

**Harmonic cross-correlation decomposition for multivariate time series**Tanja Zerenner<sup>1</sup>,\* Marc Goodfellow<sup>2</sup>, and Peter Ashwin<sup>2</sup>*EPSRC Centre for Predictive Modeling in Healthcare, University of Exeter, Exeter EX4 4PY, United Kingdom and College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4PY, United Kingdom*

(Received 18 December 2020; accepted 25 May 2021; published 21 June 2021)

We introduce harmonic cross-correlation decomposition (HCD) as a tool to detect and visualize features in the frequency structure of multivariate time series. HCD decomposes multivariate time series into spatiotemporal harmonic modes with the leading modes representing dominant oscillatory patterns in the data. HCD is closely related to data-adaptive harmonic decomposition (DAHD) [Chekroun and Kondrashov, *Chaos* **27**, 093110 (2017)] in that it performs an eigendecomposition of a grand matrix containing lagged cross-correlations. As for DAHD, each HCD mode is uniquely associated with a Fourier frequency, which allows for the definition of multidimensional power and phase spectra. Unlike in DAHD, however, HCD does not exhibit a systematic dependency on the ordering of the channels within the grand matrix. Further, HCD phase spectra can be related to the phase relations in the data in an intuitive way. We compare HCD with DAHD and multivariate singular spectrum analysis, a third related correlation-based decomposition, and we give illustrative applications to a simple traveling wave, as well as to simulations of three coupled Stuart-Landau oscillators and to human EEG recordings.

DOI: [10.1103/PhysRevE.103.062213](https://doi.org/10.1103/PhysRevE.103.062213)**I. INTRODUCTION**

The identification of periodic components in observations of complex systems can help to improve our understanding of system dynamics and aid in predictive modeling. Alterations of the oscillatory electrical activity of the human brain, for instance, are linked to a diverse range of neurological conditions [1,2]. The identification and quantification of meaningful features in oscillatory brain activity has been shown to aid EEG-based diagnostics [3–6]. Another example are periodic variations in atmospheric and oceanic circulations, which affect weather on both regional and global scales [7–9]. The identification and quantification of these periodicities can aid in empirical-statistical forecasting [10,11]. Data from high-dimensional complex systems such as the climate system or the human brain, however, are typically noisy and contain periodicities over a wide range of frequencies, which can make it challenging to extract meaningful features.

One approach to tackle this challenge is singular spectrum analysis (SSA). SSA is a nonparametric spectral estimation method utilizing a delay embedding of a time series sampled from some dynamical system. SSA has its roots in dynamical systems theory. Since the embedding theorem of Takens [12], it has been known that it is possible to make faithful reconstructions of attractors for finite-dimensional dynamical systems using a delay embedding of a generic observable from the system. The use of SSA for understanding aspects of the qualitative dynamics, such as attractor structure and dimension, is discussed, for instance, in [13,14]. SSA has further been utilized in signal processing and time-series analysis

[15,16]. The ability to extract predictable trends and periodic components from noisy series make SSA and its multivariate extension M-SSA useful tools in empirical-statistical forecasting [10,17–19], noise filtering [20,21], and gap filling of time-series data with missing values, especially within the geosciences [22–25]. Also in the biomedical sciences, SSA and M-SSA have found diverse applications, for instance in the classification of sleep EEG [26–29] or to separate different types of murmur from normal heart sounds [30]. Sliding window versions of SSA or M-SSA provide tools for detecting nonstationarities in time-series data [31], which have found applications in the detection of epileptic seizures and seizure-type identification [32,33].

SSA is closely linked to principal component analysis (PCA) of multivariate time series. PCA is also known as Karhunen-Loeve transform, Hotelling transform, or, especially in the atmospheric sciences literature, empirical orthogonal function (EOF) analysis. PCA of multivariate time series performs an eigendecomposition of the zero-lag correlation matrix. The dimension of the PCA eigenvectors corresponds to the number of series, in the following also referred to as channels. Each entry of an eigenvector is associated with a particular channel. For time-series data for which the channels are associated with points in (Euclidian) space, EEG recordings of brain electrical activity measured at electrodes placed on the scalp, or temperature recordings at stations all over the globe, for instance, the EOFs are thus spatial. By projecting the original data onto the EOF basis, one obtains the principal components (PCs), an alternate representation of the given data set in terms of statistically independent (uncorrelated at zero lag) series. The associated eigenvalue provides the variance explained by the respective PC. The EOF basis is optimal in terms of explained variance, which makes PCA a popular tool both for data analysis and

\*t.zerenner@exeter.ac.uk

compression in various disciplines including biomedical sciences [34–36] and geosciences [37–39].

SSA may be viewed as the temporal equivalent of PCA. SSA performs an eigendecomposition of the correlation matrix of a lag-embedded time series. The eigenvectors are thus temporal and represent statistically persistent temporal patterns in the data, typically trends, oscillatory components, and noise [40,41]. Multivariate (or multichannel) SSA is thus a combination of SSA and PCA. Accordingly, M-SSA is also known as extended EOF (EEOF) analysis in the atmospheric sciences literature [42]. In M-SSA and EEOF analysis, the eigenvectors, also called modes, have both a spatial and a temporal dimension. For the simple traveling wave shown in Fig. 1, the relation between time-series data [Fig. 1(a)] and M-SSA modes [Fig. 1(e), left] is apparent. For a simple traveling wave in which each channel time series is given by a pure sine, the M-SSA modes take the shape of pure sines along their temporal dimension. For more complex real-world data, however, this is typically not the case in M-SSA. Thus one can usually not uniquely assign an M-SSA mode to a single frequency.

Aiming at a particular data-driven approach to modeling observables of real-world complex systems, Ref. [43] recently introduced data-adaptive harmonic decomposition (DAHD). Like M-SSA, DAHD is based on an eigendecomposition of a grand matrix of lagged cross-correlations. In the DAHD method, this grand matrix is built from (circulant) Hankel matrices. Unlike M-SSA modes, DAHD modes are pure sines along their temporal dimension. Each mode can thus be uniquely assigned to a particular frequency, and by projecting the data onto the DAH modes one obtains an alternative representation of the data in terms of oscillatory pairs that are in a narrow band around the frequency of the corresponding mode and approximately in phase quadrature. In [43–46], the transformed series are then modeled in terms of linearly coupled Stuart-Landau oscillators, where each pair of oscillatory series is identified with real and imaginary parts of an oscillator. The harmonic nature of the DAH modes further allows the definition of high-dimensional DAH power and phase spectra, which can provide frequency-specific insight into phase relations and correlation structure in multivariate time series [43–48]. DAHD is, however, not invariant under the permutation of the ordering of the individual time series. In other words, DAH modes and spectra are systematically affected by the ordering of the channels within the DAHD grand matrix. For data not embedded in one dimension in Euclidean space, the decomposition is typically not unique. The described dependency on the channel ordering is caused by a particular step in the construction of the grand matrix, which is required to ensure symmetry and thereby ensure that the eigenvalues and the eigenvectors are real. Figure 1(c) shows the leading DAH modes for the traveling wave. Unlike for M-SSA, the relation between DAH modes and the traveling wave data is not immediately obvious. The symmetrization step has a distinct effect on the shape of the modes.

In this paper, we propose a spectral decomposition based on lagged cross-correlations that is closely related to DAHD and M-SSA. We refer to this technique as harmonic cross-correlation decomposition (HCD). Like DAHD, HCD allows the definition of multivariate power and phase spectra and

allows for a data transformation into paired oscillatory series which are narrowband and approximately in phase quadrature. HCD is invariant under channel permutation, which is achieved by some modifications in the construction of the grand matrix. The HCD grand matrix is built from circulant Toeplitz blocks, and therefore it can be constructed as a symmetric matrix without any explicit symmetrization step. Similarly to the M-SSA modes, the HCD modes of the traveling wave [Fig. 1(b)] well approximate the data. However, in HCD they are aligned to the Fourier frequencies of the embedding interval.

It should be pointed out here that HCD as introduced in this paper is the time domain equivalent of the Hermitian DAHD formulation in the frequency domain recently introduced in [49]. The Hermitian DAHD formulation is not equivalent to the original DAHD from [43–48]. In the Supplemental Material, we demonstrate the systematic difference between the frequency domain formulation of the original DAHD and the Hermitian DAHD formulation using an example from [49]. In the body of the manuscript, “DAHD” refers to the original DAHD method as introduced in [43] and applied in [44–48]. A list of all decomposition methods discussed in this paper is provided in Table I.

The rest of the paper is organized as follows. In Sec. II we introduce HCD assuming no prior knowledge of M-SSA or DAHD. The simple traveling wave from the Introduction is used to illustrate HCD modes, power, and phase spectra. Section III briefly recalls DAHD and M-SSA and discusses the computational differences and their implications in applications to data using the traveling wave as an example. Section IV provides two further example applications of HCD: observations of a system of three linearly coupled Stuart-Landau oscillators and resting-state EEG data. For the latter, we also provide a comparison between HCD, M-SSA, and DAHD. In Sec. V we summarize the intercomparison of HCD, DAHD, and M-SSA, and we discuss our findings and perspectives for future work.

## II. HARMONIC CROSS-CORRELATION DECOMPOSITION

Harmonic cross-correlation decomposition (HCD) is inspired by data-adaptive harmonic decomposition (DAHD) first introduced in [43]. This section, therefore, follows [43] in large parts. We discuss the differences between HCD and DAHD and their implications in Sec. III A. As outlined in the Introduction, HCD is indeed the time domain equivalent of the Hermitian DAHD formulation from [49]. The original DAHD and the Hermitian DAHD formulation systematically differ (see the Supplemental Material).

For illustrative purposes, we apply the HCD method to a simple traveling sine wave with constant frequency and constant traveling speed recorded at  $d = 40$  equidistant channels. The time series of each channel are thus pure sinusoids at a fixed frequency that are shifted in phase across channels according to the traveling speed [Fig. 1(a)]. Section IV provides further example applications of the HCD method to multivariate time-series data of increasing complexity, time-series data sampled from three coupled Stuart-Landau oscillators, and finally human EEG recordings.

