LETTER Communicated by Robert Gilmore

Delay Differential Analysis of Time Series

Claudia Lainscsek
claudia@salk.edu
Terrence J. Sejnowski
terry@salk.edu
Howard Hughes Medical Institute, Computational Neurobiology Laboratory, Salk Institute for Biological Studies, La Jolla, CA 92037, U.S.A. and Institute for Neural Computation, University of California San Diego, La Jolla, CA 92093, U.S.A.

Nonlinear dynamical system analysis based on embedding theory has been used for modeling and prediction, but it also has applications to signal detection and classification of time series. An embedding creates a multidimensional geometrical object from a single time series. Traditionally either delay or derivative embeddings have been used. The delay embedding is composed of delayed versions of the signal, and the derivative embedding is composed of successive derivatives of the signal. The delay embedding has been extended to nonuniform embeddings to take multiple timescales into account. Both embeddings provide information on the underlying dynamical system without having direct access to all the system variables. Delay differential analysis is based on functional embeddings, a combination of the derivative embedding with nonuniform delay embeddings. Small delay differential equation (DDE) models that best represent relevant dynamic features of time series data are selected from a pool of candidate models for detection or classification.

We show that the properties of DDEs support spectral analysis in the time domain where nonlinear correlation functions are used to detect frequencies, frequency and phase couplings, and bispectra. These can be efficiently computed with short time windows and are robust to noise. For frequency analysis, this framework is a multivariate extension of discrete Fourier transform (DFT), and for higher-order spectra, it is a linear and multivariate alternative to multidimensional fast Fourier transform of multidimensional correlations. This method can be applied to short or sparse time series and can be extended to cross-trial and cross-channel spectra if multiple short data segments of the same experiment are available. Together, this time-domain toolbox provides higher temporal resolution, increased frequency and phase coupling information, and it allows an easy and straightforward implementation of higher-order spectra across time compared with frequency-based methods such as the DFT and cross-spectral analysis.

Neural Computation **27, 594–614** (2015)
doi:10.1162/NECO_a_00706
© 2015 Massachusetts Institute of Technology

1 Introduction

Delay differential analysis (DDA) uses delay differential equations (DDE) to reveal dynamical information from time series. A DDE unfolds timescales and couplings between those timescales. In nonlinear time series analysis, an embedding converts a single time series into a multidimensional object in an embedding space that reveals valuable information about the dynamics of the system without having direct access to all the system's variables (Whitney, 1936; Packard, Crutchfield, Farmer, & Shaw, 1980; Takens, 1981; Sauer, Yorke, & Casdagli, 1991). A DDE is a functional embedding, an extension of a nonuniform embedding with linear or nonlinear functions of the time series. DDA therefore connects concepts of nonlinear dynamics with frequency analysis and higher-order statistics.

A Fourier series decomposes periodic functions or periodic signals into the sum of a (possibly infinite) set of simple oscillating functions (Fourier, 1822). The Fourier transform $F(f)$ of the function $f(t)$ has many applications in physics and engineering and is referred to as frequency domain and time-domain representations, respectively. The relationship between frequency analysis and the analysis of frequency or phase couplings in the time domain is poorly understood (see Hjorth, 1970; Chan & Langford, 1982; Raghuveer & Nikias, 1985, 1986; Stankovic, 1994).

This study introduces spectral DDE methods and applies them to simulated data. In a companion paper (Lainscsek, Hernandez, Poizner, & Sejnowski, 2015), the methods are applied to electroencephalography (EEG) data.

The letter is organized as follows. In section 2, DDA is introduced as a nonlinear time series classification tool. In section 3, functional embeddings serve as a link between nonlinear dynamics and traditional frequency analysis and higher-order statistics. The time-domain spectrum (TDS) and the time-domain bispectrum (TDB) are introduced. These are extended in section 4 to sparse data and in section 5 to the cross-trial spectrogram (CTS) and the cross-trial bispectrogram (CTB). The results are compared to traditional wavelet analysis. The conclusion in section 6 compares the complementary strength and weakness of time domain and frequency domain approaches to pattern detection and classification of time series.

2 DDA as a Nonlinear Classification Tool

Figure 1 illustrates the connections of DDA to nonlinear dynamics and global modeling. In nonlinear time series analysis, a single variable measurement (the blue box in Figure 1) comes from an unknown nonlinear dynamical system (pink box in Figure 1). In Figure 1 the Rössler system (Rössler, 1976) is used for illustration. A general existence theorem for embeddings in an Euclidean space was given by Whitney (1936). A generic map from an n-dimensional manifold to a $2n+1$ dimensional

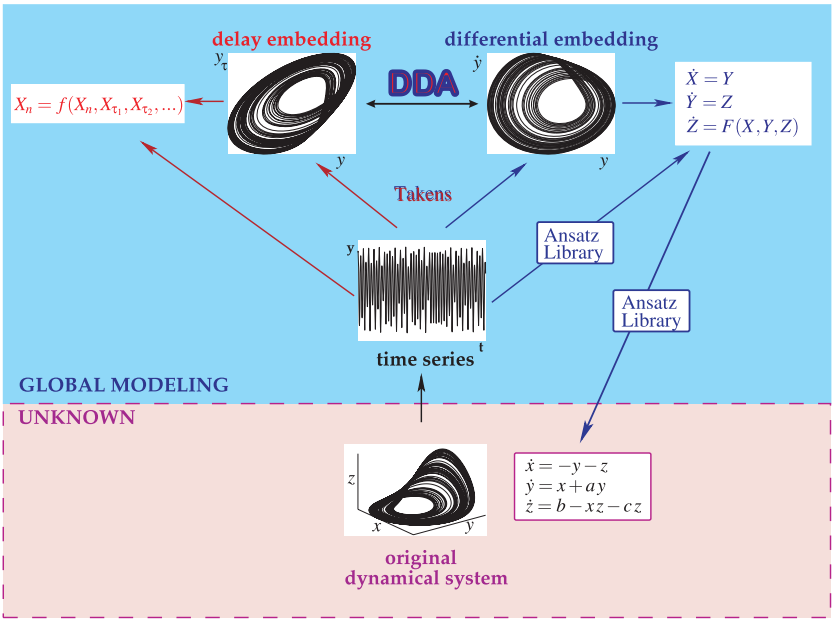

Figure 1: Delay differential analysis (DDA) and global modeling.

Euclidean space is an embedding. Whitney's theorem implies that each state can be identified uniquely by a vector of $2n+1$ measurements, thereby reconstructing the phase space. This theorem is the basis for the embedding theorems of Packard et al. (1980) and Takens (1981) for reconstruction of an attractor from single time series measurements. Takens proved that instead of $2n+1$ generic signals, the time-delayed versions $[y(t), y(t-\tau), y(t-2\tau), \ldots, y(t-n\tau)]$ or the successive derivatives of one generic signal would suffice to embed the n-dimensional manifold. Although the reconstruction preserves the invariant dynamical properties of the dynamical system, the embedding theorems provide no tool to reconstruct the original dynamical system in the original phase space from the single measurement or the embedding.

