Optimized spectral estimation for nonlinear synchronizing systems

Linda Sommerlade,[1,2,3,4,*] Malenka Mader,[2,3,5] Wolfgang Mader,[2,3] Jens Timmer,[2,3,4] Marco Thiel,[1]
Celso Grebogi,[1,4] and Björn Schelter[1,3]

[1]*Institute for Complex Systems and Mathematical Biology, University of Aberdeen, Aberdeen AB24 3UE, United Kingdom*
[2]*Department of Physics, University of Freiburg, Hermann-Herder-Strasse 3, 79104 Freiburg, Germany*
[3]*Freiburg Center for Data Analysis and Modeling (FDM), University of Freiburg, Eckerstrasse 1, 79104 Freiburg, Germany*
[4]*Freiburg Institute for Advanced Studies (FRIAS), University of Freiburg, Albertstrasse 19, 79104 Freiburg, Germany*
[5]*Hospital for Neuropaediatrics and Muscular Diseases, University Medical Center of Freiburg, Germany*

In many fields of research nonlinear dynamical systems are investigated. When more than one process is measured, besides the distinct properties of the individual processes, their interactions are of interest. Often linear methods such as coherence are used for the analysis. The estimation of coherence can lead to false conclusions when applied without fulfilling several key assumptions. We introduce a data driven method to optimize the choice of the parameters for spectral estimation. Its applicability is demonstrated based on analytical calculations and exemplified in a simulation study. We complete our investigation with an application to nonlinear tremor signals in Parkinson's disease. In particular, we analyze electroencephalogram and electromyogram data.

I. INTRODUCTION

Investigating networks as an inverse problem, i.e., based on observations of the systems dynamics, is one of the key challenges faced in physics. Data-based modeling has opened new avenues to address this challenge. In particular, various data analysis techniques have been suggested to infer interactions between processes that build up the network. Among the most robust and therefore most frequently applied techniques are synchronization measures and cross-spectral analysis. This resulted from the fact that both approaches are robust with respect to noise, one of the key requirements in most applications. The cross-spectral analysis approach additionally provides information on the frequency at which the interaction is strongest, a feature that is strongly desired in applications to noisy nonlinear oscillatory systems; these are, for instance, often observed in the neurosciences [1,2]. Incidentally, in a recent publication, a striking similarity between one of the most frequently applied measures for phase synchronization, the mean phase coherence, and cross-spectral analysis was demonstrated [3]; the mean phase coherence is a special case of the wider class of cross-spectral analysis techniques.

Given the long history of applications of cross-spectral analysis to linear and nonlinear noisy systems [4–7], it seems surprising that one of the key problems in cross-spectral analysis has not been addressed before. Namely, cross-spectral analysis techniques rely on at least one parameter that has to be chosen beforehand, for the different proposed spectral estimators [5]. For instance, for two of the most commonly used methods, which are averaging and smoothing of the cross periodograms, the crucial parameters are the appropriate segment length for averaging and the window width for smoothing. In order to establish an asymptotically consistent estimator [8], the segment length or window width as well as the length of the time series have to be infinitely large

while their ratio tends to zero [9]. While these parameters are usually determined *ad hoc*, a poor choice will violate the implicit assumptions of the statistics.

The choice of the smoothing window determines whether or not a statistically significant result is obtained. Analytical significance levels can be derived under some assumptions [7,10,11]. Especially for narrowband signals these assumptions, which typically assume that an asymptote is reached, are difficult to achieve [12]. In case the asymptote is not reached, false positive conclusions may be drawn. As shown in this paper for two independent processes a suboptimal choice of the width of the smoothing window falsely indicates an interaction between the two processes. Thus, a suboptimal choice of the width of the smoothing window in this case would result in wrong conclusions about the interaction structure. As in applications the optimal width of the smoothing window is, of course, unknown, such erroneous conclusions are inevitable. The same problem occurs for other spectral estimators as well. For the approach based on averaging periodograms, often the choice of the segment length is based on the desired frequency resolution. When making this choice, the assumption that the segments are independent is often disregarded, resulting in exactly the same problem as discussed above.

In this paper, we overcome these limitations and derive an optimized width of the smoothing window, based on the optimized segment length, using the asymptotic equivalence of the distribution of the estimators. Thereby, we present an approach to optimize the choice of the window width in spectral estimation and thus make the application of coherence reliable. Thereby, we minimize false positive conclusions about the interactions between processes. We test our analytical results in simulation studies using linear and nonlinear synchronizing noisy processes. To demonstrate its applicability, the method is then applied to nonlinear tremor data affirming that the coherence of brain and muscle is significant for a trembling patient suffering from Parkinson's disease.

This paper is organized as follows. In Sec. II, univariate spectral estimation is reviewed for the averaging estimator

*l.sommerlade@abdn.ac.uk

and the smoothing estimator, respectively. In Sec. III bivariate coherence is reviewed. The parameters that need to be chosen for spectral estimation are the segment length and the smoothing window width. In Sec. IV the analytical derivation for the optimized choice of these parameters is summarized. This is followed by a simulation study based on a linear process for which the univariate spectrum is estimated (Sec. V). A bivariate simulation study (Sec. VI) of two synchronizing Rössler oscillators for which the coupling strength as well as the width of the smoothing window are systematically varied follows. Finally, we apply our method to nonlinear tremor data in Sec. VII.

II. UNIVARIATE SPECTRAL ESTIMATION

To optimize the choice of the key parameters in cross-spectral analysis when applied to nonlinear noisy systems, we first briefly introduce key concepts from the theory of spectral and cross-spectral analysis.