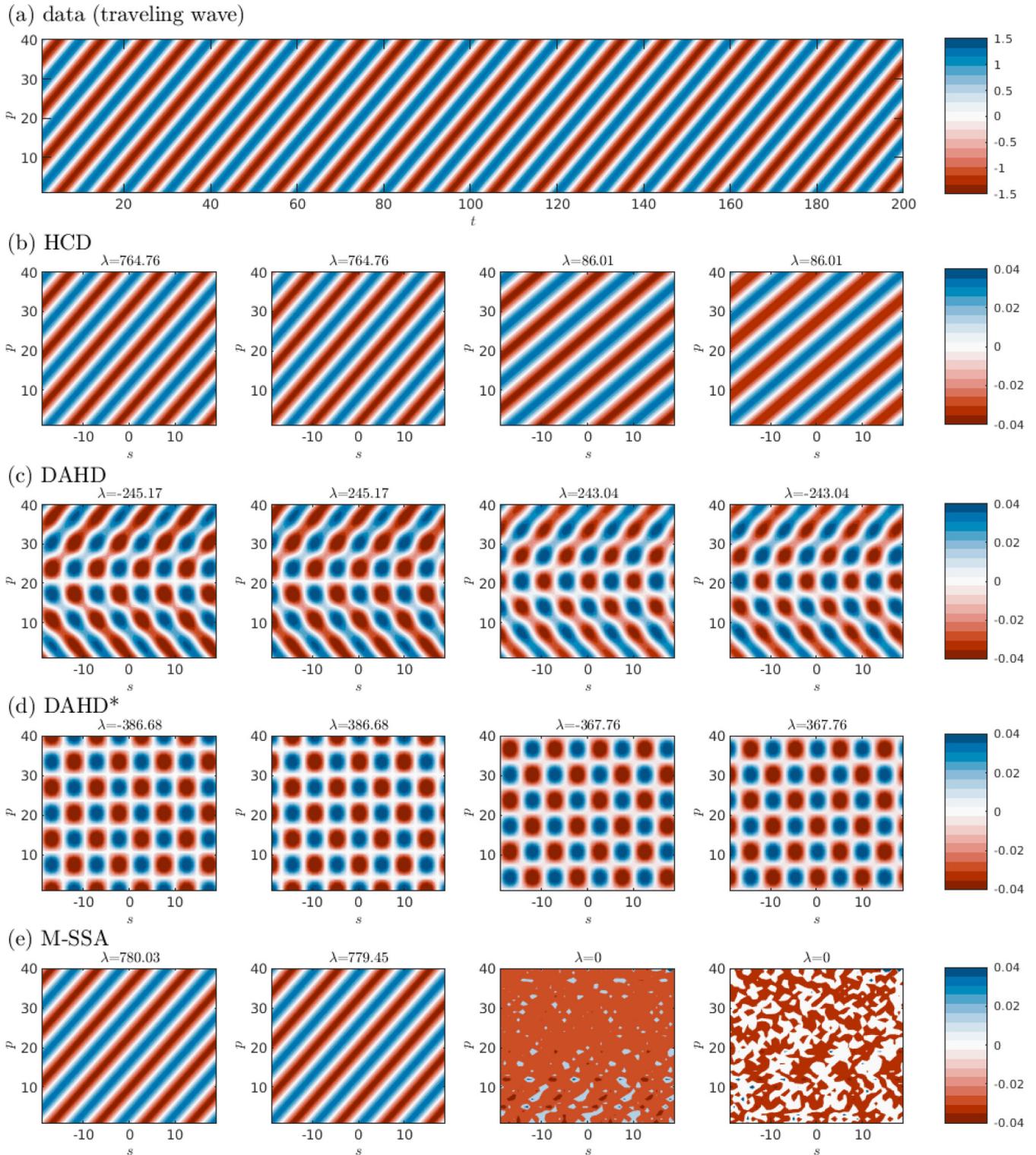


FIG. 1. A traveling wave sampled at 40 equidistant channels  $p$  (a) and the leading modes of its spectral decomposition. Shown are the two leading mode pairs across all frequencies (i.e., the modes associated with the largest absolute eigenvalues  $|\lambda|$ ) from the different decomposition techniques: (b) HCD, (c) DAHD, (d) DAHD\*, and (e) M-SSA. DAHD\* refers to a modified version of the DAHD method where the grand matrix is symmetrized by Eq. (18).

**A. Eigendecomposition of the HCD grand matrix**

The HCD method decomposes multivariate time series. Therefore, let  $\mathbf{x}$  be some multivariate time series, i.e.,

$\mathbf{x} = \{x_i^p : p = 1, \dots, d, i = 1, \dots, N\}$ , where  $d$  denotes the number of time series, which we also refer to as channels, and  $N$  is the number of data points in each time series. Each series is centralized to zero mean, such that the lagged

TABLE I. List of correlation (or covariance) -based decompositions discussed in this paper. The left column contains the abbreviation used in this paper, the middle column additionally lists alternative names for the respective method.

PCA	Principal component analysis, empirical orthogonal function (EOF) analysis, proper orthogonal decomposition, discrete Karhunen-Loeve transform, discrete Hotelling transform	(e.g., Refs. [68,69])
SSA	Singular spectrum analysis	(e.g., Refs. [40,41,50])
M-SSA	Multivariate-singular spectrum analysis, extended EOF analysis	(e.g., Refs. [40-42])
DAHD	Data-adaptive harmonic decomposition	[43-48]
HCD	Harmonic cross-correlation decomposition, Hermitian DAHD	(time domain [this paper]) (frequency domain [49])

cross-covariances  $\rho_{+s}^{(p,q)}$ ,  $\rho_{-s}^{(p,q)}$  between channel pairs  $(p, q)$  can be computed as

$$\rho_{+s}^{(p,q)} = \frac{1}{N-s} \sum_{i=1}^{N-s} x_i^p x_{i+s}^q, \quad \rho_{-s}^{(p,q)} = \frac{1}{N-s} \sum_{i=s+1}^N x_i^p x_{i-s}^q, \quad (1)$$

where series  $x^q$  is shifted with respect to  $x^p$  by  $\pm s$ . In this paper, precisely in the examples therein, the *HCD grand matrix* is built from lagged cross-covariances. By additionally nor-

malizing the time series to unit variance, one can alternatively use the cross-correlations to perform an HCD. This is up to the user as it might depend on the data set studied, if correlations or covariances are more suitable. The covariances  $\rho_{\pm s}^{(p,q)}$  up to a maximum lag of  $\pm M$  provide the entries of the HCD grand matrix. The choice of  $M$  can be guided by the decay of the autocorrelation function as suggested in [43].

The HCD grand matrix is built in two steps. For each channel pair  $(p, q)$  we first build a Toeplitz matrix  $\mathfrak{T}^{(p,q)}$  as

$$\mathfrak{T}^{(p,q)} = \begin{pmatrix} \rho_{+0}^{(p,q)} & \rho_{+1}^{(p,q)} & \cdots & \rho_{+M-1}^{(p,q)} & \rho_{-M+1}^{(p,q)} & \cdots & \rho_{-1}^{(p,q)} \\ \rho_{-1}^{(p,q)} & \rho_{+0}^{(p,q)} & \cdots & \rho_{+M-2}^{(p,q)} & \rho_{-M+1}^{(p,q)} & \cdots & \rho_{-2}^{(p,q)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \rho_{-M+1}^{(p,q)} & \rho_{-M+2}^{(p,q)} & \cdots & \rho_{+0}^{(p,q)} & \rho_{+1}^{(p,q)} & \cdots & \rho_{+M-1}^{(p,q)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \rho_{+1}^{(p,q)} & \rho_{+2}^{(p,q)} & \cdots & \rho_{-M+1}^{(p,q)} & \rho_{-M+2}^{(p,q)} & \cdots & \rho_{+0}^{(p,q)} \end{pmatrix}. \quad (2)$$

The  $\mathfrak{T}^{(p,q)}$  are circulant. That is, starting from the top each row is obtained by a cyclic shift of the row above to the right. From the  $d$  channel time series we obtain in total  $d^2$  such matrices, each of dimension  $(2M-1) \times (2M-1)$ .

Circulant matrices are diagonalized by the discrete Fourier transform, or in other words one can obtain the discrete Fourier transform of the first row of  $\mathfrak{T}^{(p,q)}$  by performing an eigendecomposition of  $\mathfrak{T}^{(p,q)}$ . The Fourier transform of auto- and cross-correlation functions is further linked to the Fourier transform of the time series through the Wiener-Khinchin theorem (for  $p=q$ ) and the cross-correlation theorem (for  $p \neq q$ ).

The  $\mathfrak{T}^{(p,q)}$  serve as building blocks to construct the HCD grand matrix  $\mathfrak{T}$  as

$$\mathfrak{T} = \begin{pmatrix} \mathfrak{T}^{(1,1)} & \mathfrak{T}^{(1,2)} & \cdots & \mathfrak{T}^{(1,d)} \\ \mathfrak{T}^{(2,1)} & \mathfrak{T}^{(2,2)} & \cdots & \mathfrak{T}^{(2,d)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathfrak{T}^{(d,1)} & \mathfrak{T}^{(d,2)} & \cdots & \mathfrak{T}^{(d,d)} \end{pmatrix}. \quad (3)$$

As  $\rho_{+s}^{(p,p)} = \rho_{-s}^{(p,p)}$ , the blocks on the diagonal which contain the autocovariances are symmetric, i.e.,  $\mathfrak{T}^{(p,p)} = \mathfrak{T}^{(p,p)T}$ .

The off-diagonal blocks which contain the cross-covariances are in general not symmetric, but as  $\mathfrak{T}^{(p,q)} = \mathfrak{T}^{(q,p)T}$  the HCD grand matrix is symmetric and thus has real eigenvalues and eigenvectors.