The Ansatz Library is a tool to reconstruct a general system from a single measurement that contains the original dynamical system as a subset (Lainscsek, 1999, 2011; Lainscsek, Letellier, & Gorodnitsky, 2003; Lainscsek, Letellier, & Schürrer, 2001). For example, for the Rössler system (Rössler, 1976),

$$\begin{aligned}
\dot{x} &= a_{1,2}\, y + a_{1,3}\, z \\
\dot{y} &= a_{2,1}\, x + a_{2,2}\, y \\
\dot{z} &= a_{3,0} + a_{3,3}\, z + a_{3,6}\, x z,
\end{aligned} \tag{2.1}$$

the standard form (or the functional form of the differential embedding) is

$$\begin{aligned}\dot{X} &= Y,\\ \dot{Y} &= Z,\\ \dot{Z} &= F(X, Y, Z),\end{aligned} \tag{2.2}$$

where $X = x_2$ and its successive derivatives define the new state-space variables X, Y, and Z. The function $F(X, Y, Z)$ is explicitly

$$\begin{aligned}F(X, Y, Z) = \alpha_1 + X\alpha_2 + X^2\alpha_3 + Y\alpha_4 + XY\alpha_5\\ + Y^2\alpha_6 + Z\alpha_7 + XZ\alpha_8 + YZ\alpha_9.\end{aligned} \tag{2.3}$$

When we use the relations between the coordinates α_r of the differential embedding and the coordinates $a_{i,*}$ the model of the Rössler system becomes

$$\begin{aligned}\dot{\varphi}_x &= d^2\,\frac{1}{b}\,a_{1,2}\,y + d^3\,\frac{1}{b}\,\frac{1}{c}\,a_{1,3}\,\varphi_z,\\ \dot{y} &= b\,a_{2,1}\,\varphi_x + d\,a_{2,2}\,y,\\ \dot{\varphi}_z &= c\,a_{3,0} + d\,a_{3,3}\,\varphi_z + b\,a_{3,6}\,\varphi_x\,\varphi_z,\end{aligned} \tag{2.4}$$

where the functions $\varphi_x = \varphi_x(x, z)$ and $\varphi_z = \varphi_z(x, z)$ are generic linear independent functions of x and z (see Lainscsek et al., 2003, and Lainscsek, 2011, for details). The parameters b and c specify the structure of the parameter space $a_{i,*}$ and d is an additional time-scaling parameter.

Such a reconstruction is possible only for simple dynamic systems with no noise and can be seen as the theoretical limit for global modeling. It also shows that measurements from the same dynamical system on different timescales (and consequently frequency shifted) can be identified as belonging to the same dynamical system. Further, the time-scaling parameter can be reconstructed as a separate parameter. This is consistent with Takens's theorem, which does not specify the sampling rate.

Since real-world data are noisy and generally originate from highly complex dynamical systems, global modeling from data is in general not feasible. Often it is sufficient to detect dynamical differences between data classes and quantify the difference. DDA can distinguish between heart conditions in EKG (electrocardiography) data (Lainscsek & Sejnowski, 2013) and EEG data between Parkinson's patients and control subjects (Lainscsek, Weyhenmeyer, Hernandez, Poizner, & Sejnowski, 2013). Dynamical differences can be detected and analyzed by combining derivative embeddings with delay embeddings as functional embeddings in the DDA framework.

DDA in the time domain can be related to frequency analysis as shown in the next section. In linear DDEs (Falbo, 1995), the estimated parameters

and delays of a DDE relate to the frequencies of a signal, and in nonlinear DDEs, the estimated parameters are connected to higher-order statistical moments.

3 Functional Embedding as a Connection Between Nonlinear Dynamics and Frequency Analysis

3.1 Linear DDE and Frequency Analysis. Lainscsek & Sejnowski (2013) show that the linear DDE (Falbo, 1995),

$$\dot{x} = a x_\tau, \tag{3.1}$$

where $x_\tau = x(t-\tau)$, can be used to detect frequencies in the time domain: equation 3.1 with a harmonic signal with one frequency, $x(t) = A\ \cos(\omega t + \varphi)$ can be written as

$$\begin{aligned} &-aA\cos(\omega(t-\tau)+\varphi) - A\omega\sin(t\omega+\varphi) = 0 \\ &\sin(\omega t)(-aA\cos(\varphi)\sin(\tau\omega) + aA\sin(\varphi)\cos(\tau\omega) \\ &\quad - A\omega\cos(\varphi)) + \cos(\omega t)(-aA\sin(\varphi)\sin(\tau\omega) \\ &\quad - aA\cos(\varphi)\cos(\tau\omega) - A\omega\sin(\varphi)) = 0. \end{aligned} \tag{3.2}$$

This equation has the solutions

$$\tau = \frac{(2n-1)\pi}{2\omega} \overset{\omega=2\pi f}{=} \frac{2n-1}{4f}, \quad a = (-1)^n\omega, \quad n \in \mathbb{N}. \tag{3.3}$$

This means that the delays are inversely proportional to the frequency f and the coefficient a is directly proportional to the frequency. For linear DDEs,

$$\dot{x} = \sum_{k=1}^{N} a_k x_{\tau_k}, \tag{3.4}$$

a special solution is a harmonic signal with $2N-1$ frequencies,

$$x(t) = \sum_{k=1}^{2N-1} A_k\ \cos(\omega_k t + \varphi_k), \tag{3.5}$$

and

$$\tau_i = \frac{\pi(2n-1)}{2\omega_j}, \quad n \in \mathbb{N}, \tag{3.6}$$

where all delays τ_i are related to one of the frequencies. The expressions for the coefficients a are more complicated than in equation 3.3, and each of the coefficients a depends on all the frequencies in the signal.

Equation 3.1 can be expanded as a Yule-Walker equation (Yule, 1927; Walker, 1931; Boashash, 1995; Kadtke & Kremliovsky, 1999), $\dot{x}\, x_\tau = a\, x_\tau^2$. Applying the expectation operator $\langle F(t)\rangle \equiv \lim_{T\to\infty}(\frac{1}{T}\int_0^T F(t)\ \mathrm{d}t)$, we get

$$a = \frac{\langle \dot{x}\, x_\tau\rangle}{\langle x_\tau^2\rangle} = \frac{\langle \dot{x}\, x_\tau\rangle}{\langle x^2\rangle}. \tag{3.7}$$

For a harmonic signal with one frequency (this is a special solution of the linear DDE in equation 3.1; see Falbo, 1995, for details), $x(t) = A\ \cos(\omega\, t + \varphi)$. This becomes:

$$\begin{aligned} \langle \dot{x}\, x_\tau\rangle &= -\frac{1}{2}\omega A^2 \sin(\omega\tau) \\ & \qquad\qquad\qquad\qquad \Rightarrow \quad a = -\omega \sin(\omega\,\tau) \\ \langle x^2\rangle &= \frac{1}{2}A^2 \end{aligned} \tag{3.8}$$

and for $\sin(\omega\,\tau) = \pm 1$ the special solutions for a and τ are

$$\tau = \frac{(2\,n-1)\,\pi}{2\,\omega} \overset{\omega=2\pi f}{=} \frac{2\,n-1}{4\,f}, \quad a = (-1)^n\,\omega, \tag{3.9}$$

the same as in equation 3.3. For DDEs without explicit analytical solutions, Yule-Walker equations can be applied.

The numerator in equation 3.7 can be rewritten as delay derivatives of the autocorrelation function in the case of a bounded stationary signal,

$$\begin{aligned} \langle \dot{x}x_\tau\rangle &= \lim_{T\to\infty}\frac{1}{T}\int_0^T \dot{x}x_\tau\ dt = \lim_{T\to\infty}\frac{1}{T}\left[xx_\tau\right]_0^T - \lim_{T\to\infty}\frac{1}{T}\int_0^T x\frac{dx_\tau}{dt}\ dt = \\ &= \lim_{T\to\infty}\frac{1}{T}\int_0^T x\frac{dx_\tau}{d\tau}\ dt = \frac{d}{d\tau}\langle xx_\tau\rangle; \end{aligned} \tag{3.10}$$

therefore,

$$a = \frac{\langle \dot{x}\, x_\tau\rangle}{\langle x^2\rangle} = \frac{\frac{d}{d\tau}\langle x\, x_\tau\rangle}{\langle x^2\rangle}. \tag{3.11}$$

Thus, the numerator in equation 3.11 is the delay derivative of the autocorrelation function $\langle x\, x_\tau\rangle$. For a harmonic signal $x(t) = \sum_{k=1}^N A_k\ \cos(\omega_k\, t + \varphi_k)$

with N frequencies, it is $\langle x\, x_\tau \rangle = \frac{1}{2} \sum_{k=1}^{N} A_k^2 \cos(\omega_k \tau)$ and is used for pitch detection (Boersma, 1993; Roads & Curtis, 1996).