The power spectrum $S(\omega)$ of a process $x(t)$ is the Fourier transform of its autocovariance function $F_x(\tau) = \langle x(t)x(t-\tau) \rangle$, where $\langle \cdot \rangle$ denotes expected value. A naive estimator for the power spectrum is the so-called *periodogram*,

$$\hat{S}_{\text{naive}}(\omega) = P(\omega_k) = |f(\omega_k)|^2, \tag{1}$$

where $f(\omega_k)$ denotes the Fourier transform of the time series. Here $\omega_k = \frac{2\pi k}{N \Delta t}$ are the natural Fourier frequencies with $k = -\frac{N}{2}, \ldots, 0, \ldots, \frac{N}{2}$, where Δt is the sampling step and N refers to the number of data points.

In a well established procedure to deal with leakage, the time series is tapered [13] by multiplying it with an adequate window function before taking the Fourier transform. The variance of the tapered periodogram is then given by [5]

$$\text{var}\left[\frac{\hat{S}_{\text{naive}}^{\text{tap}}(\omega)}{S(\omega)} \right] = \frac{2}{\nu}, \tag{2}$$

where ν denotes the equivalent number of degrees of freedom. Here $\nu = 2\nu^{\text{tap}}$ depends on the taper window [13],

$$\nu^{\text{tap}} = \frac{q_2^2}{q_4}, \tag{3}$$

with

$$q_2 = \frac{1}{N} \sum_{i=1}^{N} W_{\text{tap}}^2(i) \quad \text{and} \quad q_4 = \frac{1}{N} \sum_{i=1}^{N} W_{\text{tap}}^4(i), \tag{4}$$

where $W_{\text{tap}}(i)$ is the window function. The variance is independent of the number of data points. Since tapering is used for bias reduction, the tapered periodogram is an asymptotically unbiased but inconsistent estimator. To overcome this, the periodogram can, for example, be either averaged over segments or smoothed by convolution with a smoothing kernel of appropriate width and shape. Both approaches are explained in the following sections.

A. Averaged periodograms

The first spectral estimator is based on the idea that for multiple measurements of the same stationary process, the spectrum can be estimated by averaging of the respective periodograms. The assumption of a decaying autocovariance function is implicitly made when averaging periodograms, since it is assumed that the segments are interpretable as independent realizations of the process and thus uncorrelated.

For a single measurement the given time series of length N is cut into M segments of length $L = \frac{N}{M}$. For practical purposes L is rounded down to an integer. The M segments are then assumed to form individual measurements of the process. Averaging over the M segments yields an asymptotically consistent estimator for the spectrum [13],

$$\hat{S}_{\text{ap}}(\omega_i) = \frac{1}{M} \sum_{l=1}^{M} P_l(\omega_k), \tag{5}$$

where the subscript "ap" stands for averaged periodograms. The variance of this spectral estimator is again given by Eq. (2). However, here ν depends on the number of segments

$$\nu_{\text{ap}} = 2M, \tag{6}$$

multiplied by the equivalent number of degrees of freedom of the taper window [Eq. (3)], i.e., $\nu = \nu_{\text{ap}} \nu_{\text{ap}}^{\text{tap}}$.

Assuming a constant length of segments, the variance decreases with increasing number of segments, i.e., increasing the total number of data points N; thus, this estimator is consistent. More precisely, it is asymptotically consistent for all frequencies if $N \to \infty$, $M \to \infty$, and $\frac{M}{N} \to 0$ holds [5]. For any fixed number N of measurements, the trade-off between a high frequency resolution and a small variance has to be dealt with.

B. Smoothed periodogram

The so-called smoothed periodogram uses the assumption of a smooth spectrum due to a decaying autocovariance function of the process directly. The periodogram of the whole time series is smoothed by convolution with a kernel K_j,

$$\hat{S}_{\text{sp}}(\omega_i) = \sum_{j=-h}^{h} K_j P(\omega_{i+j}), \tag{7}$$

where the subscript "sp" stands for smoothed periodogram. The kernel K_j needs to be normalized to one in order to preserve the variance of the process. Among several other possibilities, a triangular kernel,

$$K_j = \begin{cases} \frac{1}{h} - \frac{|j|}{h^2} & \text{for} |j| \leqslant h, \\ 0 & \text{else}, \end{cases} \tag{8}$$

can be used. Again, the variance of the spectral estimator is given by Eq. (2). In this case, ν depends on the smoothing kernel,

$$\nu_{\text{sp}} = \frac{2}{\sum_{j=-h}^{h} K_j^2}, \tag{9}$$

multiplied by the equivalent number of degrees of freedom of the taper window [Eq. (3)], i.e., $\nu = \nu_{\text{sp}} \nu_{\text{sp}}^{\text{tap}}$.

Using a triangular kernel [Eq. (8)] leads to

$$\nu_{\text{sp}} = \frac{6h^3}{(2h^2+1)} \tag{10}$$

[see Eqs. (A4) to (A7) in Appendix A]. Thus, increasing the width of the smoothing window h decreases the variance of the estimator. At the same time the number of data points has to increase in order to achieve a high frequency resolution. Therefore, smoothing the periodogram is an asymptotically consistent estimator for the spectrum if $N \to \infty$, $h \to \infty$, and $\frac{h}{N} \to 0$ holds [5].

To summarize, both spectral estimators based on the periodogram require the choice of a parameter such that the condition of consistency is fulfilled. For the averaged periodograms [Eq. (5)], the number of segments M or the segment length L, respectively, have to be chosen. For the smoothed periodogram [Eq. (7)], the width of the smoothing window h based on the shape of the smoothing kernel is mandatory. We optimize the width of the smoothing window based on the optimal segment length proposed by [9] in Sec. IV.