We compute the sets of eigenvalues  $\{\lambda_j\}$  and the associated eigenvectors  $\{\mathbf{W}_j\}$  of the HCD grand matrix  $\mathfrak{T}$ . Each eigenvector is of length  $(2M-1)d$  and can be rearranged into a spatiotemporal  $d \times (2M-1)$  *HCD mode* where the spatial dimension  $d$  corresponds to the  $d$  channels. The temporal dimension  $2M-1$  is determined by the maximum lag  $M$  considered. Each mode  $\mathbf{W}_j$  can thus be split into  $d$  snippets  $\mathbf{E}_p^j$  of length  $2M-1$ , such that each snippet vector is associated with a particular channel  $p$ . The full mode  $\mathbf{W}_j$  can then be written as

$$\mathbf{W}_j = (\mathbf{E}_1^j, \dots, \mathbf{E}_p^j, \dots, \mathbf{E}_d^j)^T. \quad (4)$$

Similarly to the DAHD modes, we observe that each snippet vector  $\mathbf{E}_p^j$  of an HCD mode takes the following shape:

$$\mathbf{E}_p^j(s) = B_p^j \cos(f_j s + \theta_p^j) \quad \text{with } 1 \leq s \leq 2M-1, \quad (5)$$

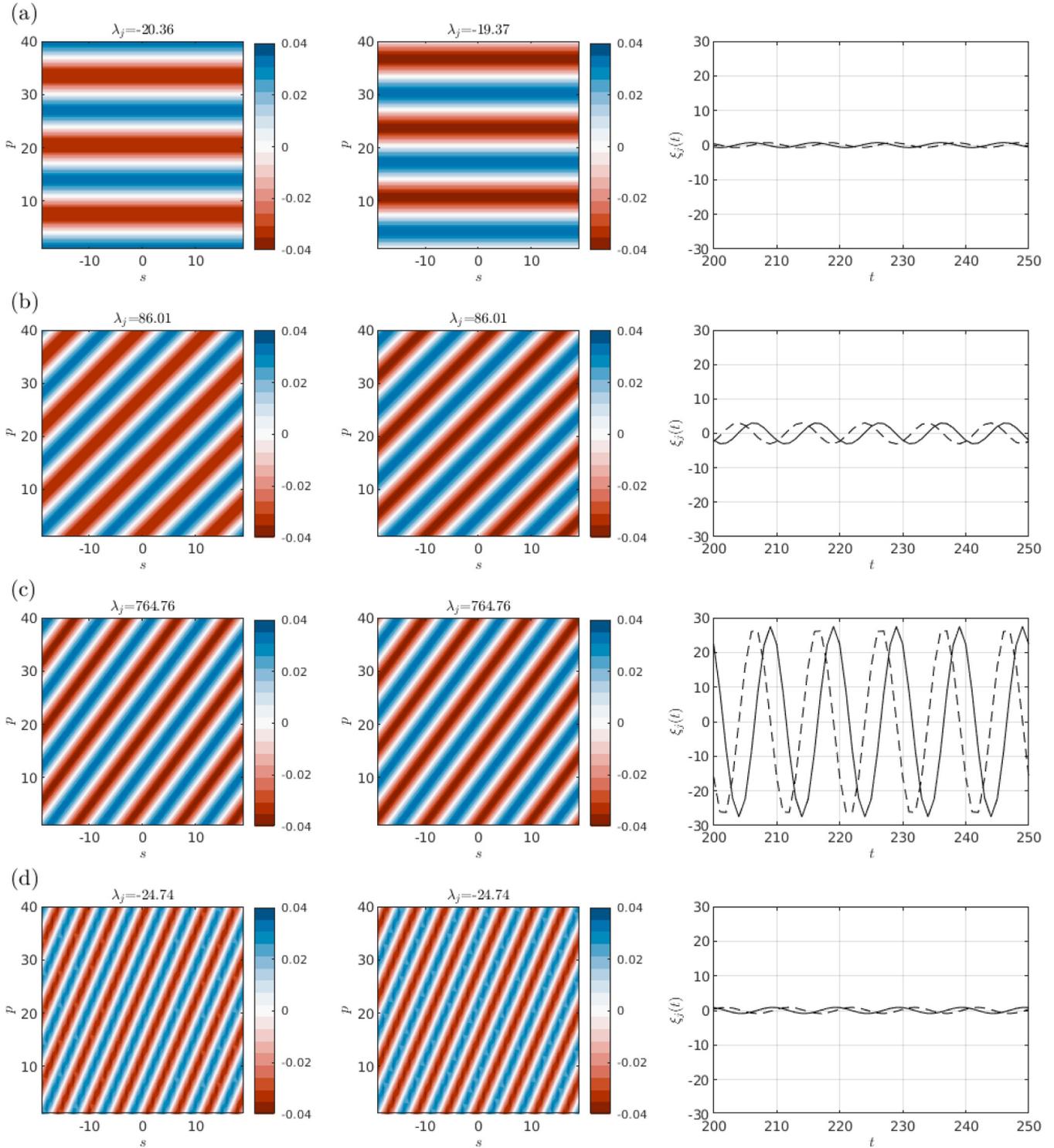


FIG. 2. HCD of a traveling wave with constant frequency and traveling speed sampled at 40 equidistant channels. Shown is the leading mode pair (i.e., the pair of modes associated with the largest absolute eigenvalues  $|\lambda|$ ) and the corresponding coefficient time series  $\xi_j(t)$  at four different frequencies: (a)  $f_1$ , (b)  $f_4$ , (c)  $f_5$ , and (d)  $f_8$ . The frequency  $f_5$  is closest to the wave frequency and hence corresponds to the peak in the HCD power spectrum shown in Fig. 3(a).

where  $f_j$  is the mode-specific frequency, meaning  $f_j$  is the same across all snippets of a single mode  $\mathbf{W}_j$ . The amplitudes  $B_p^j$  and phases  $\theta_p^j$  are snippet-specific, meaning amplitudes and phases can vary across the snippets of a single mode  $\mathbf{W}_j$ . Amplitudes and phases depend on the data.

Figure 2 shows the leading two HCD modes for the traveling wave at  $f_1 = 0$  and the leading mode pairs at three different  $f_i \neq 0$ , i.e., the modes corresponding to the largest eigenvalues at the respective frequency. The oscillatory modes show constant amplitudes  $B_j$  across all snippets. The

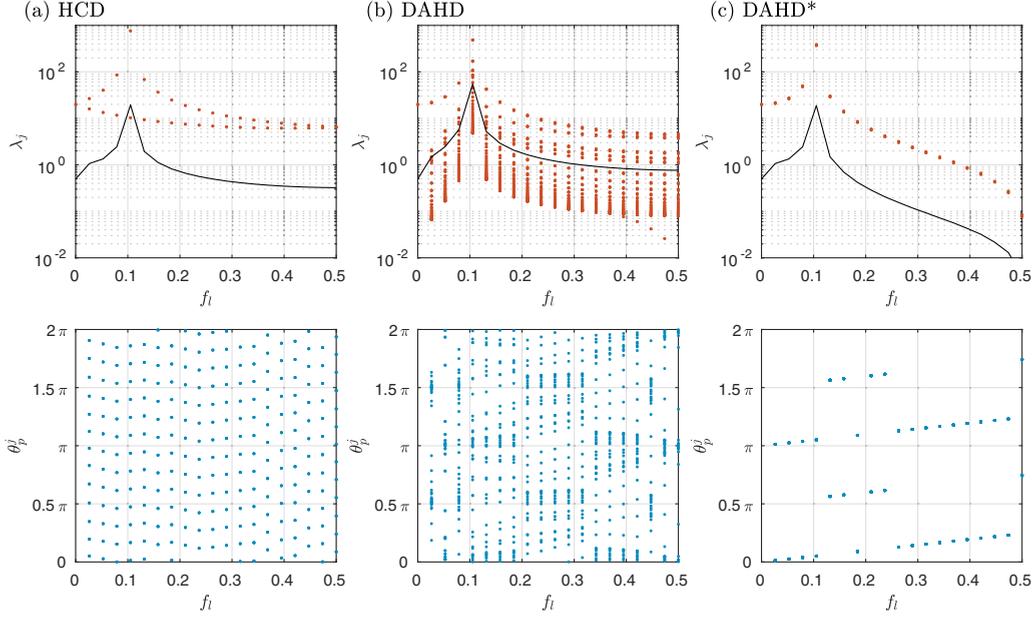


FIG. 3. Multidimensional power spectra (top) and leading mode phase spectra (bottom) from the spectral decomposition of the traveling wave: (a) shows the HCD spectra, (b) shows the DAHD spectra, and (c) shows the spectra from the modified version of the DAHD method where the grand matrix is symmetrized by Eq. (18). The back line in the upper panels shows the average over the eigenvalues at the respective frequency.

constant traveling speed induces constant phase differences  $\Delta\theta = \theta_{p+1} - \theta_p$  between neighboring channels.

### B. HCD power and phase spectra

The frequencies of the eigenvectors of the HCD grand matrix  $f_j$  align with the Fourier frequencies  $f_l$ ,

$$f_l = \frac{\pi(l-1)}{M-1} \quad \text{with } l = 1, \dots, M \quad (6)$$

such that we obtain  $2d$  eigenvalues and vectors at each  $f_l > 0$  (similar to Theorem V.1 of [43]). These  $2d$  eigenvalues come in pairs of equal values with the corresponding eigenvectors differing solely by a phase shift of  $\pi/2$ . At  $f_l = 0$  we obtain  $d$  unpaired eigenvalues and vectors. Because each mode  $\mathbf{W}_j$  is associated with a Fourier frequency through Eq. (5), this allows the definition of multivariate power and phase spectra. The eigenvalues sorted according to the frequencies of the associated modes form the *HCD power spectrum*. Because of the eigenvalue pairing, the power spectrum is of dimension  $d \times M$  and linearly spaced in frequency [Eq. (6)].

Computationally, the HCD power spectrum is derived by rearranging the eigenvalues  $\{\lambda_j\}$  into  $M$  sets

$$\mathcal{P}_l := \{|\lambda_j| : j \in \mathcal{I}(f_l)\}, \quad (7)$$

each containing all eigenvalues associated with the respective  $f_l$ , i.e.,  $\mathcal{I}(f_l) := \{j : \text{s.t. } f_j = f_l\}$ . This can be realized by carrying out the following successive steps for each mode  $\mathbf{W}_j$ : (i) Compute the discrete Fourier transform  $\widehat{\mathbf{E}}_p^j$  of each mode snippet  $\mathbf{E}_p^j$ . (ii) Take the average of the power spectral density over all snippets of  $\mathbf{W}_j$ , i.e.,  $\overline{\Gamma}_j = \frac{1}{d} \sum_{p=1}^d |\widehat{\mathbf{E}}_p^j|$ . (iii) Identify the frequency  $f_l = f_j$  at which the average power

spectral density  $\overline{\Gamma}_j$  exhibits a dominant peak. (iv) Sort the corresponding eigenvalue  $\lambda_j$  into the respective set  $\mathcal{P}_l$ . Note that eigenvectors associated with eigenvalues that are numerically zero do not fulfill Eq. (5) and thus cannot be assigned a frequency.

The HCD power spectrum of the traveling sine wave is shown in Fig. 3. It peaks at the frequency  $f_l$  closest to the frequency of the wave. The steepness of the decay from the peak is related to the distance between wave frequency and closest  $f_l$ . The smaller the difference, the more narrow the power spectrum becomes. For strongly correlated data, such as a traveling wave, a great portion of the  $d(2M-1)$  eigenvalues are numerically zero. For the traveling wave we find two nonzero pairs of eigenvalues at each  $f_l$ . One might expect only one nonzero pair per frequency. In fact, for a standing wave with the same frequency we indeed find only one nonzero pair of eigenvalues at each frequency (not shown). The occurrence of two nonzero pairs for the traveling wave appears to be a consequence of the wave being sampled at different phases at different channels.