The denominator $\langle x^2 \rangle$ will be used as frequency detector in the following way: Consider a signal

$$x(t) = \underbrace{\sum_{i=1}^{N} A_i \cos(\omega_i t + \varphi_i)}_{\mathcal{S}} + \underbrace{D \cos(\Omega t + \phi)}_{\mathcal{P}}, \tag{3.12}$$

where $\mathcal{S}$ is the signal under investigation and $\mathcal{P} = D \cos(\Omega t + \phi)$ is a probe signal. The expectation value $\langle x^2 \rangle = \lim_{T\to\infty} \frac{1}{T} \int_0^T x^2 \, \mathrm{dt}$ can be written in the form

$$\begin{aligned}
\langle x^2 \rangle &= \langle (\mathcal{S} + \mathcal{P})^2 \rangle = \langle \mathcal{S}^2 \rangle + 2\langle \mathcal{S}\mathcal{P} \rangle + \langle \mathcal{P}^2 \rangle, \\
\langle \mathcal{S}^2 \rangle &= \frac{1}{2} \sum_{i=1}^{N} A_i^2, \\
\langle \mathcal{P}^2 \rangle &= \frac{1}{2} D^2, \\
\langle \mathcal{S}\mathcal{P} \rangle &= \begin{cases} 0 & \text{for} \quad \omega_i \neq \Omega \\ A_i D \cos(\varphi_i - \phi) & \phantom{\text{for}} \quad \omega_i = \Omega \end{cases}.
\end{aligned} \tag{3.13}$$

The term $\langle \mathcal{S}\mathcal{P} \rangle = \langle \mathcal{S}\, D \cos(\Omega t + \phi) \rangle$ combines $\mathcal{R} = \langle \mathcal{S} \cos(\Omega t) \rangle$ and $\mathcal{I} = \langle \mathcal{S} \sin(\Omega t) \rangle$ used in discrete frequency analysis.

3.1.1 Time Domain Spectrum. Let the amplitude of the probe signal be $D = 1$ (arbitrary parameter). Then one way to define the time domain spectrum (TDS) is

$$L(\Omega) = \max_{\phi} \left(\langle x^2 \rangle - \langle \mathcal{S}^2 \rangle - \frac{1}{2} \right). \tag{3.14}$$

$L(\Omega)$ is zero for frequencies Ω that are not part of the signal $\mathcal{S}$ and equal to the amplitude A_i of a frequency $f_i = \frac{\omega_i}{2\pi}$ if $\Omega = \omega_i$, which follows directly from $\langle x^2 \rangle = \langle \mathcal{S}^2 \rangle + \frac{1}{2} + A_i \cos(\varphi_i - \phi)$. The amplitude A_i is always positive, and $\cos(\varphi_i - \phi)$ ranges between -1 and 1 with a maximum when $\phi = \varphi_i$. For a given signal $\mathcal{S}$, we loop over a range of ϕ values between 0 and 2π and take the maximal value of $\langle x^2 \rangle - \langle \mathcal{S}^2 \rangle - \frac{1}{2} = A_i$. The value ϕ that maximizes $\cos(\varphi_i - \phi)$ is then equal to the phase φ_i. Therefore, this simultaneously detects the amplitudes A_i and their phases φ_i in a signal. This method can

detect frequencies and couplings only at lower than half the sampling rate since $\Omega = \frac{2\pi f}{f_s}$, where f_s is the sampling rate, in $\cos(\Omega t + \phi)$ with the time t given in time steps needs to be smaller than π. This is the same as the Nyquist frequency limit for traditional methods (Nyquist, 1928).

3.1.2 Comparison to the Goertzel Algorithm and Discrete Frequency Analysis. In discrete frequency analysis for a signal $\mathcal{S}(t) = \sum_{i=1}^{N} A_i \cos(\omega_i t + \varphi_i)$, the expectation values $\mathcal{R} = \langle \mathcal{S} \cos(\Omega t) \rangle$ and $\mathcal{I} = \langle \mathcal{S} \sin(\Omega t) \rangle$ are computed. $\mathcal{R}$ and $\mathcal{I}$ are nonzero only if Ω is contained in the signal. The amplitude of that frequency is then $A_i \overset{\Omega=\omega_i}{=} 4\sqrt{\mathcal{R}^2 + \mathcal{I}^2}$, and the phase is $\varphi_i = \tan^{-1} \frac{\mathcal{I}}{\mathcal{R}}$. The Goertzel algorithm (Goertzel, 1958; Jacobsen & Lyons, 2003) is the numerical implementation of this principle and gives the same results as the computation of the TDS.

The TDS $L(\Omega)$ differs from discrete frequency analysis in two ways. First, $D\cos(\Omega t + \phi)$ is multiplied with the signal $\mathcal{S}$ instead of $\cos(\Omega t)$. Second, only $\langle \mathcal{SP} \rangle = \langle \mathcal{S} D \cos(\Omega t + \phi) \rangle$ is computed instead of $\mathcal{R} = \langle \mathcal{S} \cos(\Omega t) \rangle$ and $\mathcal{I} = \langle \mathcal{S} \sin(\Omega t) \rangle$. Amplitudes and phases are extracted from $\langle \mathcal{SP} \rangle$ instead of $A_i \overset{\Omega=\omega_i}{=} 2\sqrt{\mathcal{R}^2 + \mathcal{I}^2}$ and $\varphi_i = \tan^{-1} \frac{\mathcal{I}}{\mathcal{R}}$. An advantage of $L(\Omega)$ is that it can be more easily expanded to the computation of higher-order spectra (e.g., the bispectrum).

3.2 Nonlinear DDE and Bispectral Analysis

3.2.1 Quadratic Phase and Frequency Couplings. Frequency components do not always appear completely independent of one another. Nonlinear interactions of frequencies and their phases (e.g., quadratic phase coupling) cannot be detected by a power spectrum—the Fourier transform of the autocorrelation function (second-order cumulant)—since phase relationships and frequency couplings of signals are lost. Such couplings are usually detected by bispectral analysis (Tukey, 1953; Kolmogorov & Rozanov, 1960; Leonov & Shiryeav, 1959; Rosenblatt & Van Ness, 1965; Brillinger & Rosenblatt, 1967; Swami, 2003; Mendel, 1991; Fackrell & McLaughlin, 1995). The bispectrum or bispectral density is the Fourier transform of the third-order cumulant (bicorrelation function).

Consider the signal $x(t) = A_1 \cos(\omega_1 t + \varphi_1) + A_2 \cos(\omega_2 t + \varphi_2)$, which is passed through a quadratic nonlinear system $h(t) = b\,x^2(t)$ where b is a nonzero constant. The output of the system will include the harmonic components: $(2\omega_1, 2\varphi_1)$, $(2\omega_2, 2\varphi_2)$, $(\omega_1 + \omega_2, \varphi_1 + \varphi_2)$, and $(\omega_1 - \omega_2, \varphi_1 - \varphi_2)$. These phase relations are called quadratic phase coupling (QPC) in general and quadratic frequency coupling (QFC) when the phases are zero ($\varphi_1 = \varphi_2 = 0$). For the signal $x(t) = \cos(\omega_1 t) + \cos(\omega_2 t) + \cos(\omega_3 t)$, QFC occurs when ω_3 is a multiple of one of the frequencies or of the sum or difference of the two frequencies.