III. BIVARIATE SPECTRAL ESTIMATION: COHERENCE

For two stationary processes $x(t)$ and $y(t)$, the cross spectrum $S_{xy}(\omega)$ is defined as the Fourier transform of the cross-covariance function $F_{xy}(\tau) = \langle x(t)y(t-\tau)\rangle$ [14]. The cross spectrum is a complex valued quantity from which different real valued quantities can be derived. Here we investigate the normalized modulus of the cross spectrum $S_{xy}(\omega)$, called coherence [14],

$$C_{xy}(\omega) = \frac{|S_{xy}(\omega)|}{\sqrt{S_x(\omega)S_y(\omega)}} \quad \in [0,1]. \tag{11}$$

It describes the degree of linear predictability of $x(t)$ from $y(t)$ and vice versa in the frequency domain.

Estimation of the cross spectrum can be achieved analogously to the power spectrum. Consistent estimators can be obtained by averaging or smoothing cross periodograms. The variance of this estimator is given by [13]

$$\mathrm{var}[\hat{C}_{xy}(\omega)] = \frac{1}{\nu}[1 - |C_{xy}(\omega)|^2], \tag{12}$$

where, as above, ν refers to the equivalent number of degrees of freedom, which, in turn, depends on the estimation procedure [Eqs. (3), (6), and (9)], as before.

In order to estimate the variance of the coherence, the true coherence $C_{xy}(\omega)$ in Eq. (12) has to be substituted by its estimate. Under the null hypothesis of zero coherence, a critical value,

$$s = \sqrt{1 - \alpha^{\frac{1}{\frac{\nu}{2}-1}}}, \tag{13}$$

for a given α-significance level is obtained [13].

As both spectral as well as cross-spectral analysis rely on periodograms, optimization in terms of the choice of the parameters applied to univariate as well as to bivariate spectral analysis is crucial.

IV. OPTIMIZING SPECTRAL ESTIMATION

The segment length for averaging can be estimated from the measured data if the process is exponentially mixing [9,15].

To achieve this, an exponentially decaying function,

$$f(k) = \varphi^k, \varphi < 1, \tag{14}$$

is fitted to the envelope of the autocorrelation function $F_x^{\mathrm{norm}}(\tau) = \frac{1}{\sigma^2}\langle x(t)x(t-\tau)\rangle$, which is the autocovariance function normalized by the variance σ^2 of the process. The optimal segment length is [9]

$$L_{\mathrm{opt}} = \sqrt[3]{\frac{4\left[\frac{\varphi}{1-\varphi} + \frac{\varphi^2}{(1-\varphi)^2}\right]^2}{\left(1 + 2\frac{\varphi}{1-\varphi}\right)^2} N}. \tag{15}$$

This optimal segment length holds for the estimator of the variance, as the spectrum is the variance per frequency; it is a suitable choice for (cross-) spectral analysis as well. Typically in nonlinear dynamics the spectrum is dominated by the oscillation frequency of the process. The segment length is then determined by this dominant frequency. As this frequency contains most of the variance, the segment length can indeed be considered as "optimal." If the spectrum contains several peaks, the peak that dominates the decay time of the autocovariance function determines the optimal segment length. In some applications this dominant peak may differ from the frequency of interest. If the data are filtered according to the frequencies of interest, optimal segment lengths can be determined for the filtered signal in order to determine segment lengths for each peak. This in turn results in optimal segment lengths for spectral estimation in the presence of several spectral peaks. In the following we use the term optimal in the sense discussed above.

Since both the averaged and the smoothed periodogram are χ_ν^2-distributed random variables that are asymptotically unbiased estimators of the spectrum, they are equivalent. Using this equivalence, their degrees of freedom are set equal,

$$\nu = \nu_{\mathrm{ap}}\nu_{\mathrm{ap}}^{\mathrm{tap}} = \nu_{\mathrm{sp}}\nu_{\mathrm{sp}}^{\mathrm{tap}}, \tag{16}$$

where $\nu_{\mathrm{ap}}^{\mathrm{tap}}$ and $\nu_{\mathrm{sp}}^{\mathrm{tap}}$ differ since the length of the tapered segments differ. This yields the relation

$$\frac{\left[\frac{1}{L}\sum_{i=1}^{L}\tilde{W}_{\mathrm{tap}}^2(i)\right]^2}{\frac{1}{L}\sum_{i=1}^{L}\tilde{W}_{\mathrm{tap}}^4(i)} 2M = \frac{\left[\frac{1}{N}\sum_{i=1}^{N}W_{\mathrm{tap}}^2(i)\right]^2}{\frac{1}{N}\sum_{i=1}^{N}W_{\mathrm{tap}}^4(i)} \frac{2}{\sum_{j=-h}^{h}K_j^2}, \tag{17}$$

and, thus,

$$\underbrace{M\frac{\left[\frac{1}{L}\sum_{i=1}^{L}\tilde{W}_{\mathrm{tap}}^2(i)\right]^2}{\frac{1}{L}\sum_{i=1}^{L}\tilde{W}_{\mathrm{tap}}^4(i)}\frac{\frac{1}{N}\sum_{i=1}^{N}W_{\mathrm{tap}}^4(i)}{\left[\frac{1}{N}\sum_{i=1}^{N}W_{\mathrm{tap}}^2(i)\right]^2}}_{M'=M'(M,L)} = \frac{1}{\sum_{j=-h}^{h}K_j^2}. \tag{18}$$