The snippet phases  $\theta_p^j$  [compare Eq. (5)] form the *HCD phase spectrum*, i.e.,

$$\Phi_l := \{\theta_p^j(f_l) : j \in \mathcal{I}(f_l)\}. \quad (8)$$

The snippet phases  $\theta_p^j$  are computed by taking the angle of the discrete Fourier transform at the mode frequency  $f_l$ , i.e.,

$$\theta_p^j(f_l) = \arg(\widehat{\mathbf{E}}_p^j(f_l)). \quad (9)$$

The full phase spectrum is of dimension  $d \times d \times (M-1)$ . When plotting the phase spectra, we restrict ourselves to the leading modes, i.e., the modes with the largest eigenvalue at

each  $f_i$ . The part of the phase spectrum plotted is thus of dimension  $d \times (M - 1)$ . As the leading modes contain the dominant oscillatory signals present in the data, such phase spectra can provide insight into characteristic phase differences between the time series recorded at distinct channels.

Figure 3 shows the leading mode phase spectrum of the traveling wave. As expected, we find the phases distributed equidistantly within each frequency as the wave is recorded at equidistant channels and traveling with a constant speed.

**C. Data transformation**

HCD modes, power, and phase spectra provide tools for obtaining frequency-specific insight into correlation structure and phase relations in multivariate time-series data. The HCD modes further provide an orthogonal basis allowing an alternative representation of a multivariate time series  $\mathbf{x}$  though a projection onto the HCD modes  $\{\mathbf{W}_j\}$  as

$$\xi_i^j := \sum_{s=1}^{2M-1} \sum_{p=1}^d x_{i+s-1}^p \mathbf{E}_p^j(s) \quad \text{with } 1 \leq i \leq N - 2M + 2. \tag{10}$$

$$\begin{aligned} m_i &= 2M - 1, \quad l_i = 1, \quad u_i = 2M - 1 \quad \text{for } 2M - 1 \leq i \leq N - 2M + 2, \\ m_i &= i, \quad l_i = 1, \quad u_i = i \quad \text{for } 1 \leq i \leq 2M - 2, \\ m_i &= N - i + 1, \quad l_i = i - N + (2M - 1), \quad u_i = 2M - 1 \quad \text{for } N - 2M + 3 \leq i \leq N. \end{aligned} \tag{12}$$

The multivariate time series represented by a single coefficient series  $\xi^j$  is thus  $R^j = [R_1^j, \dots, R_p^j, \dots, R_d^j]$ . Summation over all  $j$  gives back the original data time series  $\mathbf{x}$ , i.e.,

$$\mathbf{x} = \sum_{j=1}^{d(2M-1)} R^j. \tag{13}$$

Summation over a subset of the coefficient series such as the series corresponding to the leading  $n$  modes or a particular

$$\mathbf{H}^{(p,q)} = \begin{pmatrix} \rho_{-M+1}^{(p,q)} & \rho_{-M+2}^{(p,q)} & \dots & \rho_0^{(p,q)} & \rho_1^{(p,q)} & \dots & \rho_{M-1}^{(p,q)} \\ \rho_{-M+2}^{(p,q)} & \rho_{-M+3}^{(p,q)} & \dots & \rho_1^{(p,q)} & \rho_2^{(p,q)} & \dots & \rho_{-M+2}^{(p,q)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \rho_{M-1}^{(p,q)} & \rho_{-M+1}^{(p,q)} & \dots & \rho_{-1}^{(p,q)} & \rho_0^{(p,q)} & \dots & \rho_{M-2}^{(p,q)} \end{pmatrix}, \tag{14}$$

where again  $\rho_{+s}^{(p,q)}$ ,  $\rho_{-s}^{(p,q)}$  denote the lagged cross-covariances (or correlations) between the time series recorded at channels  $p$  and  $q$  [Eq. (1)]. The Hankel blocks  $\mathbf{H}^{(p,q)}$  are left-circulant. That is, starting from the top each row is obtained by a cyclic shift of the row above to the left. As shown in Theorem IV.1 of [43], left-circulants  $\mathbf{H}^{(p,q)}$  and Fourier transform are closely

The  $\xi_j^i$  are referred to as *HCD coefficients*. From the original time series  $\mathbf{x}$  one obtains  $d(2M - 1)$  HCD coefficient series  $\xi_j^i$ . The representation of  $\mathbf{x}$  in terms of HCD coefficients, therefore, comes with an increase of dimension. The potential usefulness of this representation lies in the particular characteristics of the coefficient series. For frequency-rich data with a non-narrow (Fourier) power spectrum (such as the EEG recordings discussed in Sec. IV B), the paired oscillatory coefficient series are typically narrowband around the mode frequency  $f_j$  and approximately in phase quadrature. For the traveling sine wave, however, which only contains a single frequency, the coefficient series oscillate with the frequency of the wave (Fig. 2).

The back transformation of the coefficient time series into channels space can accordingly be obtained by

$$R_{pi}^j := \frac{1}{m_i} \sum_{s=l_i}^{u_i} \xi_{i-s+1}^j \mathbf{E}_p^j(s) \tag{11}$$

with normalization factor  $\frac{1}{m_i}$  and summation bounds  $l_i, u_i$  differing between the center of the time series and its edges (see, e.g., Ref. [50]) as

frequency  $j \in \mathcal{I}(f_i)$  (or a range of frequencies) allows us to obtain filtered versions of  $\mathbf{x}$ .

**III. RELATION TO EXISTING SPECTRAL DECOMPOSITIONS**

**A. DAHD**

Computationally, DAHD as introduced in [43] differs from HCD introduced in this paper in the construction of the grand matrix. The DAHD grand matrix  $\mathfrak{H}$  is built from Hankel blocks  $\mathbf{H}^{(p,q)}$  given by

linked, in the sense that one can recover the Fourier transform of the first row of  $\mathbf{H}^{(p,q)}$  from the eigendecomposition of  $\mathbf{H}^{(p,q)}$ .

Left-circulant matrices are by construction symmetric. Each  $\mathbf{H}^{(p,q)}$  thus has real eigenvalues and eigenvectors. It may therefore appear favorable to build a grand matrix from

left-circulant blocks  $\mathbf{H}^{(p,q)}$  when aiming at a real valued decomposition. However, the construction of a symmetric grand matrix from left-circulants is not straightforward. The *DAHD grand matrix*  $\mathfrak{H}$  is a block matrix similar to the HCD grand matrix, i.e.,

$$\mathfrak{H} = \begin{pmatrix} \mathfrak{H}^{(1,1)} & \mathfrak{H}^{(1,2)} & \dots & \mathfrak{H}^{(1,d)} \\ \mathfrak{H}^{(2,1)} & \mathfrak{H}^{(2,2)} & \dots & \mathfrak{H}^{(2,d)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathfrak{H}^{(d,1)} & \mathfrak{H}^{(d,2)} & \dots & \mathfrak{H}^{(d,d)} \end{pmatrix}, \quad (15)$$

but to ensure eigenvalues and eigenvectors of  $\mathfrak{H}$  are real, the grand matrix blocks  $\mathfrak{H}^{(p,q)}$  are identified with the  $\mathbf{H}^{(p,q)}$  from Eq. (14) as

$$\mathfrak{H}^{(p,q)} = \mathbf{H}^{(p,q)} \quad \text{for } p \leq q, \quad \mathfrak{H}^{(p,q)} = \mathbf{H}^{(q,p)} \quad \text{for } q < p. \quad (16)$$

This permutation of the indices in the lower triangle is needed to ensure that the DAHD grand matrix is symmetric, because  $\mathbf{H}^{(p,q)} \neq \mathbf{H}^{(q,p)^T}$  for nonsymmetric  $\rho^{(p,q)}$ .

The eigendecomposition of the DAHD grand matrix gives  $(2M - 1)d$  eigenvectors  $\mathbf{W}_j$ . The eigenvectors are rearranged into spatiotemporal DAHD modes of dimension  $d \times (2M - 1)$ , such that each mode can be split into snippets [compare Eq. (4)]. The snippets can then be written as cosines of the Fourier frequency  $f_i$  of the respective mode with a snippet-specific amplitude and phase [compare Eq. (5)]. According to the variational principle of [43], there are  $2d$  pairs of eigenvalues and eigenvectors at each Fourier frequency  $f_i$ , and  $d$  unpaired eigenvalues at  $f_i = 0$ , which gives rise to DAHD power and phase spectra. The projection of the multivariate time series onto the DAHD modes provides an alternative representation of the data in terms of oscillatory pairs narrow-band around the mode frequency and approximately in phase quadrature [compare Eq. (10)]. The only immediately obvious difference between DAHD and HCD here lies in the signs of the eigenvalues. The eigenvalues of the DAHD grand matrix come in pairs with the same value but opposite sign, whereas the eigenvalue pairs of the HCD grand matrix are equal in value and sign.

When aiming for frequency-specific analyses of correlation and phase relation structure in multivariate time-series data, DAHD comes with a drawback. DAHD modes, power, and phase spectra are systematically influenced by the ordering of the channels within the DAHD grand matrix. Or in other words, DAHD modes, power, and phase spectra are not invariant under a permutation of channels. The dependency of the DAHD on channel ordering is a consequence of the symmetrization of the DAHD grand matrix according to Eq. (16). We illustrate this using time series from EEG recordings. Details about data and further analysis are given in Sec. IV B. For the illustrative example in this section, we construct a reduced data set  $\tilde{\mathbf{x}}$  from the original 64-channel EEG recordings  $\mathbf{x}$  by setting

$$\begin{aligned} \tilde{x}^p(t) &= x^1(t) \quad \text{for } p = 1, \dots, 29, 41, \dots, 64, \\ \tilde{x}^p(t) &= x^2(t) \quad \text{for } p = 30, \dots, 40. \end{aligned} \quad (17)$$

That is, the reduced 64-channel data practically contain only two different time series.