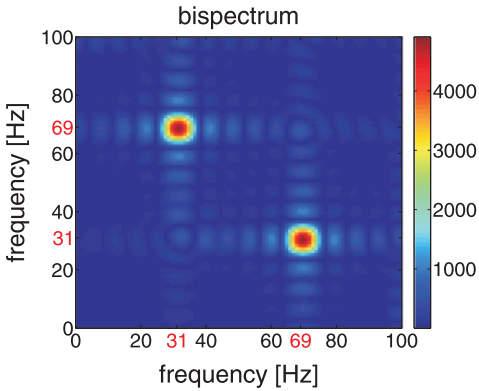

Figure 2: Bispectrum magnitude (see equation 3.15) for the signal $\mathcal{S} = \sum_{i=1}^{3} A_i \cos(\omega_i t + \varphi_i) + \cos((\omega_1 + \omega_3)t + (\varphi_1 + \varphi_3))$ with the frequencies, amplitudes, and phases listed in Table 1.

Table 1: Frequencies f_i, Amplitudes A_i, and Phases φ_i for the Numerical Experiments Shown in Figures 2 and 3.

i	f_i	A_i	φ_i
1	31	1.14	0.53
2	40	0.92	1.14
3	69	1.13	0.09

The third-order cumulant is $C_3(\tau_1, \tau_2) = \langle x(t)x(t-\tau_1)x(t-\tau_2)\rangle$, and the traditional bispectrum is the double Fourier transform of C_3,

$$B_x(f_1, f_2) = \mathcal{F}^\star(f_1 + f_2)\,\mathcal{F}(f_1)\,\mathcal{F}(f_2), \tag{3.15}$$

where $\mathcal{F}(f)$ is the Fourier transform of $x(t)$. An example is given in Figure 2.

3.2.2 Nonlinear DDE. The simplest nonlinear DDE,

$$\dot{x} = a x_\tau^2, \tag{3.16}$$

where $x_\tau = x(t-\tau)$, extends linear spectral analysis to quadratic coupling analysis or bispectral analysis. This equation can be solved in the same way as the linear DDE:

$$a = \frac{\langle \dot{x}\, x_\tau^2\rangle}{\langle x_\tau^4\rangle} = \frac{\langle \dot{x}\, x_\tau^2\rangle}{\langle x^4\rangle}. \tag{3.17}$$

For a harmonic signal with three frequencies $x(t) = A_1 \cos(\omega_1 t + \varphi_1) + A_2 \cos(\omega_2 t + \varphi_2) + A_3 \cos(\omega_3 t + \varphi_3)$, the coefficient a is nonzero only if there is quadratic coupling between the frequencies: $\omega_3 = \omega_1 \pm \omega_2, 2\omega_1, 2\omega_2, \frac{\omega_1}{2}$, or $\frac{\omega_2}{2}$. Only the numerator $\langle \dot{x}\, x_\tau^2 \rangle$ contains quadratic coupling information and can be rewritten for a bounded stationary signal:

$$\langle \dot{x} x_\tau^2 \rangle = \lim_{T\to\infty} \frac{1}{T} \int_0^T \dot{x} x_\tau^2 \, dt = \lim_{T\to\infty} \frac{1}{T} \left[x x_\tau^2 \right]_0^T - \lim_{T\to\infty} \frac{1}{T} \int_0^T x \frac{d x_\tau^2}{dt} \, dt =$$
$$= \lim_{T\to\infty} \frac{1}{T} \int_0^T x \frac{d x_\tau^2}{d\tau} \, dt = \frac{d}{d\tau} \langle x x_\tau^2 \rangle. \tag{3.18}$$

The denominator in equation 3.17 is a constant independent of up to quadratic couplings in the signal. Therefore equation 3.17 can be rewritten as

$$a \langle x^4 \rangle = \langle \dot{x}\, x_\tau^2 \rangle$$
$$\frac{d}{d\tau} \langle x x_\tau^2 \rangle = \text{const.} \tag{3.19}$$

For $x(t) = \sum_{i=1}^{\infty} A_i \cos(\omega_i t + \varphi_i) + A_{j,k} \cos((\omega_j + \omega_k) t + \varphi_j + \varphi_k)$, the expectation value is

$$\langle x x_\tau^2 \rangle = \frac{1}{2} A_j A_k A_{j,k} \left(\cos(\omega_j \tau) + \cos(\omega_k \tau) + \cos((\omega_j + \omega_k)\tau) \right). \tag{3.20}$$

Therefore, $\langle x x_\tau^2 \rangle$ is zero for a signal without QFC ($A_{j,k} = 0$) and nonzero if there is QFC in the signal. Without loss of generality, the delay τ can be set to zero in equation 3.20 and $\langle x x_\tau^2 \rangle \overset{\tau=0}{=} \langle x^3 \rangle = \frac{3}{2} A_j A_k A_{j,k}$. $\langle x^3 \rangle$, and the coefficient a of the nonlinear DDE in equation 3.16 is then nonzero only if QFC is present in the signal ($A_{j,k} \neq 0$) and can be used as a bispectral detector. $\langle x^3 \rangle$ can detect the presence or absence of QFC but not the frequencies involved. To determine the frequencies, the time domain bispectrum is introduced.

3.2.3 Time Domain Bispectrum. To determine the frequencies of QFC, consider a signal

$$x(t) = \mathcal{S} + \mathcal{P},$$
$$\mathcal{S} = \sum_{i=1}^{N} A_i \cos(\omega_i t + \varphi_i) + A_{j,k} \cos((\omega_j + \omega_k)t + (\varphi_j + \varphi_k)),$$
$$\mathcal{P} = D \cos(\Omega t + \phi), \tag{3.21}$$

where $\mathcal{S}$ is the signal under investigation and $\mathcal{P} = D\cos(\Omega\, t + \phi)$ is a probe signal. The frequencies f_j and f_k are coupled. $\langle x^3 \rangle$ is then

$$\begin{aligned}
\langle x^3 \rangle &= \langle \mathcal{S}^3 \rangle + 3\,\langle \mathcal{S}^2\,\mathcal{P} \rangle + 3\,\langle \mathcal{S}\,\mathcal{P}^2 \rangle + \langle \mathcal{P}^3 \rangle, \\
\langle \mathcal{S}^3 \rangle &= \frac{3}{2}\,A_{j,k}\,A_j A_k, \\
\langle \mathcal{S}^2\,\mathcal{P} \rangle &= \begin{cases} \frac{1}{2} A_j A_k D\cos(\phi - \varphi_j \mp \varphi_k) & \Omega = \omega_j \pm \omega_k \\ \frac{1}{2} A_k\,A_{j,k} D\cos(\phi - \varphi_j) & \Omega = \omega_j \\ \frac{1}{4} A_i^2 D\cos(\phi - 2\varphi_i) & \Omega = 2\omega_i \\ 0 & \Omega \text{ otherwise} \end{cases}, \\
\langle \mathcal{S}\,\mathcal{P}^2 \rangle &= \begin{cases} \frac{1}{4} A_i D^2\cos(2\phi - \varphi_i) & \Omega = \frac{\omega_i}{2} \\ 0 & \Omega \text{ otherwise} \end{cases}, \\
\langle \mathcal{P}^3 \rangle &= 0.
\end{aligned} \tag{3.22}$$

The term $\langle \mathcal{S}\,\mathcal{P}^2 \rangle$ is nonzero only when $\widetilde{\Omega} = 2\,\Omega = \omega_i$. It therefore can be used as an alternative definition of the time domain spectrum TDS in equation 3.14:

$$L(\widetilde{\Omega}) = \max_{\phi} \langle \mathcal{S}\,\mathcal{P}^2 \rangle. \tag{3.23}$$

To define the time domain bispectrum (TDB), we set in our analysis $D = 1$:

$$B(\Omega) = L(\widetilde{\Omega}) \cdot \max_{\phi} \langle \mathcal{S}^2\,\mathcal{P} \rangle. \tag{3.24}$$

The expression $\max\langle \mathcal{S}^2\,\mathcal{P} \rangle$ is zero for frequencies Ω that are not coupled and non-zero according to equation 3.22 for the coupling frequencies $\omega_j \pm \omega_k$, ω_j, ω_k, and $2\omega_i$. To get only a nonzero expression for $\omega_j + \omega_k$, ω_j, and ω_k, we use $L(\widetilde{\Omega})$ in equation 3.24 to filter out $2\omega_i$.