Inserting the optimal segment length and the respective optimal number of segments such that $M'_{\mathrm{opt}} = M'_{\mathrm{opt}}(M_{\mathrm{opt}},L_{\mathrm{opt}})$, and solving for the smoothing window width h yields the optimal window width h_{opt}. For a triangular kernel we obtain

$$h_{\mathrm{opt}} = \sqrt[3]{\left(\frac{8}{729}M'^2 + \frac{1}{6} + \sqrt{\frac{8}{2187}M'^2 + \frac{1}{36}}\right)M'}$$

$$+ \sqrt[3]{\left(\frac{8}{729}M'^2 + \frac{1}{6} - \sqrt{\frac{8}{2187}M'^2 + \frac{1}{36}}\right)M'}$$

$$+ \frac{2}{9}M'. \tag{19}$$

The derivation is shown in Appendix A for $M' = M'_{\text{opt}}$, depending on the optimal number of segments, i.e., the optimal segment length,

$$M_{\text{opt}} = \frac{N}{L_{\text{opt}}}, \tag{20}$$

in the averaging approach.

If more than eight independent segments of optimal length, i.e., $M_{\text{opt}} > 8$ could be obtained from the data, the approximation

$$h_{\text{opt}} \approx \frac{2}{3}M_{\text{opt}} \tag{21}$$

for the optimal width of the smoothing window holds, as shown in Appendix B.

Interestingly, if the estimated optimal segment length is longer than the duration of the sampled time series, this is a strong indication that no meaningful statistics can be found. This is a major and almost unique advantage of our approach: If the statistics is not applicable, our approach suggests a smoothing window of zero width. Almost all other approaches will provide a result irrespective of the reliability of the results.

V. UNIVARIATE SIMULATION STUDY

To exemplify the above results in a simulation study, we consider an autoregressive process of order two (AR[2]),

$$x(t) = a_1 x(t-1) + a_2 x(t-2) + \varepsilon(t), \tag{22}$$

with $a_1 = 1.6$, $a_2 = -0.98$, and where $\varepsilon(t)$ is Gaussian white noise. We vary the number of data points N after removal of transients between 5000 and 50 000 in steps of 5000; additionally we used $N = 1000$. For 100 realizations of the AR[2] process [Eq. (22)], with different N the optimal width of the smoothing window is computed according to Eq. (19). The optimal segment length L_{opt} according to Eq. (20) is estimated by fitting an exponentially decaying function [Eq. (14)] to the envelope of the estimated autocorrelation function of the time series.

For the AR[2] given by Eq. (22), the decay constant,

$$\varphi = \sqrt{-a_2}, \tag{23}$$

of the envelope of the autocorrelation function can be derived analytically [16]. Therefore, the optimal segment length L_{opt} given in Eq. (15) and, from this, the optimal width of the smoothing window h_{opt} [Eq. (19)] can be obtained analytically.

In Figs. 1 and 2, the mean estimated optimal width of the smoothing window (black crosses) of the simulation study with varying data length N is compared to the analytical one [red (gray) line]. The standard deviations of the mean optimal width of the smoothing windows for the 100 realizations of length N are indicated by the black error bars. The results for the width of the smoothing window in frequency bins are shown in Fig. 1. From these bins, the width of the smoothing window
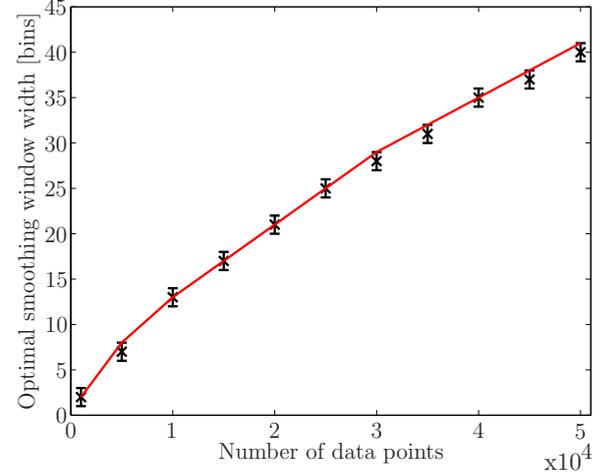

FIG. 1. (Color online) Optimal width of the smoothing window h_{opt} for AR[2] realizations given in Eq. (22) dependent on the length of the realization. The black crosses and error bars denote the mean and its standard deviation of h_{opt} for 100 realizations computed from the estimated segment lengths. The optimal width of the smoothing window derived from the analytic optimal segment length is shown in red (gray).

in Hz is calculated and depicted in Fig. 2. The simulated mean values agree with the analytical values.

VI. BIVARIATE SIMULATION STUDY

In order to investigate the impact of the choice of the width of the smoothing window on cross-spectral estimation using the smoothing approach, we simulate a system of two identical
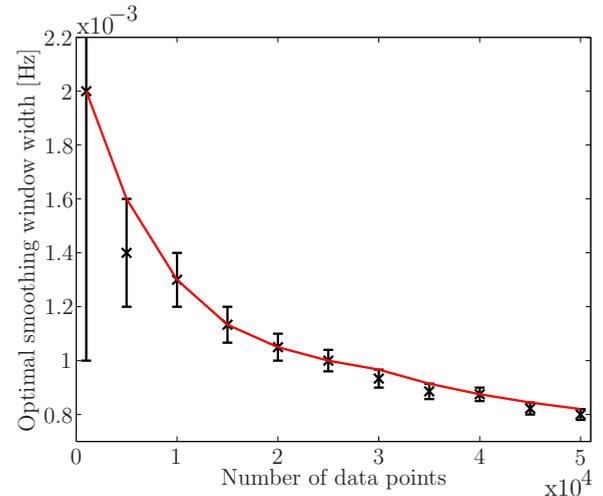