Figure 4(b) shows the DAHD power and (leading mode) phase spectra from the reduced EEG data  $\tilde{\mathbf{x}}$ . For each  $f_i \neq 0$  we find three pairs of nonzero eigenvalues. This is surprising. As  $\tilde{\mathbf{x}}$  contains only two different series, we would expect two pairs of nonzero eigenvalues at each  $f_i \neq 0$ . Looking at the DAHD modes in Fig. 5 provides a bit more insight. In Fig. 5 we show the DAHD modes associated with the four largest eigenvalues at  $f_5 = 2.02$  Hz. The series at  $p = 1, \dots, 29, 41, \dots, 64$ , though all exactly the same [Eq. (17)], appear differently in the DAHD modes for  $p > 30$  and  $p < 40$ . Most prominently, the mode snippets of channels 1–29 are shifted in phase with respect to the snippets of channels 41–64 in panels (b) and (c), which is not meaningful as the corresponding time series are exactly the same. The position of the channels within the grand matrix affects their relation (relative amplitude and phase) within the modes. In the HCD method this is not the case (Fig. 6). Without symmetrization, i.e., by building the full DAHD grand matrix as  $\mathfrak{H}^{(p,q)} = \mathbf{H}^{(p,q)}$ , one obtains only two nonzero pairs at each frequency (not shown), but as  $\mathbf{H}^{(q,p)} \neq \mathbf{H}^{(p,q)^T}$  for nonsymmetric  $\rho^{(p,q)}$  such a grand matrix is not symmetric and can therefore have nonreal eigenvalues and eigenvectors.

Looking at the DAHD spectra for the traveling sine wave (Fig. 3) allows us to obtain more insight into the above finding. For constant traveling speed and equidistant channels, one would expect the constant phase differences between neighboring channels to be reproduced in the phase spectrum. However, we do find a qualitatively different shape in the DAHD phase spectrum. The DAHD phase spectrum converges for  $f_i \rightarrow f_{\max}$ .

By looking at the DAHD modes of the traveling wave in Fig. 1, it becomes obvious that the modes do not reflect the phase structure of the original data. In fact, the upper half (channels 1–10) is found to be the negative mirror image of the lower half of the mode (channels 11–20). Only at the bottom end of the mode ( $\approx$  channels 16–20) do we observe approximately the expected structure. The top ( $\approx$  channels 1–4) approximately resembles a traveling wave in the opposite direction.

When we go back to the symmetrization of the DAHD grand matrix [Eq. (16)] and the definition of the Hankel blocks [Eq. (14)], we see that the permutation of the indices  $p$  and  $q$  in the blocks on the lower triangle of the DAHD grand matrix is in fact equivalent to reversing the cross-correlation function. A shift of  $+1$  of channel  $p$  with respect to channel  $q$  is equivalent to shifting  $q$  by  $-1$  with respect to  $p$ . It is therefore unsurprising that the symmetrization of the grand matrix systematically impacts the representation of phase relationships.

When building a grand matrix from Toeplitz blocks, or more precisely right-circulant matrices, as done in the HCD method, there is no need for an explicit symmetrization. Unlike the Hankel blocks, the (off-diagonal) Toeplitz blocks are not symmetric, but as already noted in Sec. II due to  $\mathfrak{T}^{(p,q)} = \mathfrak{T}^{(q,p)^T}$  the HCD grand matrix constructed according to Eq. (3) is symmetric and thereby has real eigenvalues and eigenvectors. It thus appears that we can keep the advantageous

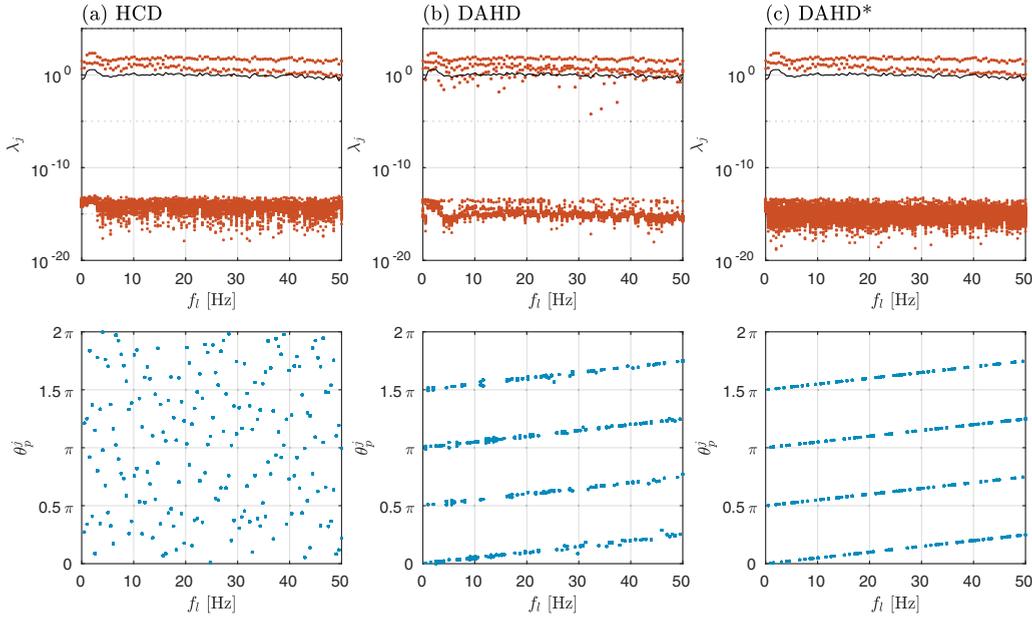


FIG. 4. Multidimensional power spectra (top) and leading mode phase spectra (bottom) from the spectral decomposition of the reduced EEG recordings constructed according to Eq. (17): (a) shows the HCD spectra from the EEG recording, (b) shows the DAHD spectra, and (c) shows the spectra from the modified version of the DAHD method where the grand matrix is symmetrized by Eq. (18). The back line in the upper panels shows the average over the eigenvalues at the respective frequency.

features of the DAHD method in the sense that we obtain a real multivariate harmonic decomposition that allows us to express a given data set in terms of narrowband oscillatory pairs. But by ensuring invariance under channel permutation, we achieve an intuitive relation between the HCD phases and the phase relations in the actual traveling wave data. Further, in the HCD method one obtains only two pairs of nonzero eigenvalues at each nonzero  $f_i$  [Fig. 3(a)], meaning the traveling wave can be fully described by  $2(2M - 1)$  coefficient series in the space spanned by the HCD modes, while in the space spanned by the DAHD modes all  $d(2M - 1)$  coefficient series are needed.

Equation (16) is not the only way to obtain a symmetric grand matrix from Hankel blocks. Alternatively, one may identify the grand matrix blocks  $\mathcal{H}^{(p,q)}$  from Eq. (15) with the Hankel blocks  $\mathbf{H}^{(q,p)}$  from Eq. (14) as

$$\mathcal{H}^{(p,q)} = \frac{1}{2}(\mathbf{H}^{(p,q)} + \mathbf{H}^{(q,p)}). \quad (18)$$

The eigendecomposition of such a grand matrix is invariant under channel permutation. However, the symmetrization according to Eq. (18) is equivalent to symmetrizing the cross-correlation functions. Accordingly, the snippets of the modes obtained from the eigendecomposition of such a grand matrix are all perfectly in phase or antiphase and are therefore not capable of capturing phase relations between time series. Example modes from the traveling wave illustrating this behavior are shown in Fig. 1(d). For the traveling wave one obtains, as for HCD, two pairs of nonzero eigenvalues at each nonzero  $f_i$ , but as the actual phase relations of the data are not captured in the modes, the two pairs are almost equal in value [Figs. 3(c) and 1(d)]. As expected, the corresponding phase spectra are not meaningful. We solely find phase differences of  $\Delta\theta \in \{0, \pi\}$  across all frequencies  $f_i > 0$  [Fig. 3(c)].

### B. M-SSA

Like DAHD and HCD, multichannel (or multivariate) singular spectrum analysis (M-SSA) utilizes an eigendecomposition of a grand matrix containing lagged cross-correlations. M-SSA can be performed by building a grand matrix from Toeplitz blocks or via a delay embedding. Note that in contrast to the blocks of the HCD grand matrix, the Toeplitz blocks occurring in M-SSA are not circulant. Again, let  $\mathbf{x} = \{x_i^p : p = 1, \dots, d, i = 1, \dots, N\}$  be some multivariate time series, where  $d$  denotes the number of channels,  $N$  denotes the length of the series, and each series is centralized to zero mean. In the delay embedding approach to M-SSA, each channel of the multivariate time series is embedded into an  $M$ -dimensional phase space as

$$X_i^p = [x_i^p, x_{i+1}^p, \dots, x_{i+M-1}^p] \quad \text{with } i = 1, \dots, N - M + 1. \quad (19)$$

The full trajectory matrix of the delay embedded series  $\mathbf{X} = (X^1, X^2, \dots, X^d)$  is thus of dimension  $(N - M + 1) \times dM$ . The covariance matrix

$$\mathcal{C} = \mathbf{X}^T \mathbf{X} / (N - M + 1) \quad (20)$$

from the trajectory matrix is thus of dimension  $dM \times dM$ . When writing out the covariance matrix, we see that  $\mathcal{C}$  consists of  $d^2$  square blocks  $\mathbf{C}^{(p,q)}$ ,

$$\mathcal{C} = \begin{pmatrix} \mathbf{C}^{(1,1)} & \mathbf{C}^{(1,2)} & \dots & \mathbf{C}^{(1,d)} \\ \mathbf{C}^{(2,1)} & \mathbf{C}^{(2,2)} & \dots & \mathbf{C}^{(2,d)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}^{(d,1)} & \mathbf{C}^{(d,2)} & \dots & \mathbf{C}^{(d,d)} \end{pmatrix}, \quad (21)$$