Note that unlike the bispectrum, which is a function of two frequencies (see Figure 2), the TDB is a function of a single frequency. In Figure 3, the TDB is shown for the same signal used to illustrate the traditional bispectrum in Figure 2. Note that the 40 Hz component, which is not coupled, is absent in both the TBS and the bispectrum but is present in the TDS.

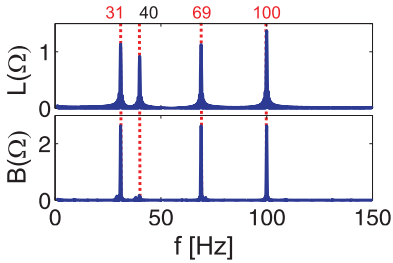

Figure 3: Time domain spectrum TDS $L(\Omega)$ and time domain bispectrum TDB $B(\Omega)$ for the signal $\mathcal{S} = \sum_{i=1}^{3} A_i \cos(\omega_i t + \varphi_i) + \cos((\omega_1 + \omega_3)t + (\varphi_1 + \varphi_3))$ with the frequencies, amplitudes, and phases listed in Table 1.

3.3 General Nonlinear DDE as Generalized Nonlinear Spectral Tool. Combining linear and nonlinear terms in a DDE makes all coefficients nonlinear. If we, for example, combine the DDEs of sections 3.1 and 3.2,

$$\dot{x} = a_1 x_\tau + a_2 x_\tau^2, \tag{3.25}$$

the solution can be written as

$$\begin{aligned} a_1 &= \frac{\langle x_\tau^4\rangle\langle x_\tau \dot{x}\rangle - \langle x_\tau^3\rangle\langle x_\tau^2 \dot{x}\rangle}{\langle x_\tau^2\rangle\langle x_\tau^4\rangle - \langle x_\tau^3\rangle^2} = \frac{\langle x^4\rangle \frac{d}{d\tau}\langle x_\tau x\rangle - \langle x^3\rangle \frac{d}{d\tau}\langle x_\tau^2 x\rangle}{\langle x^2\rangle\langle x^4\rangle - \langle x^3\rangle^2}, \\ a_2 &= \frac{\langle x_\tau^2\rangle\langle x_\tau^2 \dot{x}\rangle - \langle x_\tau^3\rangle\langle x_\tau \dot{x}\rangle}{\langle x_\tau^2\rangle\langle x_\tau^4\rangle - \langle x_\tau^3\rangle^2} = \frac{\langle x^2\rangle \frac{d}{d\tau}\langle x_\tau^2 x\rangle - \langle x^3\rangle \frac{d}{d\tau}\langle x_\tau x\rangle}{\langle x^2\rangle\langle x^4\rangle - \langle x^3\rangle^2}. \end{aligned} \tag{3.26}$$

The coefficient a_1 as well as the coefficient a_2 contain both linear and nonlinear statistical moments. The moments and coefficients for combinations of sinusoids are given in Table 2 with and without couplings.

3.3.1 Example 1: Data with No Couplings. To explore the properties of equation 3.26, we start with a harmonic signal with four independent frequencies and no couplings:

$$x(t) = \sum_{k=1}^{4} A_k \cos(\omega_k t + \varphi_k). \tag{3.27}$$

Using equation 3.26, we get for the two coefficients,

$$\begin{aligned} a_1 &= \frac{3}{16}\left(A^{(1)} + 4A_2^2A_3^2 + 4A_4^2A_3^2 + 4A_2^2A_4^2\right)\sum_{k=1}^{4} A_k^2\omega_k \sin\left(\tau\omega_k\right), \\ a_2 &= 0, \end{aligned} \tag{3.28}$$

Table 2: Moments and Coefficients of Equations 3.25 and 3.26.

	Couplings Present	No Couplings
	$x(t) = \sum_{k=1}^{N} A_k \cos(\omega_k t + \varphi_k) + \sum_{i=1}^{N} \sum_{j=i+1}^{N} (A_{i,j} \cos((\omega_i + \omega_j) t + (\varphi_i + \varphi_j)))$	$x(t) = \sum_{k=1}^{N} A_k \cos(\omega_k t + \varphi_k)$
$\langle x^2 \rangle$	Const.	Const.
$\langle x^3 \rangle$	Const.	0
$\langle x^4 \rangle$	Const.	Const.
$\langle x_\tau x \rangle$	$f(\omega_{i,j,k}, \tau)$	$f(\omega_k, \tau)$
$\langle x_\tau^2 x \rangle$	$f(\omega_{i,j,k}, \tau)$	0
a_1	$f(\omega_{i,j,k}, \tau)$	$f(\omega_k, \tau)$
a_2	$f(\omega_{i,j,k}, \tau)$	0

Note: The coefficient a_2 is nonzero only if QFC is present in the data.

where $A^{(1)} = A_1^4 + A_2^4 + A_3^4 + A_4^4 + 4A_1^2 \left(A_2^2 + A_3^2 + A_4^2\right)$. The nonlinear coefficient a_2 vanishes in this case.

3.3.2 Example 2: Data with Couplings. Replace the fourth term in equation 3.27 with a coupling frequency,

$$x(t) = \sum_{k=1}^{3} A_k \cos(\omega_k t) + A_4 \cos((\omega_2 + \omega_3) t + (\varphi_2 + \varphi_3)). \tag{3.29}$$

The two coefficients are:

$$\begin{aligned}
a_1 = {} & \frac{3}{16}((A^{(1)} + 4A_2^2A_3^2 + 4A_4^2A_3^2 + 4A_2^2A_4^2)A_1^2\omega_1 \sin(\tau\omega_1) \\
& + (4A_2^2A_3^2 + A^{(1)})(A_2^2\omega_2 \sin(\tau\omega_2) + A_3^2\omega_3 \sin(\tau\omega_3)) \\
& + 4A_4^2(A_2^4\omega_2 \sin(\tau\omega_2) + A_3^4\omega_3 \sin(\tau\omega_3)) \\
& + A_4^2(\omega_2 + \omega_3)(4A_2^2A_4^2 + 4A_3^2A_4^2 + A^{(1)})(\sin(\tau\omega_2)\cos(\tau\omega_3) \\
& + \sin(\tau\omega_3)\cos(\tau\omega_2)), \\
a_2 = {} & \frac{1}{4}A^{(2)}(-3A_1^2\omega_1 \sin(\tau\omega_1) \\
& + (A_1^2 - 2A_2^2 + A_3^2 + A_4^2)\omega_2 \sin(\tau\omega_2) \\
& + (A_1^2 + A_2^2 - 2A_3^2 + A_4^2)\omega_3 \sin(\tau\omega_3) \\
& + (A_1^2 + A_2^2 + A_3^2 - 2A_4^2)(\omega_2 + \omega_3)(\sin(\tau\omega_2)\cos(\tau\omega_3) \\
& + \sin(\tau\omega_3)\cos(\tau\omega_2))),
\end{aligned} \tag{3.30}$$

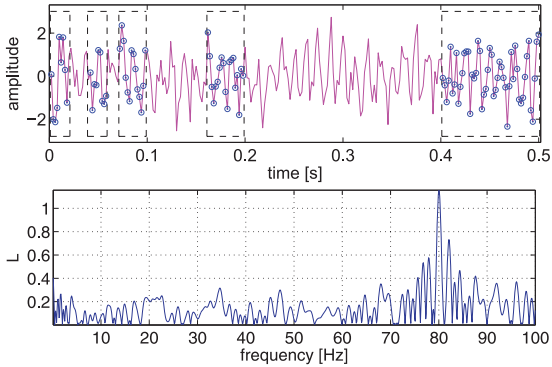

Figure 4: Frequency detection from sparse data. Only 101 data points circled from a signal $\mathcal{S} = \cos(2\pi f t)$ with $f = 80$ Hz and a sampling rate of 500 Hz and added white noise (SNR = 0 dB) (upper plot) were used to detect the frequency f (lower plot).

where $A^{(1)} = A_1^4 + A_2^4 + A_3^4 + A_4^4 + 4A_1^2\left(A_2^2 + A_3^2 + A_4^2\right)$ and $A^{(2)} = A_2 A_3 A_4$.