FIG. 2. (Color online) Optimal width of the smoothing window h_{opt} in Hz for AR[2] realizations given in Eq. (22) dependent on the length of the realization. The black crosses and error bars denote the mean and its standard deviation of h_{opt} for 100 realizations computed from the estimated segment lengths. The optimal width of the smoothing window derived from the analytic optimal segment length are shown in red (gray). For $N = 1000$ the standard deviation corresponds to one frequency bin or 10^{-3} Hz.
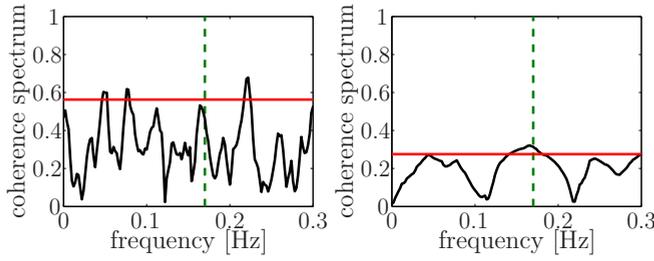
FIG. 3. (Color online) Coherence spectra of two independent Rössler oscillators [Eq. (24)]. (Left) Width of the smoothing window chosen optimally (here 5 frequency bins). (Right) Smoothing window five times as wide as optimal (here 25 frequency bins). Critical values for a 5% significance level are depicted in red (gray). The oscillation frequency of 0.17 Hz is highlighted by the horizontal green dashed lines.

stochastic Rössler oscillators,

$$\dot{x}_i(t) = -\omega y_i(t) - x_i(t) + \gamma[x_j(t) - x_i(t)] + \varepsilon_i(t),$$
$$\dot{y}_i(t) = \omega x_i(t) + ay_i(t), \tag{24}$$
$$\dot{z}_i(t) = b + [x_i(t) - c]z_i(t),$$

with $i, j = 1, 2$, $i \neq j$, $\omega = 1$, $a = 0.2$, $b = 0.2$, $c = 6.3$, and Gaussian distributed white noise ε_i of unit variance. We use $N = 50\,000$ data points after removal of transients. The rate of mixing of this system is rather low [12] and thus the decay of the autocovariance function is slow.

Figure 3 shows the coherence of the uncoupled ($\gamma = 0$) Rössler system [Eq. (24)], as derived from optimal smoothing (left) and five times too wide smoothing (right). The oscillator's main frequency is highlighted (green dashed vertical line). In this example, a spurious interaction is detected if the smoothing window is chosen too wide.

In a simulation study we vary the bidirectional coupling strength γ between 0 and 0.5 in steps of 0.025 for 100 realizations of Eq. (24) each. For each realization we estimate the coherence [Eq. (11)] using different windows. Significance is tested at the main frequency of the oscillators according to Eq. (13) for optimal smoothing as well as for five times too wide smoothing and for half of the optimal width of the smoothing window. The performance of the three smoothing scenarios for varying coupling strength γ is shown in Fig. 4 (top). For each coupling strength the rate of significant coherence is shown.

Additionally to the coherence we estimated mean phase coherence (MPC) [3,17] for each realization. The mean MPC for 100 realizations, depending on the coupling strength, is shown in Fig. 4 (bottom). The increase of MPC shows that for increased coupling the two Rössler oscillators are synchronized.

A detailed investigation of the coherence estimates for the synchronizing oscillators shows that if the width of the smoothing window is chosen too small (green dashed line in Fig. 4), the estimation procedure loses power. An interaction is, thus, detected less reliably than for optimal smoothing. In contrast, if the width of the smoothing window is chosen too wide (blue dashed dotted line in Fig. 4), the size of the significance test exceeds the significance level, i.e., for a given

5% significance level; more than 5% false positive conclusions are drawn if there is no interaction, i.e., $\gamma = 0$. Choosing the optimal width of the smoothing window results in the best power while preserving the size of the test.

The same analysis was carried out for the Rössler system in the funnel regime. We used Eq. (24) with $\omega = 1$, $a = 0.3$, $b = 0.4$, and $c = 7.5$. We again varied the bidirectional coupling strength γ between 0 and 0.5 in steps of 0.025. Results are shown in Fig. 5. The Rössler system in the funnel regime is a broadband chaotic system. For broadband signals the
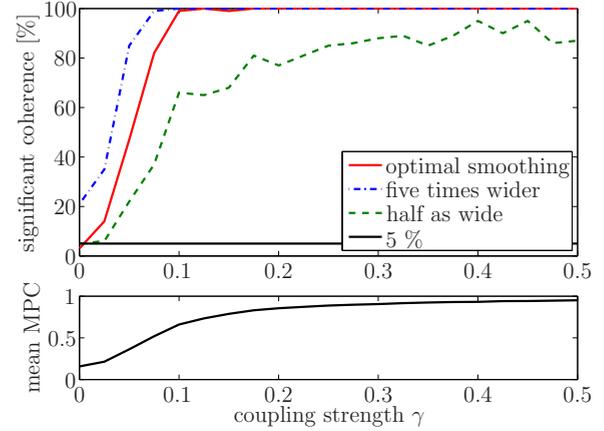