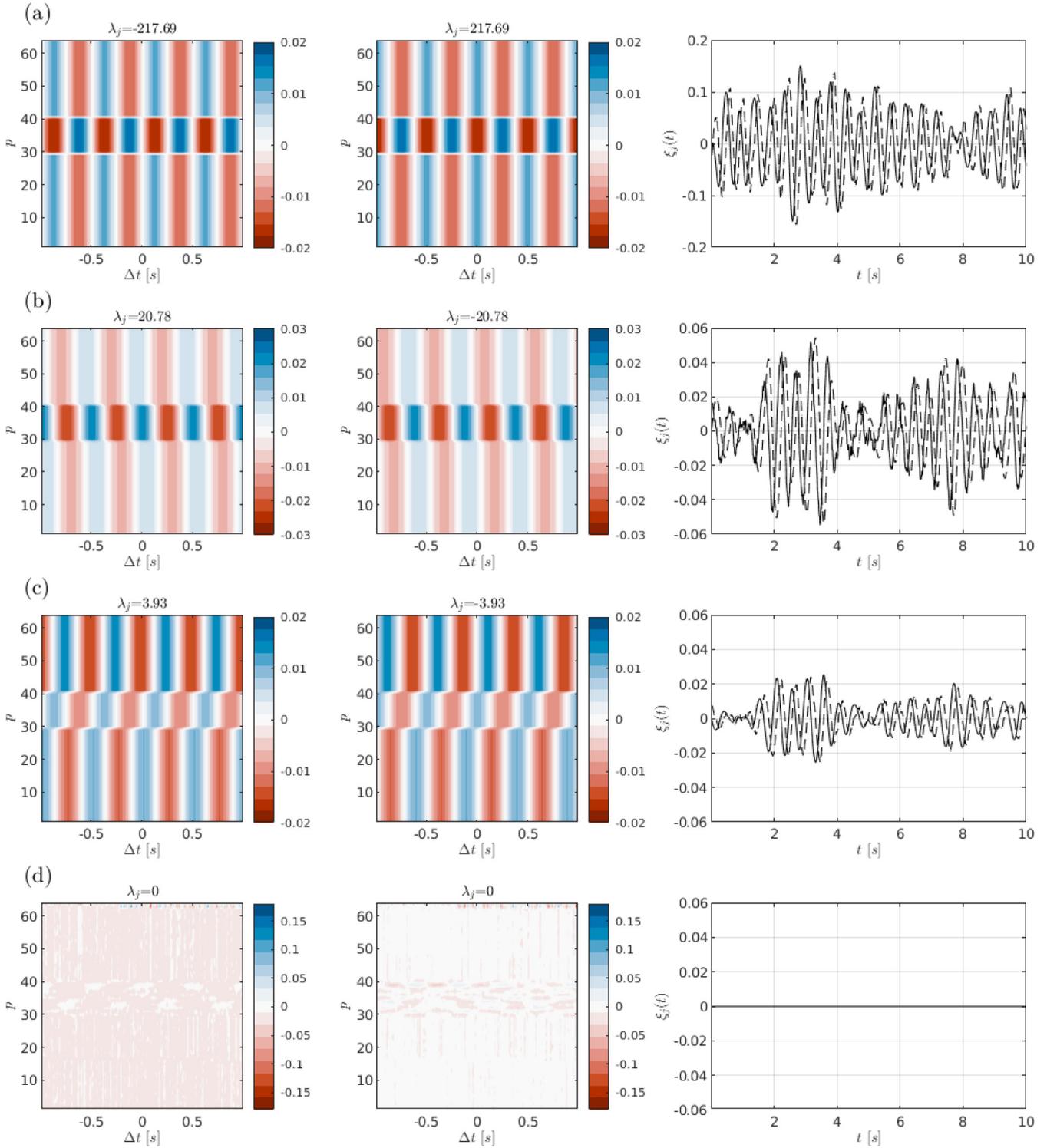


FIG. 5. DAHD of the reduced EEG recordings constructed according to Eq. (17): Shown are the leading four mode pairs (i.e., the mode pairs associated with the four largest absolute eigenvalues  $|\lambda|$ ) and the corresponding coefficient time series  $\xi_j(t)$  (right) at  $f_5 = 2.02$  Hz.

with each  $M \times M$  block being of the following shape:

$$\mathbf{C}^{(p,q)} = \begin{pmatrix} \rho(x_{+0}^p, x_{+0}^q) & \rho(x_{+0}^p, x_{+1}^q) & \cdots & \rho(x_{+0}^p, x_{+M-1}^q) \\ \rho(x_{+1}^p, x_{+0}^q) & \rho(x_{+1}^p, x_{+1}^q) & \cdots & \rho(x_{+1}^p, x_{+M-1}^q) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(x_{+M-1}^p, x_{+M-1}^q) & \rho(x_{+M-1}^p, x_{+1}^q) & \cdots & \rho(x_{+M-1}^p, x_{+M-1}^q) \end{pmatrix} \quad (22)$$

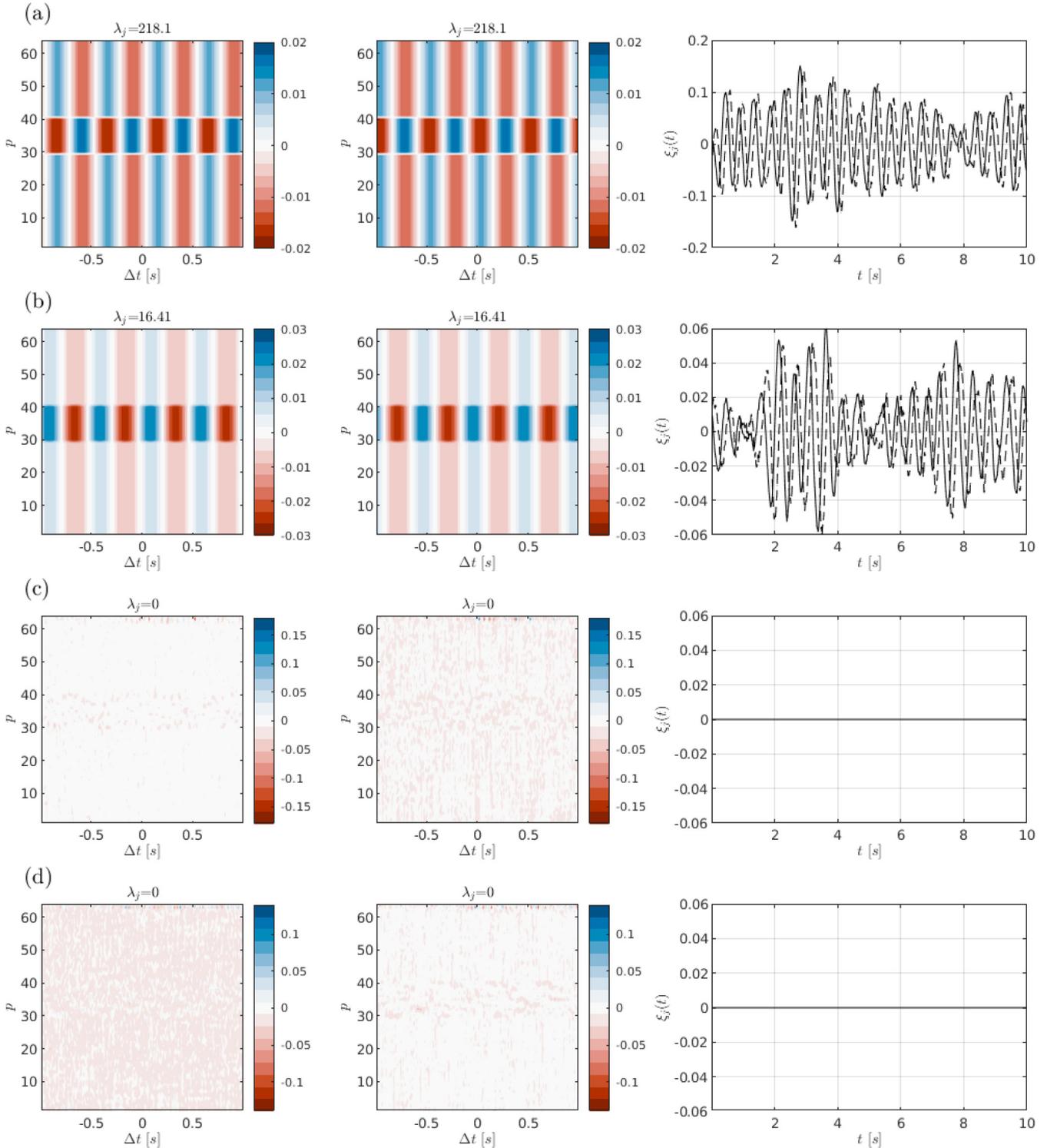


FIG. 6. HCD of the reduced EEG recordings constructed according to Eq. (17): Shown are the leading four mode pairs (i.e., the mode pairs associated with the four largest absolute eigenvalues  $|\lambda_j|$ ) and the corresponding coefficient time series  $\xi_j(t)$  (right) at  $f_5 = 2.02$  Hz.

with entries given by

$$\rho(x_{+s_1}^p, x_{+s_2}^q) = \sum_{i=1}^{N-M+1} \frac{x_{i+s_1}^p x_{i+s_2}^q}{N-M+1}. \quad (23)$$

The M-SSA modes can then be obtained either from an eigendecomposition of the above covariance matrix or from

a singular value decomposition  $\mathfrak{C} = U \Lambda V^T$  as the columns of  $U$  (see, e.g., Ref. [40]). Due to Eq. (20),  $\mathfrak{C}$  is symmetric and thereby has real eigenvalues and eigenvectors.

The block structure of  $\mathfrak{C}$  [Eq. (21)] reminds one of HCD and DAHD grand matrices [Eqs. (3) and (15)]. With an embedding dimension of  $M$  the M-SSA blocks are of

dimension  $M \times M$  instead of  $(2M - 1) \times (2M - 1)$  as in HCD and DAHD.

Accordingly, the M-SSA modes are of lower (spatial) dimension than in HCD and DAHD when choosing  $M$  to be equal. For all example applications of M-SSA we pick a larger  $M$  for M-SSA, such that the M-SSA modes we obtain have the same dimension as DAHD and HCD modes.

The blocks  $\mathbf{C}^{(p,q)}$  [Eq. (22)] of the M-SSA grand matrix obtained from the embedding approach are not Toeplitz. They can, however, be made Toeplitz by replacing the covariance (or correlation) estimates  $\rho(x_{+s_1}^p, x_{+s_2}^q)$  from Eq. (23) by  $\rho(x_0^p, x_s^q)$  as defined in Eq. (1) with  $s = s_2 - s_1$ ,

$$\mathbf{C}^{(p,q)} = \begin{pmatrix} \rho_{+0}^{(p,q)} & \rho_{+1}^{(p,q)} & \cdots & \rho_{+M-1}^{(p,q)} \\ \rho_{-1}^{(p,q)} & \rho_0^{(p,q)} & \cdots & \rho_{+M-2}^{(p,q)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{-M+1}^{(p,q)} & \rho_{-M+2}^{(p,q)} & \cdots & \rho_0^{(p,q)} \end{pmatrix}. \quad (24)$$

Due to  $\rho_s^{(p,q)} = \rho_{-s}^{(q,p)}$  and thereby  $\mathbf{C}^{(p,q)} = \mathbf{C}^{(q,p)^T}$ , the M-SSA grand matrix built from the Toeplitz blocks is by construction symmetric. The correlation estimates from Eqs. (23) and (1) are different for finite time series. Therefore, embedding M-SSA and Toeplitz M-SSA are not fully equivalent.

