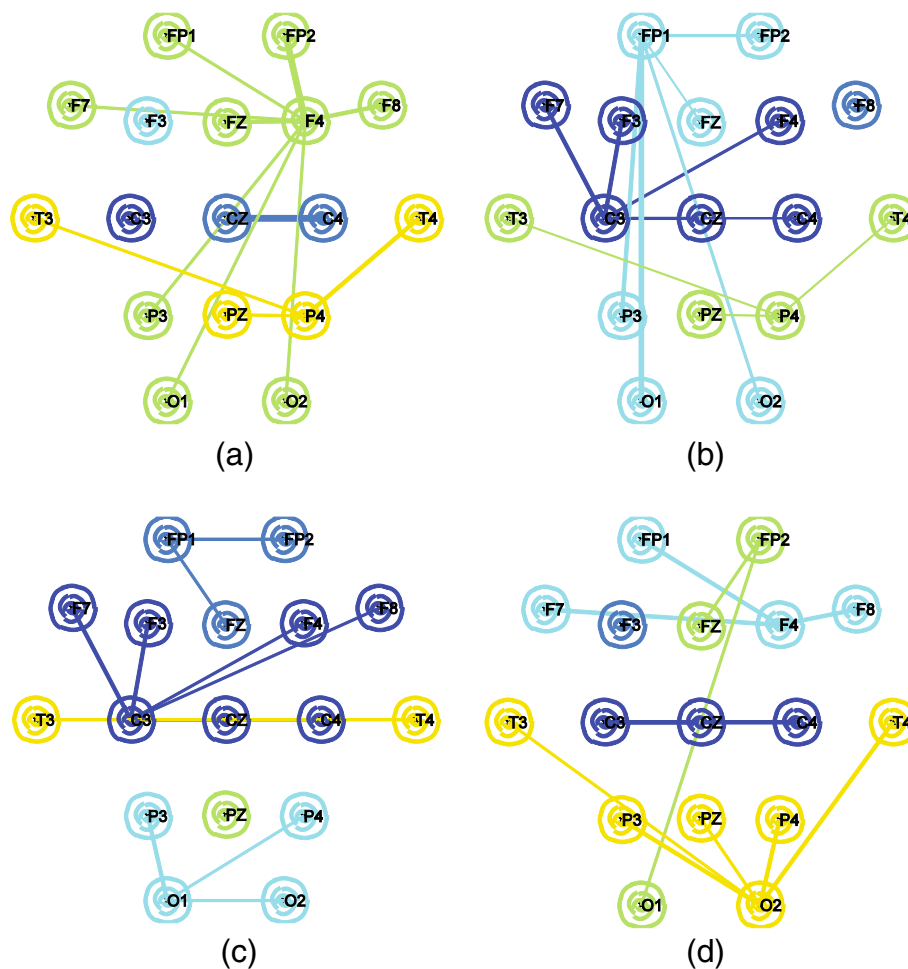


Figure 10 Segmentation maps (20, 20). Segmentation of electrodes based on bicoherence at (20, 20): **(a)** for E01 and the left eye, **(b)** for E01 and the right eye, **(c)** for E05 and the left eye, **(d)** for E05 and the right eye.

representative electrode and the procedure is repeated until a suitable threshold is found to create at least five segments if this is possible. The maps are only a rough guide. More information is obtained by closer examination of bicoherence as a function of time at electrodes. This is done for the map of Figure 10c in Figure 11a through Figure 11d.

At the beta band frequency of 20Hz and its second harmonic, frontal electrodes have bicoherence evolution highly correlated with central electrodes as shown in Figure 11a for F4 and C4. $\text{BMAD}(\text{F4}, \text{C4})$ is 0.003. Although they appear in different segments, F4 also is highly correlated with Fp2 as shown in Figure 11b and Fp2 is only separated because the segmentation algorithm attempts to find at least five segments if possible with a single threshold. $\text{BMAD}(\text{F4}, \text{Fp2})$ is 0.003. Figure 11c,d shows a comparison between bicoherence evolutions at F4 and O2 and F4 and O1, respectively. Some phenomena do not last the entire time duration used in the computation

of the BMAD. F4 is highly correlated with O2 for 700 milliseconds following the stimulus as shown in Figure 11c. $\text{BMAD}(\text{F4}, \text{O2})$ is 0.043. The BMAD between them is primarily dependent on the duration from 0.7 to 1.2s. Bicoherence at O1 on the other hand is rarely above the 95% statistically significant level of 0.15 and is not highly correlated with F4. $\text{BMAD}(\text{F4}, \text{O1})$ is 0.178.