FIG. 4. (Color online) (Top) Rate of significant coherence between two Rössler oscillators with increasing bidirectional coupling. We estimated coherence with optimal width of the smoothing window [red (gray) line], five times wider window width (blue dash-dotted line), and a window width half as wide as optimally (green dashed line), respectively. Significance was evaluated at the oscillation frequency of 0.17 Hz with a 5% significance level. (Bottom) Averaged MPC for the respective 100 realizations for each coupling strength.
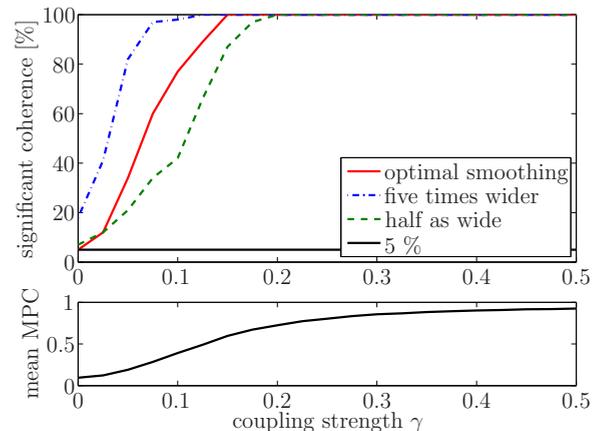

FIG. 5. (Color online) (Top) Rate of significant coherence between two Rössler oscillators in the funnel regime with increasing bidirectional coupling. We estimated coherence with optimal width of the smoothing window [red (gray) line], five times wider window width (blue dash-dotted line), and a window width half as wide as optimally (green dashed line), respectively. Significance was evaluated at the oscillation frequency of 0.14 Hz with a 5% significance level. (Bottom) Averaged MPC for the respective 100 realizations for each coupling strength.

autocovariance function is decaying faster than for narrowband signals. This is due to their respective mixing properties. It is thus expected that for broadband signals the periodogram can be smoothed with a wider smoothing window width. Therefore, it is less likely to accidentally choose the smoothing window width too wide. Nevertheless, the choice can be optimized with respect to the power and size of significance testing using our approach.

VII. APPLICATION TO TREMOR TIME SERIES

One of the core symptoms of Parkinson's disease is tremor, which is an involuntary rhythmical movement predominantly of the upper limbs. In Parkinson's disease the tremor frequency is mainly between 4 and 6 Hz [18]. The pathophysiology of the human tremor is still under debate. One hypothesis that is widely accepted [19–21] states that an abnormal oscillatory activation in the brain may be generating the tremor. Hypotheses on the location of the abnormal oscillatory activation differ [18,22–26].

We used electroencephalogram (EEG) data from one electrode located over the hand area of the left sensorimotor cortex [20,27]. Using electromyography (EMG), the muscle activity from the flexor muscle of the right wrist was simultaneously recorded to obtain a readout of the tremor activity. Data were recorded at a sampling rate of 1 kHz and EMG recordings were rectified. Data were downsampled to 100 Hz prior to analysis.

Parkinsonian tremor signals were shown to be nonlinear [28]. We applied the proposed method of optimally choosing the width of the smoothing window for spectral estimation to these signals. The results of coherence estimation of EEG and EMG, with the optimal width of the smoothing window, is shown in Fig. 6. Choosing the width of the smoothing window optimally, i.e., 22 frequency bins and therefore 0.072 Hz, a significant coherence at the tremor frequency of 4.4 Hz is identified. Thus, we conclude that, in fact, there is coherence between muscle and brain at the tremor frequency. This finding supports the hypothesis that an abnormal oscillatory activation in the brain is generating the tremor.
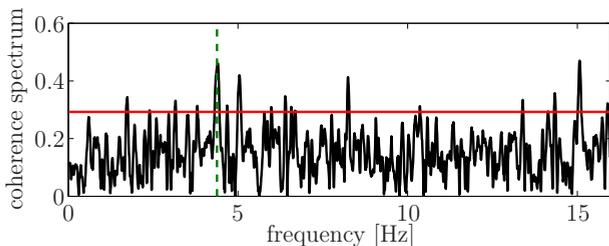

FIG. 6. (Color online) Coherence spectrum of right EMG and contralateral, i.e., left EEG time series of one patient suffering from Parkinsonian tremor estimated using the optimal width of the smoothing window of 22 frequency bins and 0.072 Hz, accordingly. The critical value for a 5% significance level is depicted in red (gray). The tremor frequency of 4.4 Hz is highlighted by the horizontal green dashed line.

VIII. DISCUSSION AND CONCLUSION

The method presented here offers a broad opportunity to employ unbiased and consistent spectral estimation with a data driven choice of the width of the spectral smoothing window. The optimal width of the smoothing window is chosen based on the mixing properties of the process. Exemplarily, we used a triangular kernel for smoothing of the periodogram. For other smoothing kernels the optimal width of the smoothing window can be derived analogously. Based on the approach presented here, spectral estimation using an adaptive width of the smoothing window [29] can be improved. This can be achieved by replacing the heuristic choice of the width of the smoothing window at the peak frequency with the width of the smoothing window presented here.

We investigated the usefulness for univariate and bivariate spectral estimation based on simulations both for linear and nonlinear synchronizing systems. In an application to EMG and EEG data of a Parkinsonian tremor patient we detected a significant interaction of the muscle and brain activity at the tremor frequency. We demonstrated that the choice of the width of the smoothing window is crucial in cross-spectral analysis. On the one hand, if the smoothing window is chosen too wide, interactions may be detected spuriously. On the other hand, if the smoothing window is chosen too narrow, the sensitivity to detect actual interactions is hampered. Choosing the width of the smoothing window optimally prevents false positive conclusions while preserving a high sensitivity.