As in HCD and DAHD, a portion of the M-SSA modes are typically oscillatory pairs approximately in phase quadrature. But unlike for HCD and DAHD modes, there is no exact pairing in M-SSA. Pairs can typically only be made in terms of similar eigenvalues and eigenvectors. The identification of pairs in M-SSA is thus not necessarily straightforward. One may view M-SSA as the most data-adaptive of three correlation-based decompositions discussed in this paper in the sense that not only phases and amplitudes of the modes but also their temporal structure is data-adaptive. M-SSA modes do not necessarily fall onto Fourier frequencies. The frequencies of the M-SSA modes depend on the data.

M-SSA of a traveling sine wave yields only two nonzero eigenvalues belonging to a single oscillatory pair of modes with the frequency of the wave (Fig. 1). For the traveling wave data, the M-SSA modes are sinusoids along the temporal dimension. This is, however, due to the sinusoidal channel time series for the traveling wave. Unlike HCD and DAHD modes, the oscillatory M-SSA modes are typically neither sinusoids nor uniquely assigned to a specific frequency, as can be seen in our subsequent example application to EEG data. The identification of phase relation from M-SSA modes is therefore less straightforward.

#### IV. EXAMPLES

##### A. Stuart-Landau oscillators

As a test scenario of intermediate complexity, we consider a system of coupled Stuart-Landau (Hopf normal form) oscillators that are linearly combined to produce channel observations. A system of  $n$  linearly coupled Stuart-Landau oscillators  $z_k \in \mathbb{C}$  with  $k = 1, \dots, n$  is governed by

$$\dot{z}_k = (\rho_k + i\omega_k)z_k - |z_k|^2 z_k + \epsilon \left( E + \sum_{j \neq k, j=1}^n A_{kj} \text{Re}(z_j) \right). \quad (25)$$

We simulate this system of  $n = 3$  Stuart-Landau oscillators with natural angular frequencies  $\omega_k$  chosen to be

$$\omega_1 = 3, \quad \omega_2 = 6.04, \quad \omega_3 = 9.11$$

close to a 1 : 2 : 3 resonance. We choose linear growth rates  $\rho_k = 1$  and identical coupling  $A_{kj} = 1$  for all  $j, k$ . The constant  $E = 1$  breaks the rotational symmetry  $z \mapsto e^{i\psi} z$  of the Stuart-Landau system and ensures a generic breakdown of the three-frequency motion on increasing the coupling strength  $\epsilon$ . We choose two different values of  $\epsilon$  to obtain two qualitatively different attractors for the coupled system:

$$\epsilon = 1 \text{ (QP1)}, \quad \epsilon = 0.15 \text{ (QP3)}.$$

On increasing coupling strength  $\epsilon$ , there is a sequence of bifurcations (not shown here) leading to full mode-locking [51]. For strong coupling ( $\epsilon = 1$ ), all three oscillators are fully mode-locked into a periodic behavior (QP1) with a 1 : 2 : 3 ratio of frequencies. For small coupling ( $\epsilon = 0.15$ ), there is no locking of oscillators and the attractor is three frequency quasiperiodic (QP3).

To obtain time series  $z_i^k$  with  $i = 1, \dots, N$  we integrate Eq. (25) for all three scenarios over a time span of  $T = 5000$  ( $dt = 0.05$ ). Thus we obtain series of length  $N = 10000$  from which we generate a three-channel observation ( $d = 3$ ) and an eight-channel observation ( $d = 8$ )  $x_i^p \in \mathbb{R}$  as

$$x_i^p = \sum_{k=1}^n P_{pk} \text{Re}(z_i^k) + Q_{pk} \text{Im}(z_i^k). \quad (26)$$

See the Supplemental Material [52] for figures showing the observation time series. In the three-channel observation, each channel records signals from a single oscillator only. In the eight-channel observation, the channels record mixed signals. Precisely, our three-channel recording corresponds to picking  $P_{pk}$  and  $Q_{pk}$  as

$$P_{pk} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad Q_{pk} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (27)$$

That is, the three-channel observations are given by the real part of the oscillator time series  $x_i^p = \text{Re}(z_i^p)$ .

To generate an eight-channel observation, we pick

$$P_{pk} = \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 0.2 & 0.4 \\ 0 & 0.1 & 0.3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0.1 & 0.4 & 0.8 \\ 0.9 & 0.6 & 0.2 \end{pmatrix}, \quad Q_{pk} = \begin{pmatrix} 0 & 0 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 0.7 & 1.2 \\ 0 & 0.3 & 0.4 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0.3 & 0.7 & 2 \\ 0.9 & 1.4 & 0.3 \end{pmatrix}. \quad (28)$$

That is, channels 1–4 record signals from two oscillators. Channels 5–6 record signals from a single oscillator only. Channels 7–8 record signals from all three oscillators. The oscillators contribute to the different channel observations with different phases, as we use different weightings of real and imaginary parts. The above choices of  $P_{pk}$  and  $Q_{pk}$  are arbitrary and solely aimed at emulating observations in which

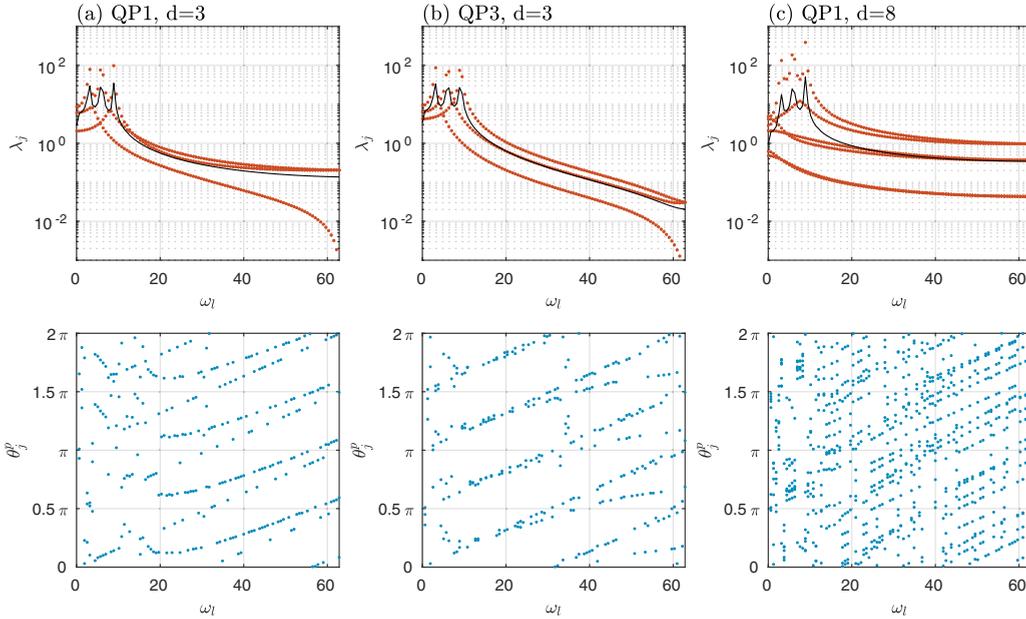


FIG. 7. HCD power spectra (top) and leading mode phase spectra (bottom) of the three-oscillator system: (a) shows the spectra from the three-channel recording of configuration QP1, (b) shows the spectra from the three-channel recording of configuration QP3, (c) shows the spectra from the eight-channel recording of configuration QP1 (for details, see Sec. IV A).

channels record different combinations of the three oscillators.

We perform an HCD with  $M = 100$ . With  $M \approx 100$  (or larger) we clearly identify three separate distinct peaks of approximately equal size in the HCD power spectra (Fig. 7). These peaks correspond to the dominant frequencies of the three oscillators. From the peaks the spectrum decays in a similar manner to that for the traveling wave (compare Fig. 2), such that for the three oscillators with three distinct frequencies, we find all three eigenvalue pairs to be greater than 0 for all  $\omega_l > 0$ . HCD power and phase spectra from the three-channel recordings of QP1 and QP3 exhibit the same qualitative shape. The periodic behavior of all three oscillators in QP1 does not lead to a simpler HCD spectrum compared to QP3 as the nonrandom phase relations in QP1 occur between signal components at distinct frequencies.

The HCD spectra of the eight-channel observation of QP1 is shown in Fig. 7(c). The HCD power spectrum of the eight-channel observation resembles the spectrum of the three-channel observation in the sense that again we find three peaks near the dominant frequencies of the three oscillators and a smooth decay from the peaks similar to the traveling wave spectra. At  $\omega_l = 0$  we observe six nonzero eigenvalues, and for  $\omega_l > 0$  six nonzero pairs of eigenvalues at each frequency. For the eight-channel observation this means that for  $\omega_l > 0$  two eigenvalue pairs at each frequency are numerically zero. Three pairs are expected as we are observing signals from three oscillators with distinct frequencies. The additional three smaller nonzero pairs result, as in the case of the traveling wave, from the discretization. For an eight-channel observation without phase lags, e.g., by setting  $Q_{pk} = 0$  for all  $p, k$ , the HCD power spectrum reduces to three nonzero eigenvalue pairs for  $\omega_l > 0$  (not shown). The HCD power spectrum may thus be viewed as foremost characterizing the underlying

oscillatory system rather than a particular observation (i.e., linear combination of the oscillator series). The heights of the power spectral peaks are, however, determined by the weights of the different oscillators in  $P_{pk}$  and  $Q_{pk}$ . The HCD spectrum is thus not unaffected by the choices made when observing the system.

The mixing of the oscillator time series in the channel observations becomes apparent in the HCD modes. Figure 8 shows the leading HCD modes from QP1 at four different frequencies. We first take a closer look at the mode snippets corresponding to channels 5 and 6. Channels 5 and 6 only observe oscillator number 3, i.e., the oscillator with the highest natural frequency ( $\omega_3 = 9.11$ ). Accordingly, channels 5 and 6 show up in the leading mode at the third power spectral peak [Fig. 8(d)]. In the leading modes at the first and second peak [Figs. 8(b) and 8(c)], the snippets corresponding to channels 5 and 6 exhibit an amplitude close to 0. As the observations at channels 5 and 6 are approximately in phase quadrature, the corresponding mode snippets are shifted by  $\pi/2$  accordingly. Channel 8 observes signals from all three oscillators, with particularly strong contributions from oscillators 1 and 2. Thus the channel 8 snippet has a comparably large amplitude in the leading modes at the first and second peak [Figs. 8(b) and 8(c)] and visibly smaller amplitude in the leading modes at the third peak [Fig. 8(d)]. The mode shapes are thus indicative of the presence (or absence) of an oscillator in the observations at a particular channel.