Relationship to existing knowledge

Dynamics of visual recognition in the human brain are discussed in [46] which gives a recent review of the area. A great deal is known about the visual cortex from studies on humans and primates. The occipital region is mapped in this study by electrodes O1 and O2 and the visual cortex by O2. As explained in [46] (refer Figure 1 in [46]), the cortex is further divided into regions labelled as V1, V2, V3 and V4. More is known about V1, the primary visual cortex, than any other region. Information from the right and left eye is combined in V1. There are two main

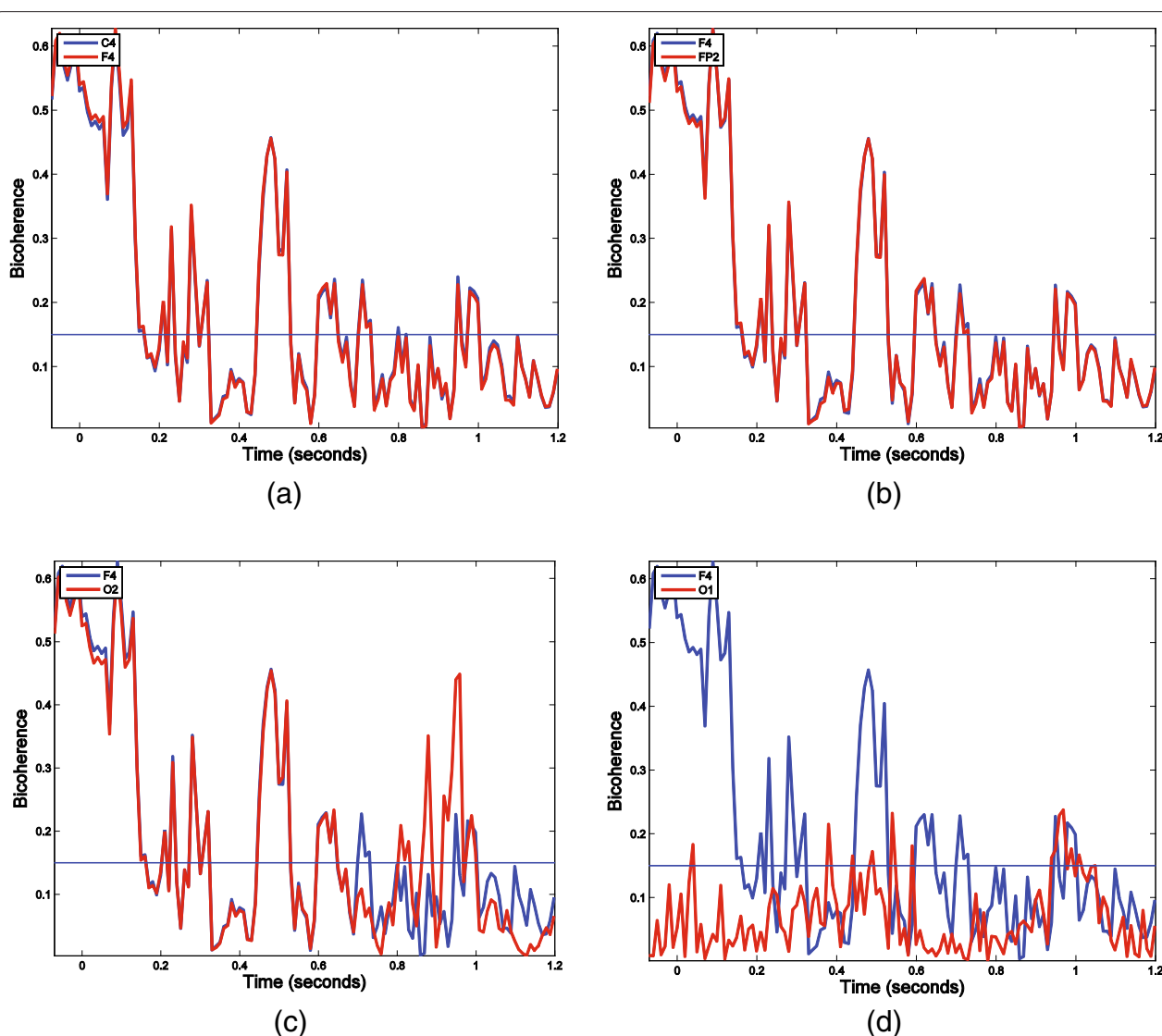


Figure 11 Bicoherence evolutions at (20, 20). Bicoherence evolutions at (20, 20) as functions of time for subject E05 for a flash visual stimulus viewed through the left eye: **(a)** over F4 and C4, **(b)** over F4 and FP2, **(c)** over F4 and O2 and **(d)** over F4 and O1. There is near perfect correlation in **(a)** and **(b)**—almost eerie for signals that are fluctuating so much. There is near perfect correlation in **(c)** for about 700ms following the stimulus. The correlation is weak in **(d)**, vastly different from the others. It serves to provide confidence that the near perfect correlations are not artefacts of data collection or processing. O1 and O2 are adjacent channels; F4 is not near either of them.