Equation 3.25 is therefore a QFC detector. For any signal without QFC, the coefficient a_2 will vanish. The detection of nonlinear couplings in the data is important in underwater acoustics (Nikias & Raghuver, 1987).

4 Sparse Data

Real-world data often contain data segments that cannot be used for the analysis, such as artifacts in electroencephalography (EEG) data. Equation 3.14 can nonetheless be used even when many data points are missing. For a sparse signal $x(T)$, where T is the vector of times for which the signal is good, the probe signal $\cos(\Omega T + \phi)$ can be used to detect the spectrum or any higher-order spectra. Equation 3.12 can be modified as

$$x(T) = \underbrace{\sum_{i=1}^{N} A_i \cos(\omega_i T + \varphi_i)}_{\mathcal{S}} + \underbrace{\cos(\Omega T + \phi)}_{\mathcal{P}}, \tag{4.1}$$

where T is a sparse time vector. The upper plot in Figure 4 shows an example where only a few data points of a signal were used. Those data and their time points were then used to compute the TDS using equation 3.14. The TDS for these sparse data is shown in the lower plot of Figure 4. Note that the sparse data have to satisfy the restrictions of the Nyquist theorem (Nyquist, 1928). Equation 3.24 can be used for the bispectrum of sparse data in the same way.

5 Cross-Trial Spectrogram and Bispectrogram

The same method can be used when very short segments of data need to be analyzed and there are multiple trials of the same experiment, as in event-related potential (ERP) analysis (Davis, 1939; Davis & Davis, 1936; Sutton, Braren, Zubin, & John, 1965; Luck, 2014). There are two different ways to define a cross-trial spectrogram (CTS). The first assumes that there is cross-trial phase coherence; the second does not.

5.1 Phase Coherence Across Trials. The first method assumes phase coherence across trials. Consider a signal $x_1(T_1), x_2(T_2), \ldots, x_n(T_n)$ where all time series $x_i(T_i)$ are centered around the same event (e.g., a stimulus S for EEG data). Then all time vectors T_i can be considered equal and the signal can be concatenated into a single vector, $x_1(T), x_2(T), \ldots, x_n(T)$; $L_i(\Omega)$ is computed for this single vector. For a short data window that would be too short for spectral analysis, the data of all trials can be combined and the spectrum can be computed by using a probe signal with a time vector that consists of n repetitions of the time vector T for n data segments. In this manner, a cross-trial spectrogram can be computed by using short sliding windows when there is some phase coherence in the relevant frequencies.

To test this, we generated a signal $\mathcal{S}_i = \mathcal{W}_i + \mathcal{Y}_i$, where $\mathcal{W}$ is white noise and $\mathcal{Y}_i$ are short-phase coherent (phases varied in a range from 0 to π for each trial) data segments of 150 ms (t =[126, 279] ms, [350, 500] ms, [576, 726] ms, and [800, 950] ms). The frequencies in the first segment are 10, 20, and 30 Hz (QFC); the single frequencies in the following three segments are 20, 30, and 24 Hz, respectively. Twenty-one second data segments ($i = 1, 2, \ldots, 20$) at a sampling rate of 500 Hz were generated to simulate data from repeated trials in an experiment.

Figure 5a shows the mean signal of these 20 trials. Sliding windows of 100 ms (50 data points) with a window shift of 10 ms (5 data points) were used to compute the CTS in Figure 5c. This CTS was compared with traditional Morlet wavelet analysis (Mallat, 2008) in Figure 5b.

In a companion paper (Lainscsek et al., 2015), we apply this method to EEG data.

5.2 No Phase Coherence Across Trials. The second cross-trial method does not assume phase coherence across trials and is defined by computing $L_i(\Omega)$ in equation 3.14 for each trial using sliding windows and then averaging over the 20 spectra $L_i(\Omega)$ ($i = 1, 2, \ldots, 20$).

Figure 5c shows the CTS assuming phase coherence in the signal, and Figure 5d shows the phase-independent CTS. All three methods identified the frequencies correctly, and both CTS methods produce similar results (the correlation coefficient between Figures 5c and 5d is 0.96).

The same experiment was repeated for completely phase-randomized data segments $\mathcal{Y}_i$. Figure 6a shows the mean of the data, Figure 6b the

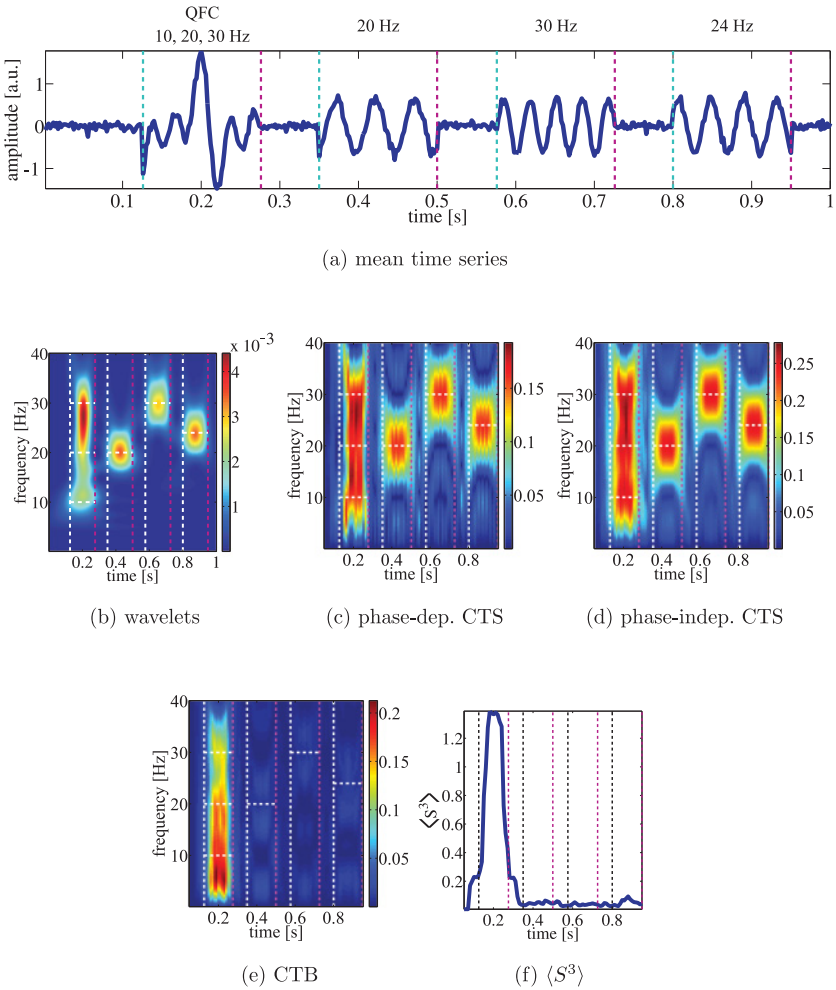

Figure 5: Cross-trial spectrogram (CTS) of phase-coherent simulated data with white noise. Four segments of noisy short signals of 10, 20, 24, and 30 Hz with some phase coherence added to the data. Horizontal white dashed lines indicate frequencies present in each segment. (a) Mean over the 20 trials. (b) Wavelet spectrogram. (c) Phase-dependent CTS. (d) Phase-independent CTS. (e) Bispectrogram CTB. (f) bispectral detector $\overline{|\langle S^3\rangle|}$.