We emphasize that the method presented here is data driven. This means that the optimal width of the smoothing window is estimated based on a given measurement. The estimated width of the smoothing window may be zero, if the measurement is not long enough to treat segments of it as independent realizations. In this case, the asymptotic is not reached and smoothing is not advisable. In case the estimated width of the smoothing window is larger than zero, the statistics can be applied, leading to a reliable (cross-) spectral estimation. Thus, we presented an approach to choose the width of the smoothing window optimally with respect to the measured data at hand.

APPENDIX A: CALCULATION OF h_{opt}

Here the optimal width of the smoothing window h_{opt} is derived for a triangular window [Eq. (8)],

$$K_j = \begin{cases} \frac{1}{h} - \frac{|j|}{h^2} & \text{for} |j| \leqslant h, \\ 0 & \text{else}. \end{cases} \quad (A1)$$

For this, the degrees of freedom of the spectral estimators via averaging and smoothing are set equal [Eq. (18)], leading to

$$M' = \frac{1}{\sum_{j=-h}^{h} \left(\frac{1}{h} - \frac{|j|}{h^2} \right)^2}, \qquad \text{(A2)}$$

with

$$M' = M \underbrace{\frac{\left[\frac{1}{L} \sum_{i=1}^{L} \tilde{W}_{\text{tap}}^2(i) \right]^2}{\frac{1}{L} \sum_{i=1}^{L} \tilde{W}_{\text{tap}}^4(i)}}_{\nu_{\text{ap}}^{\text{tap}}} \underbrace{\frac{\frac{1}{N} \sum_{i=1}^{N} W_{\text{tap}}^4(i)}{\left[\frac{1}{N} \sum_{i=1}^{N} W_{\text{tap}}^2(i) \right]^2}}_{(\nu_{\text{sp}}^{\text{tap}})^{-1}}, \qquad \text{(A3)}$$

the number of segments M scaled with the fraction of the degrees of freedom for the taper kernels. Multiplying the inverse of Eq. (A2) with h^2 after factoring out $\frac{1}{h^2}$ on the right yields

$$\frac{h^2}{M'} = \sum_{j=-h}^{h} \left(1 - \frac{2|j|}{h} + \frac{j^2}{h^2} \right) \qquad \text{(A4)}$$

$$= \sum_{j=-h}^{h} 1 - \frac{4}{h} \sum_{j=0}^{h} j + \frac{2}{h^2} \sum_{j=0}^{h} j^2 \qquad \text{(A5)}$$

$$= (2h + 1) - 2(h + 1) + \frac{(h+1)(2h+1)}{3h} \qquad \text{(A6)}$$

$$= \frac{2h^2 + 1}{3h}. \qquad \text{(A7)}$$

Solving the resulting equation,

$$3h^3 - 2M'h^2 - M' = 0, \qquad \text{(A8)}$$

for h yields the optimal width of the smoothing window since M' can be chosen optimally when the number of segments M used for averaging the periodograms is chosen optimally, such that $L_{\text{opt}} = \frac{N}{M_{\text{opt}}}$ is the optimal segment length as derived in Sec. IV.

The cubic equation is solved by a substitution step, which transforms Eq. (A8) to $x^3 + px + q = 0$ with $x = h + \frac{2}{3}M'$, $p = -\frac{2^2}{3^3}M'^2$, $q = -\frac{2^4}{3^6}M'^3 - \frac{1}{3}M'$, and Cardano's method as

a second step [30]. The solution is the optimal width of the smoothing window,

$$h_{\text{opt}} = \sqrt[3]{\left(\frac{8}{729} M_{\text{opt}}'^2 + \frac{1}{6} + \sqrt{\frac{8}{2187} M_{\text{opt}}'^2 + \frac{1}{36}} \right) M_{\text{opt}}'}$$

$$+ \sqrt[3]{\left(\frac{8}{729} M_{\text{opt}}'^2 + \frac{1}{6} - \sqrt{\frac{8}{2187} M_{\text{opt}}'^2 + \frac{1}{36}} \right) M_{\text{opt}}'}$$

$$+ \frac{2}{9} M_{\text{opt}}'. \qquad \text{(A9)}$$

APPENDIX B: APPROXIMATION OF h_{opt}

The exact solution Eq. (A9) can be approximated if

$$\frac{8}{729} M_{\text{opt}}'^2 \gg \frac{1}{6} \pm \sqrt{\frac{8}{2187} M_{\text{opt}}'^2 + \frac{1}{36}}. \qquad \text{(B1)}$$

Subtracting $\frac{1}{6}$ on both sides and squaring the equation yields

$$\frac{64}{729^2} M_{\text{opt}}'^4 - \frac{8}{2187} M_{\text{opt}}'^2 + \frac{1}{36} \gg \frac{8}{2187} M_{\text{opt}}'^2 + \frac{1}{36}. \quad \text{(B2)}$$

This leads to

$$M_{\text{opt}}'^2 \gg 2 \frac{729^2}{64} \frac{8}{729 \times 3} = 2 \frac{729}{8} \frac{1}{3} = \frac{243}{4}, \qquad \text{(B3)}$$

$$M_{\text{opt}}' \gg \frac{9}{2} \sqrt{3} \approx 7.8. \qquad \text{(B4)}$$

This means that at least $M_{\text{opt}}' = 8$ segments are necessary for the first term in the cubic roots larger than the remaining terms.