pathways from V1 leading to the frontal cortex and on to the hippocampus and surrounding structures. There are the dorsal pathway and the ventral pathway. The dorsal pathway includes the medial temporal area (MT) and the medial superior temporal area (MST) that are known to be involved in processing spatial information—object motion, position and depth perception from stereo vision. The ventral pathway includes the inferior temporal cortex (ITC) which is involved in discriminating colours and shapes. The EEG electrode array used in this study does not have the spatial resolution to differentiate all these regions—nor is it within the domain of the simple flash visual stimulus experimentation. This methodology is not

intended to produce the same level of detail in understanding as neuro-anatomical and neuro-physiological studies. However, it does have the advantage of working on scalp EEG and does not involve the use of intra-cranial electrodes or invasive surgical procedures on animals. Nevertheless, it is encouraging that segmentation of electrodes using this methodology establishes connections between the frontal and occipital regions and also indicates involvement of other areas in the parietal lobe.

The experiments here are intended to be preliminary validation of the use of temporal evolution of bicoherence in better understanding the functioning of the human brain using scalp EEG. The methodology does have the

advantage that different EEG frequency bands can be analyzed using the same set of experimentally recorded data.

Conclusions

In this work, a methodology for time-varying bispectral analysis of a multi-channel signal during a stimulus locked transient response is developed. Theoretical foundations of higher order spectral analysis are revisited in this context to examine its validity. It is applied to scalp EEG data. Time-varying bispectral analysis provides information additional to what can be obtained from grand averages (first order) or power spectral (second order) techniques. It is observed that some EEG channels exhibit time-varying bicoherence in response to a stimulus with remarkable correlation over time intervals and even similarity in magnitude lasting hundreds of milliseconds. A new methodology based on the mean absolute difference of bicoherence over such intervals is developed and applied to segment channels into sets that are more similar in their response to the stimulus than others. It is shown that channel segmentation using time-varying bicoherence of the EEG response evoked by visual stimuli is consistent with knowledge about involvement of the occipital region and frontal regions in visual processing. Future work possible includes validation of the methodology using other sensory stimuli such as auditory, the use of higher spatial resolution EEG, analysis of test data from subjects with known visual system impairments such as partial blindness, prosopagnosia (inability to recognize faces), achromatopsia (inability to perceive colours) and comparison of the markers obtained via this methodology with results obtained using functional magnetic resonance imaging. If the methodology can be developed for practical diagnostic and monitoring applications, it is a much less invasive and less expensive option than many others because it is based on scalp EEG recordings.

Competing interests

The author declares that he has no competing interests.

Acknowledgements

The author is grateful to Mr. M. Gnananathan and Mr. W. Maier at the Queensland University of Technology (QUT) for collecting visually evoked EEG data. This research was not funded by any external research grant. Computing support and publication costs were provided by QUT. The author retains copyright and rights to all commercial uses of intellectual property in accordance with employment regulations.