traditional Morlet wavelet analysis, Figure 6c the CTS assuming phase coherence, and Figure 6d the phase-independent CTS. Only the phase-independent CTS can identify the frequencies correctly. The correlation coefficient between Figures 6c and 6d drops to 0.15. Therefore, the

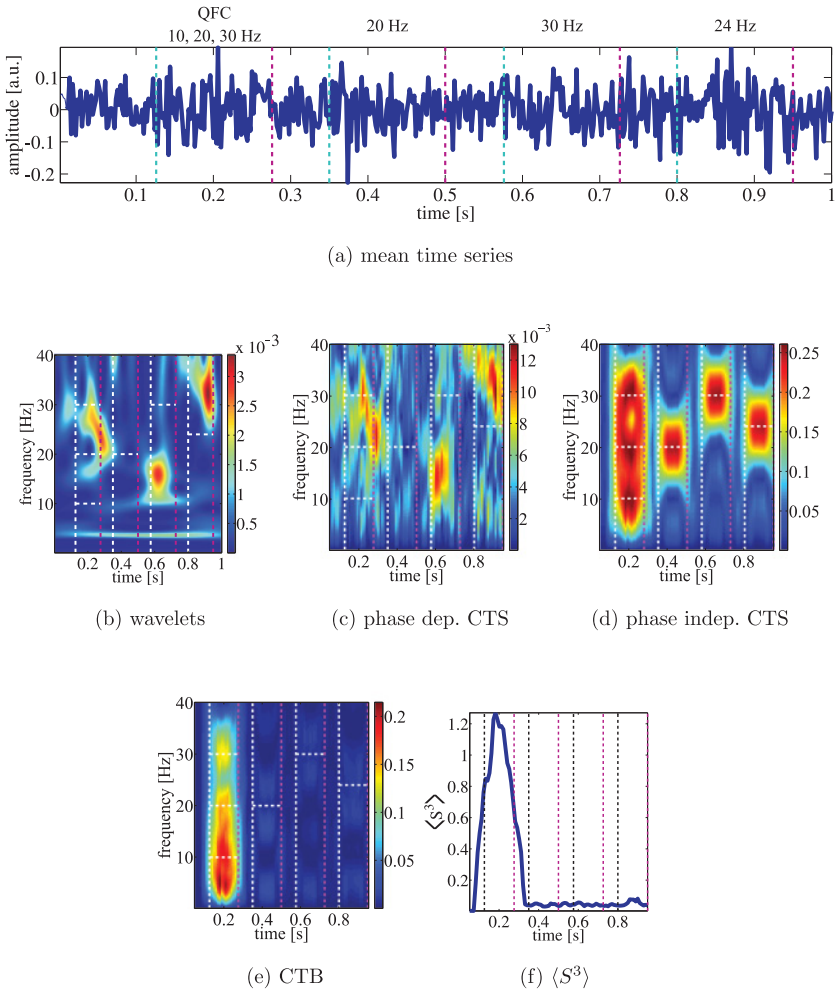

Figure 6: Cross-trial spectrogram (CTS) of nonphase-coherent simulated data with white noise. Four segments of noisy short signals of 10, 20, 24, and 30 Hz with no phase coherence added to the data. (a) The mean over the 20 trials. (b) The wavelet spectrogram. (c) The phase-dependent CTS. (d) The phase-independent CTS. (e) The bispectrogram CTB. (f) The bispectral detector $|\overline{\langle S^3 \rangle}|$.

correlation coefficient between the phase-dependent CTS and the phase independent CTS can be used to quantify the amount of phase coherence.

If multiple channels are available, data from different channels or different trials can be combined. Such cross-trial spectra are different from

the cross-spectral density (Bracewell, 1965; Papoulis, 1962) based on the cross-correlation between two signals.

Since the TDB in equation 3.24 is a function of a single frequency (one-dimensional plot instead of a two-dimensional plot for the traditional bispectrum), the TDB bispectrogram can be obtained in the same way as a spectrogram. The single-trial, cross-trial, or cross-channel TDS (or spectrogram) can be extended to a single-trial, cross-trial, or cross-channel bispectrogram (CTB) by replacing $L(\Omega)$ with $B(\Omega)$. The bispectral detector $\langle S^3 \rangle$ in Figures 5f and 6f indicates the presence of QFC in the first data segment, and CTB in Figures 5e and 6e show couplings in the first data segment with the three coupled frequencies 10, 20, and 30 Hz ($10 + 20 = 30$) present. There is no frequency coupling in other data segments.

6 Discussion

Delay differential analysis is a nonlinear dynamical time series classification and detection tool based on embedding theory. It combines aspects of delay and differential embeddings in the form of delay differential equations (DDEs), which can detect distinguishing dynamical information in data. We focused here on the connection between traditional frequency analysis and linear DDEs and between higher-order statistics and nonlinear DDEs.

We introduced a new set of time domain tools to analyze the frequency content and frequency coupling in signals: the time domain spectrum (TDS), the time domain bispectrum (TDB), the time domain cross-trial (or cross-channel) spectrogram (CTS), and the time domain cross-trial (or cross-channel) bispectrogram (CTB). In addition, the CTS has a phase-dependent and a phase-independent realization.

These new time domain methods are complementary to traditional frequency domain methods and have several advantages:

- Because these new DDA spectral methods use time itself as a variable, they can be applied cross-trial and on sparse data.
- DDA can be applied to short data segments and therefore has an improved time resolution compared with the fast Fourier transform, which was developed to be computationally fast on longer time series.
- Because some numerical components of $\frac{1}{f}$ noise are not present, the new time domain spectrogram introduced here does not need normalization across frequencies. The sources of $\frac{1}{f}$ noise will be discussed elsewhere.
- DDA is noise insensitive (Lainscsek, Weyhenmeyer et al. (2013)). Short data segments from the Rössler system with a signal-to-noise ratio between 10 dB and -5 dB (more noise than data) were well separated. This separation was also correlated to the dynamical bifurcation parameter and generalized well to new data.

- DDA can discriminate dynamical differences in data classes using a small set of features, thus avoiding overfitting. With DDA, only five features are sufficient for many applications, whereas traditional methods often rely on many more features. The goal is not to achieve the most predictive model but the most discriminative model.
- Higher-order nonlinearities are intrinsically included in DDA, in comparison with nonlinear extensions of spectral analysis, which are not straightforward and are complicated.

There are many applications of DDA, including classification of electrocardiograms (Lainscsek & Sejnowski, 2013), electroencephalograms (Lainscsek, Hernandez, Weyhenmeyer, Sejnowski, & Poizner, 2013), speech (Gorodnitsky & Lainscsek, 2004), and sonar time series (Lainscsek & Gorodnitsky, 2003).

Acknowledgments

We acknowledge support by the Howard Hughes Medical Institute, NIH (grant NS040522), the Office of Naval Research (N00014-12-1-0299 and N00014-13-1-0205), and the Swartz Foundation. We also thank Manuel Hernandez, Jonathan Weyhenmeyer, and Howard Poizner for valuable discussion and Xin Wang for helping with the wavelet analysis.