Therefore,

$$h \approx \sqrt[3]{\frac{8}{729} M_{\text{opt}}'^2 M_{\text{opt}}'} + \sqrt[3]{\frac{8}{729} M_{\text{opt}}'^2 M_{\text{opt}}'} + \frac{2}{9} M_{\text{opt}}' \quad \text{(B5)}$$

$$= \frac{2}{9} M_{\text{opt}}' + \frac{2}{9} M_{\text{opt}}' + \frac{2}{9} M_{\text{opt}}' = \frac{2}{3} M_{\text{opt}}' \qquad \text{(B6)}$$

holds if $M_{\text{opt}}' \gg 8$.

[1] F. Mormann, R. G. Andrzejak, C. E. Elger, and K. Lehnertz, Brain **130**, 314 (2007).

[2] C. A. Teixeira, B. Direito, H. Feldwisch-Drentrup, M. Valderrama, R. P. Costa, C. Alvarado-Rojas, S. Nikolopoulos, M. L. V. Quyen, J. Timmer, B. Schelter *et al.*, J. Neurosci. Methods **200**, 257 (2011).

[3] B. Schelter, M. Winterhalder, R. Dahlhaus, J. Kurths, and J. Timmer, Phys. Rev. Lett. **96**, 208103 (2006).

[4] E. J. Hannan, *Time Series Analyis*, Methuen's Monographs on Applied Probability and Statistics (Methuen, London, 1960).

[5] D. R. Brillinger, *Time Series: Data Analysis and Theory* (Holden-Day, San Francisco, 1981).

[6] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications—Multitaper and Conventional Univariate Techniques* (Cambridge University Press, Cambridge, U.K., 1993).

[7] J. Timmer, M. Lauk, S. Häußler, V. Radt, B. Köster, B. Hellwig, B. Guschlbauer, C. H. Lücking, M. Eichler, and G. Deuschl, Int. J. Bifurcation Chaos Appl. Sci. Eng. **10**, 2595 (2000).

[8] D. Thomson, Proc. IEEE **70**, 1055 (1982).

[9] M. Peifer, B. Schelter, B. Guschlbauer, B. Hellwig, C. H. Lücking, and J. Timmer, Biometrical J. **47**, 346 (2005).

[10] M. Winterhalder, B. Schelter, J. Kurths, A. Schulze-Bonhage, and J. Timmer, Phys. Lett. A **356**, 26 (2006).

[11] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods* (Springer, New York, 1998).

[12] M. Peifer, B. Schelter, M. Winterhalder, and J. Timmer, Phys. Rev. E **72**, 026213 (2005).

[13] P. Bloomfield, *Fourier Analysis of Time Series: An Introduction* (Wiley & Sons, New York, 1976).

[14] M. B. Priestley, *Spectral Analysis and Time Series* (Academic Press, London, 1989).

[15] P. Hall, J. L. Horowitz, and B.-Y. Jing, Biometrika **82**, 561 (1995).

[16] M. Mader, W. Mader, L. Sommerlade, J. Timmer, and B. Schelter, J. Neurosci. Methods **219**, 285 (2013).

[17] M. G. Rosenblum, A. S. Pikovsky, and J. Kurths, Phys. Rev. Lett. **76**, 1804 (1996).

[18] L. Timmermann, J. Gross, M. Dirks, J. Volkmann, H.-J. Freund, and A. Schnitzler, Brain **126**, 199 (2002).

[19] C. Reck, E. Florin, L. Wojtecki, H. Krause, S. Groiss, J. Voges, M. Maarouf, V. Sturm, A. Schnitzler, and L. Timmermann, Eur. J. Neurosci. **29**, 599 (2009).

[20] B. Hellwig, S. Häußler, M. Lauk, B. Köster, B. Guschlbauer, R. Kristeva-Feige, J. Timmer, and C. Lücking, Electroencephalogr. Clin. Neurophysiol. **111**, 806 (2000).

[21] D. M. Halliday, B. A. Conway, S. F. Farmer, U. Shahani, A. J. Russell, and J. R. Rosenberg, Lancet **355**, 1149 (2000).

[22] X. Liu, H. L. Ford-Dunn, G. N. Hayward, D. Nandi, R. C. Miall, T. Z. Aziz, and J. F. Stein, Clin. Neurophysiol. **113**, 1667 (2002).

[23] C. Hammond, H. Bergman, and P. Brown, Trends Neurosci. **30**, 357 (2007).

[24] J. Raethjen, M. Lindemann, H. Schmaljohann, R. Wenzelburger, G. Pfister, and G. Deuschl, Mov. Disord. **15**, 84 (2000).

[25] M. Weinberger, W. D. Hutchison, M. Alavi, M. Hodaie, A. M. Lozano, E. Moro, and J. O. Dostrovsky, Clin. Neurophysiol. **123**, 358 (2012).

[26] R. C. Helmich, M. J. R. Janssen, W. J. G. Oyen, B. R. Bloem, and I. Toni, Ann. Neurol. **69**, 269 (2011).

[27] B. Hellwig, B. Schelter, B. Guschlbauer, J. Timmer, and C. H. Lücking, Clin. Neurophysiol. **114**, 1462 (2003).

[28] J. Timmer, S. Häußler, M. Lauk, and C. H. Lücking, Chaos **10**, 278 (2000).

[29] J. Timmer, M. Lauk, and G. Deuschl, Electroencephalogr. Clin. Neurophysiol. **101**, 461 (1996).

[30] M. Artin, *Algebra* (Prentice-Hall, Upper Saddle River, NJ, 1991).