Received: 15 November 2011 Accepted: 7 June 2012

Published: 12 July 2012

References

1. G Dumermuth, P Huber, B Kleiner, T Gasser, Numerical analysis of electroencephalographic data. *IEEE Trans. Audio Electroacoust.* **18**(4), 404–411 (1970)
2. T Barnett, LC Johnson, P Naitoh, N Hicks, C Nute, Bispectral analysis of electroencephalogram signals during waking and sleeping. *Science*. **172**, 401–402 (1971)
3. T Bullock, J Achimowicz, R Duckrow, S Spencer, V Iragai-Madoz, Bicoherence of intracranial eeg in sleep, wakefulness and seizure. *Electroencephalogr. Clin. Neurophysiol.* **103**, 661–678 (1997)
4. J Muthuswamy, D Sherman, N Thakor, Higher-order spectral analysis of burst patterns in eeg. *IEEE Trans. Biomed. Eng.* **46**(1), 92–99 (1999)
5. C Pomfrett, A Pearson, in *IEEE Colloquium on New Measurements and Techniques in Intensive Care (Digest No. 1996/179)*, Dec 1996. Eeg monitoring using bispectral analysis, pp. London, UK. pp. 5/1–5/3
6. S Greenwald, C Smith, J Sigl, H Cai, P Devlin, in *[Engineering in Medicine and Biology, 1999. 21st Annual Conf. and the 1999 Annual Fall Meeting of the Biomedical Engineering Soc.] BMES/EMBS Conference, 1999. Proceedings of the First Joint*, vol. 1, The eeg bispectral index (bistm): development and utility, p. Atlanta, USA, 1999. p. 443
7. Y Tang, A Norcia, in *IEEE 17th Annual Conf. Eng. in Medicine and Biology Society*, vol. 2, Coherent bispectral analysis of the steady-state vep, pp. Montreal, Canada, 1995. pp. 901–902
8. M Shen, L Sun, in *Higher-Order Statistics, 1997. Proceedings of the IEEE Signal Processing Workshop*, July 1997, The analysis and classification of phonocardiogram based on higher-order spectra, pp. Banff, Canada. pp. 29–33
9. M Shen, Y Liu, F Chan, P Beadle, in *27th Annual International Conference of the Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005*, Jan 2005, Novel approach for time-varying bispectral analysis of non-stationary eeg signals, pp. Shanghai, China. pp. 829–832
10. S Minfen, S Lisha, X Congtao, Z Guoping, in *6th International Conference on Neural Information Processing, 1999. Proceedings. ICONIP '99*, 2, Application of higher-order statistics for the analysis of electroencephalogram in different brain functional states, pp. Perth, Australia, 1999. pp. 622–626
11. S Minfen, S Lisha, P Beadle, in *Advances in Medical Signal and Information Processing, 2000. First International Conference on (IEEE Conf. Publ. No. 476)*, Parametric bispectral estimation of eeg signals in different functional states of brain, pp. Briston, UK, 2000. pp. 66–72
12. C Tallon-Baudry, O Bertrand, C Delpuech, J Pernier, Stimulus specificity of phase-locked and non phase-locked 40 hz visual responses in humans. *Neuroscience*. **16**(13), 4240–4249 (1996)
13. A Shiryayev, Some problems in the spectral theory of higher order moments. *Theor. Prob. Appl.* **5**, 265–284 (1960)
14. D Brillinger, An introduction to polyspectra. *Ann. Math. Stat.* **36**, 1351–1374 (1965)
15. DR Brillinger, M Rosenblatt, in *Spectral Analysis of Time Series*, ed. by B. Harris. Computation and Interpretation of k-th Order Spectra (Wiley, New York, 1967) pp. 189–232
16. SS Nikias, A Petropulu, *Higher Order Spectral Analysis*. (Prentice Hall, Englewood Cliffs, 1993)
17. WA Gardner, *Introduction to Random Processes with Applications to Signals and Systems*. (MacMillan, New York, 1986)
18. K Hasselman, W Munk, G MacDonald, in *Time Series Analysis*, ed. by M. Bispectra of Ocean Waves Rosenblatt (Wiley, New York, 1963) pp. 125–130
19. DR Brillinger, M Rosenblatt, in *Spectral Analysis of Time Series*. Harris, B. Asymptotic Theory of Estimates of k-th Order Spectra (Wiley, New York, 1967) pp. 153–188
20. C Nikias, M Raghuveer, Bispectrum estimation: a digital signal processing framework. *Proc. IEEE*. **75**(7), 869–891 (1987)
21. B Boashash, E Powers, A Zoubir, *Higher Order Statistical Signal Processing* (Longman Australia, Melbourne, 1995)
22. MB Priestley, *Spectral Analysis and Time Series* (Academic Press, San Diego, 1981)
23. LE Franks, *Signal Theory* (Dowden and Culver, Stroudsburg, 1981)
24. M Hinich, H Messer, On the principal domain of the discrete bispectrum of a stationary signal. *IEEE Trans. Signal Process.* **43**(9), 2130–2134 (1995)
25. K Vixie, M Wolinsky, D Sigeti, in *Proceedings of the Fifth International Symposium on Signal Processing and Its Applications, 1999. ISSPA '99*, vol. 1, 1999, The bispectral aliasing test: a clarification and some key examples. pp. 255–258
26. P Schreier, in *Communications Theory Workshop, 2005. Proceedings. 6th Australian*. A note on aliasing in higher order spectra. Feb 2005, pp. 184–188