References

Boashash, B. (1995). *Higher-order statistical signal processing*. Hoboken, NJ: Wiley.

Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. In *Proceedings 17 of the Institute of Phonetic Sciences, University of Amsterdam* (pp. 97–110).

Bracewell, R. (1965). *The Fourier transform and its applications*. New York: McGraw-Hill.

Brillinger, D., & Rosenblatt, M. (1967). Asymptotic theory of estimates of *k*th order spectra. In B. Harris (Ed.), *Spectral analysis of time series* (pp. 153–188). Hoboken, NJ: Wiley.

Chan, Y., & Langford, R. (1982). Spectral estimation via the high-order Yule-Walker equations. *IEEE Transactions on Acoustics, Speech and Signal Processing*, *30*(5), 689–698.

Davis, H., & Davis, P. (1936). Action potentials of the brain: In normal persons and in normal states of cerebral activity. *Archives of Neurology and Psychiatry*, *36*(6), 1214–1224.

Davis, P. A. (1939). Effects of acoustic stimuli on the waking human brain. *Journal of Neurophysiology*, *2*(6), 494–499.

Fackrell, J., & McLaughlin, S. (1995). Quadratic phase coupling detection using higher order statistics. In *Proceedings of the IEE Colloquium on Higher Order Statistics* (p. 9). N.p.

Falbo, C. E. (1995). Analytic and numerical solutions to the delay differential equation $y'(t) = \alpha y(t - \delta)$. In *Proceedings of the Joint Meeting of the Northern and Southern California Sections of the MAA*, San Luis Obispo, CA. N.p.

Fourier, J. (1822). *Théorie analytique de la chaleur*. Paris: Chez Firmin Didot.

Goertzel, G. (1958). An algorithm for the evaluation of finite trigonometric series. *American Mathematical Monthly*, *65*(1), 34–35.

Gorodnitsky, I., & Lainscsek, C. (2004). Machine emotional intelligence: A novel method for spoken affect analysis. In J. Triesch & T. Jebara (Eds.), *Proc. Intern. Conf. on Development and Learning ICDL 2004*. San Diego, CA: UCSD Institute for Neural Computation.

Hjorth, B. (1970). EEG analysis based on time domain properties. *Electroencephalography and Clinical Neurophysiology*, *29*(3), 306–310.

Jacobsen, E., & Lyons, R. (2003). The sliding DFT. *IEEE Signal Processing Magazine*, *20*(2), 74–80.

Kadtke, J., & Kremliovsky, M. (1999). Estimating dynamical models using generalized correlation functions. *Physics Letters A*, *260*(3–4), 203–208.

Kolmogorov, A., & Rozanov, Y. (1960). On strong mixing conditions for gaussian processes. *Theor. Prob. Appl.*, *5*, 204–208.

Lainscsek, C. (1999). *Identification and global modeling of nonlinear dynamical systems*. Doctoral dissertation, Technische Universität Graz.

Lainscsek, C. (2011). Nonuniqueness of global modeling and time scaling. *Phys. Rev. E*, *84*, 046205.

Lainscsek, C., & Gorodnitsky, I. (2003). Characterization of various fluids in cylinders from dolphin sonar data in the interval domain. In *Oceans 2003: Celebrating the past . . . Teaming toward the future* (Vol. 2, pp. 629–632). Piscataway, NJ: Marine Technology Society/IEEE.

Lainscsek, C., Hernandez, M. E., Poizner, H., & Sejnowski, T. J. (2015). Delay differential analysis of electroencephalographic data. *Neural Computation*, *27*(3), 615–627. doi: 10.1162/NECO_a_00656

Lainscsek, C., Hernandez, M. E., Weyhenmeyer, J., Sejnowski, T. J., & Poizner, H. (2013). Non-linear dynamical analysis of EEG time series distinguishes patients with Parkinson's disease from healthy individuals. *Frontiers in Neurology*, *4*(200).

Lainscsek, C., Letellier, C., & Gorodnitsky, I. (2003). Global modeling of the Rössler system from the z-variable. *Physics Letters A*, *314*, 409–427.

Lainscsek, C., Letellier, C., & Schürrer, F. (2001). Ansatz library for global modeling using a structure selection. *Physical Review E*, *64*, 016206.

Lainscsek, C., & Sejnowski, T. (2013). Electrocardiogram classification using delay differential equations. *Chaos*, *23*(2), 023132.

Lainscsek, C., Weyhenmeyer, J., Hernandez, M., Poizner, H., & Sejnowski, T. (2013). Non-linear dynamical classification of short time series of the Rössler system in high noise regimes. *Frontiers in Neurology*, *4*(182).

Leonov, V., & Shiryeav, A. (1959). On a method of calculation of semi-invariants. *Theor. Prob. Appl.*, *4*, 319–329.

Luck, S. (2014). *An introduction to the event-related potential technique* (2nd ed.). Cambridge, MA: MIT Press.

Mallat, S. (2008). *A wavelet tour of signal processing: The sparse way* (3rd ed.). Orlando, FL: Academic Press.

Mendel, J. (1991). Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications. In *Proc. IEEE*, *79*, 278–305.

Nikias, C., & Raghuver, M. (1987). Bispectrum estimation: A digital signal processing framework. *Proc. IEEE*, *75*, 869–891.

Nyquist, H. (1928). Certain topics in telegraph transmission theory. *Trans. AIEE*, *47*, 617–644.

Packard, N. H., Crutchfield, J. P., Farmer, J. D., & Shaw, R. S. (1980). Geometry from a time series. *Phys. Rev. Lett.*, *45*, 712.

Papoulis, A. (1962). *The Fourier integral and its applications*. New York: McGraw-Hill.

Raghuveer, M., & Nikias, C. (1985). Bispectrum estimation: A parametric approach. *IEEE Transactions on Acoustics, Speech and Signal Processing*, *33*(5), 1213–1230.

Raghuveer, M. R., & Nikias, C. L. (1986). Bispectrum estimation via AR modeling. *Signal Processing*, *10*(1), 35–48.

Roads, C. (1996). *Computer music tutorial*. Cambridge, MA: MIT Press.

Rosenblatt, M., & Van Ness, J. (1965). Estimation of the bispectrum. *Annals of Mathematical Statistics*, *36*, 1120–1136.

Rössler, O. E. (1976). An equation for continous chaos. *Physics Letters A*, *57*, 397.

Sauer, T., Yorke, J. A., & Casdagli, M. (1991). Embedology. *Journal of Statistical Physics*, *65*, 579.

Stankovic, L. (1994). A method for time-frequency analysis. *IEEE Transactions on Signal Processing*, *42*, 225–229.

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*(3700), 1187–1188.

Swami, A. (2003). *HOSA—higher order spectral analysis toolbox*. Natick, MA: Matlab.

Takens, F. (1981). Detecting strange attractors in turbulence. In D. A. Rand & L.-S. Young (Eds.), Dynamical systems and turbulence, Warwick 1980. In *Lecture Notes in Mathematics* 898 (pp. 366–381). Berlin: Springer.

Tukey, J. (1953). The spectral representation and transformation properties of the higher moments of stationary time series. In D. Brillinger (Ed.), *The collected works of John W. Tukey* (Vol. 1, pp. 165–184). Belmont, CA: Wadsworth.

Walker, G. (1931). On periodicity in series of related terms. *Proceedings of the Royal Society of London, Ser. A*, *131*, 518–532.

Whitney, (1936). Differentiable manifolds. *Ann. Math.*, *37*, 645–680.

Yule, G. (1927). On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London, Ser. A*, *226*, 267–298.

Received November 20, 2013; accepted September 29, 2014.