27. W Gardner, Common pitfalls in the application of stationary process theory to time-sampled and modulated signals. *IEEE Trans. Commun.* **35**(5), 529–534 (1987)
28. V Chandran, S Elgar, Mean and variance of estimates of the bispectrum of a harmonic random process—an analysis including leakage effects. *IEEE Trans. Signal Process.* **39**(12), 2640–2651 (1991)
29. P Huber, B Kleiner, T Gasser, G Dumermuth, Statistical methods for investigating phase relations in stationary stochastic processes. *IEEE Trans. Audio Electroacoust.* **19**(1), 78–86 (1971)
30. V Chandran, S Elgar, A general procedure for the derivation of principal domains of higher-order spectra. *IEEE Trans. Signal Process.* **42**(1), 229–233 (1994)
31. R Haubrich, Earth noises, 5 to 500 millicycles per second 1. *J. Geophys. Res.* **70**, 1415–1427 (1965)
32. Y Kim, E Powers, Digital bispectral analysis and its applications to nonlinear wave interactions. *IEEE Trans. Plasma Sci.* **PS-7**(2), 120–131 (1979)
33. V Chandran, S Elgar, B Vanhoff, Statistics of tricoherence. *IEEE Trans. Signal Process.* **42**, 3430–3440 (1994)
34. S Elgar, R Guza, Statistics of bicoherence. *IEEE Trans. Acoust. Speech Signal Process.* **36**, 1667–1668 (1988)
35. W Gardner, A sampling theorem for nonstationary random processes (corresp.) *IEEE Trans. Inf. Theory.* **18**(6), 808–809 (1972)
36. F Garcia, I Lourtie, J Buescu, L2(r) nonstationary processes and the sampling theorem. *IEEE Signal Process. Lett.* **8**(4), 117–119 (2001)
37. G Frazer, B Boashash, in *1991 Conference Record of the Twenty-Fifth Asilomar Conference on Signals, Systems and Computers, 1991*, vol. 2, Nov 1991. Detection of underwater transient acoustic signals using time-frequency distributions and higher-order spectra, pp. pp. 1103–1107
38. B Boashash, B Ristic, in *Time-Frequency and Time-Scale Analysis, 1992, Proceedings of the IEEE-SP International Symposium*, Oct 1992. Polynomial wigner-ville distributions and time-varying higher-order spectra, pp. pp. 31–34
39. B Boashash, G Frazer, in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 1992. ICASSP-92*, vol. 5, Mar 1992. Time-varying higher-order spectra, generalised wigner-ville distribution and the analysis of underwater acoustic data, pp. 193–196
40. B Boashash, B Ristic, in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 1994. ICASSP-94*, vol. iv, Apr 1994. Time-varying higher-order cumulant spectra: application to the analysis of composite fm signals multiplicative and additive noise, pp. pp. IV/325–IV/328
41. B Boashash, P O'Shea, Polynomial wigner-ville distributions and their relationship to time-varying higher order spectra. *IEEE Trans. Signal Process.* **42**(1), 216–220 (1994)
42. B Ristic, G Roberts, B Boashash, in *International Conference on Acoustics, Speech, and Signal Processing, 1995. ICASSP-95*, vol. 3, May 1995. Higher-order scale spectra and higher-order time-scale distributions. pp. 1577–1580
43. P Boles, B Boashash, in *Intl. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, April 1988. The cross wigner-ville distribution—a two dimensional analysis method for the processing of vibroseis seismic signals, pp. 904–90
44. A El-Jaroudi, T Akgul, M Simaan, in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 1994. ICASSP-94*, vol. iv, Apr 1994. Application of higher order spectra to multi-scale deconvolution of sensor array signals. pp. IV/413–IV/416
45. K Gurley, T Kijewski, A Kareem, First- and higher-order correlation detection using wavelet transform. *J. Eng. Mech.* **129**, 188–201 (2003)
46. J Blumberg, G Kreiman, How cortical neurons help us see: visual recognition in the human brain. *J. Clin. Invest.* **120**(9), 3054–3063 (2010)
47. E Niedermeyer, F Lopez da Silva, *Electroencephalography: Basic Principles, Clinical Applications and Related Fields*. (Lippinkott Williams and Wilkins, Philadelphia, 2004)

doi:10.1186/1687-6180-2012-140

Cite this article as: Chandran: Time-varying bispectral analysis of visually evoked multi-channel EEG. *EURASIP Journal on Advances in Signal Processing* 2012 **2012**:140.

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

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

Submit your next manuscript at ► springeropen.com