## B. EEG and surrogate EEG data

As a real-world example, we apply HCD and DAHD to human EEG recordings. We use a 20-s epoch of a 64-channel resting-state scalp EEG recording from a healthy subject. The data have been sampled at 1 kHz and preprocessed, including artifact removal and bandpass filtering to 1–200 Hz,

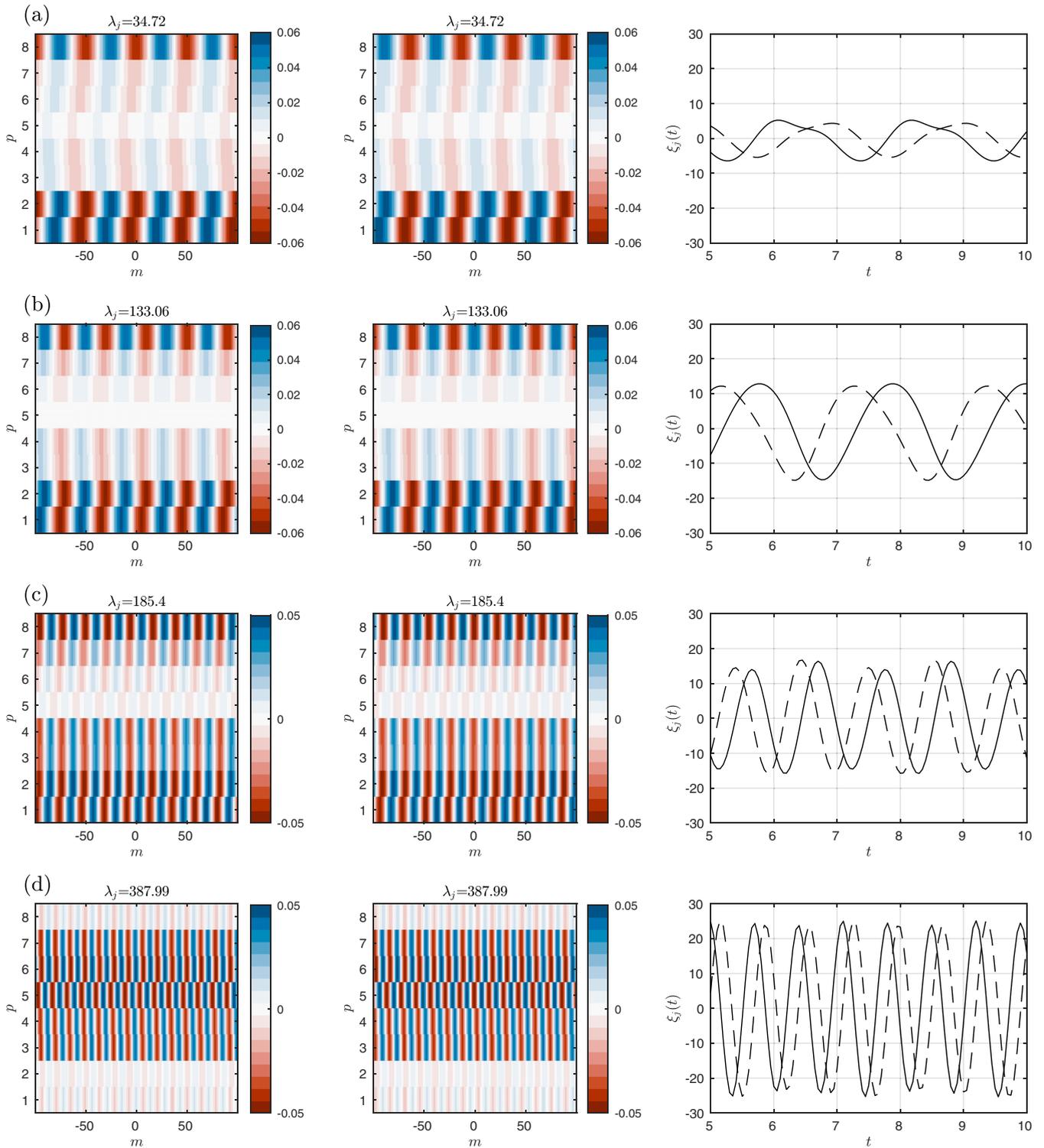


FIG. 8. HCD of the eight-channel recording of the three-oscillator system in configuration QP1 (for details, see Sec. IV A). Shown is the leading mode pair (i.e., the pair of modes associated with the largest absolute eigenvalues  $|\lambda|$ ) and the corresponding coefficient time series  $\xi_j(t)$  (right) at four different frequencies: (a)  $\omega_5 = 2.54$ , (b)  $\omega_6 = 3.17$ , (c)  $\omega_{10} = 5.71$ , and (d)  $\omega_{15} = 8.89$ . The frequencies shown in panels (b)–(d) correspond to the peaks in the HCD power spectrum shown in Fig. 7(c).

by [53]. Before we apply the spectral decompositions, we down-sample the data to 0.1 kHz. The autocorrelations of the down-sampled EEG data are shown in Fig. 9. The autocorrelations decay quickly and become flat at about 0.4 s, therefore we choose  $M = 40$ .

For comparison, we generated surrogate data following [54]. In the surrogates, Fourier power spectral density and amplitude distribution of each channel series resemble those of the actual EEG recording, but the phases are drawn randomly from a uniform distribution on  $[0, 2\pi)$ . We thus expect

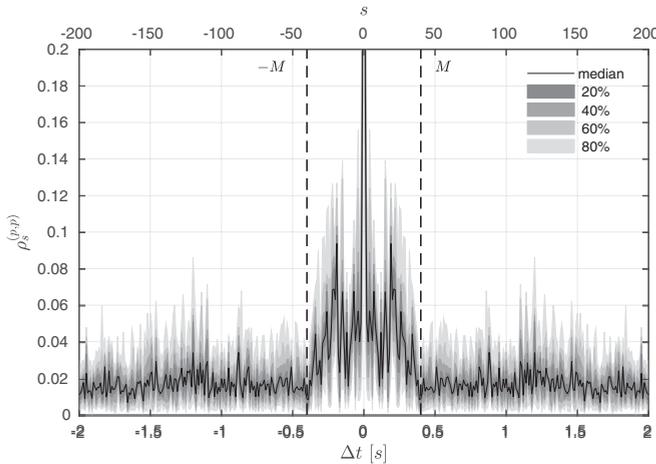


FIG. 9. Autocorrelations of the 64-channel EEG recording of a single subject (for details, see Sec. IV B). The shaded areas indicate the distribution, i.e., 20% of the plotted autocorrelations lie within the area with the darkest shading, and so on.

the phase spectra from the surrogate data to be random and to not exhibit any distinct structure. The correlations in the actual EEG recordings are a mixture of correlations in brain electrical activity, common signals being recorded at multiple channels and random correlations that arise due to finite window length. In the surrogate data, solely the random correlations, attributed to the finite length and frequency content of the series [55], remain. The correlations in the surrogate data are thus expected to be (on average) smaller than for the actual EEG data for a given window size.

With a sampling rate of 0.1 kHz and  $M = 40$ , the frequencies of the HCD modes range from  $f_1 = 0$  Hz to  $f_{40} = 50$  Hz with  $\Delta f \approx 1.28$  Hz. Figure 10 shows the leading HCD modes from the EEG data at 0, 2.56, 5.12 and 8.97 Hz. The phase differences between the snippets appear to be concentrated around 0 and  $\pi$ . In other words, the snippets are close to being in phase or antiphase. The structure of the leading modes thus appears to be dominated by nonrandom correlations, i.e., actual correlations of electrical brain activity and/or signals from common sources recorded at distinct channels.

The HCD power and phase spectra are shown in Fig. 11. The black line corresponds to the average over the eigenvalues at each frequency, i.e.,  $\Lambda_l = \sum_{j \in \mathcal{I}(f_l)} |\lambda_j| / 2d$ . As expected, the averages are of the same shape in both EEG and surrogate spectra as the surrogates are constructed to have the same Fourier power spectral density as the actual EEG. In the  $y$ -direction, EEG and surrogate spectra differ. The surrogate spectrum is narrower than the EEG spectrum.

For the zero-lag correlation matrix, the eigenvalue spectrum has been extensively studied. The eigenvalue spectrum provides information on the presence of genuine, i.e., nonrandom, cross-correlation in multivariate time-series data (see, e.g., Refs. [56,57]). The parts of the spectrum dominated by random correlations have been found to match predictions from random matrix theory [58], while genuine correlations have been found to imprint on the spectrum in the form of repulsions between particular eigenvalues [59–61]. Typically, the repulsion of eigenvalues leads to a broadening of the spectrum. Hence, the finding of a wider HCD spectrum (in

the  $y$ -direction) for the actual EEG compared to surrogates appears to be in line with [59–61].

Next we examine the HCD phase spectra (Fig. 11). The (leading mode) phase spectrum from the actual EEG has the majority of phases organized into short (vertical) lines indicating the phases of the mode snippets to be concentrated around particular phases at each frequency. Looking closely, we see that the snippets are close to being either in phase or antiphase as we have already observed in the example modes in Fig. 10. As expected, the HCD phase spectrum from the surrogate data does not exhibit any visible structure.

The DAHD phase spectrum from the EEG data shares some of the features of the HCD spectrum. In both cases, the snippet phases at each frequency are concentrated around some  $\theta$ . The DAHD phase spectrum, however, further exhibits a linelike structure in the  $x$ -direction, which is even more obvious for the modified DAHD with the alternate grand matrix symmetrization from Eq. (18), which in the following we abbreviate as DAHD\*. For the DAHD\* the phase spectrum collapses to these linelike patterns, i.e., the mode snippets are exactly in phase or antiphase, which is expected as our alternate grand matrix symmetrization is analogous to symmetrizing the cross-correlations. This affects EEG and surrogate spectra alike. Surprisingly, however, the phase spectrum for the surrogates obtained from the original DAHD exhibits a nonrandom structure as well. The surrogate spectrum appears to contain a stronger random component than the DAHD spectrum from the actual EEG, but some of the structure remains for the surrogate data. This is counterintuitive as the surrogate data have been constructed by randomly drawing the phases from a uniform distribution.

Finally, Fig. 12 compares the leading modes from HCD, DAHD, and M-SSA. This example illustrates how DAHD and HCD differ from M-SSA. DAHD and HCD modes exhibit unique frequencies, in the sense that the mode snippets are pure sines with a single frequency. The oscillatory DAHD and HCD modes come as distinct pairs in phase quadrature and associated with eigenvalues of the same absolute value and opposite sign for DAHD and the same value and same sign for HCD. The oscillatory pairs from M-SSA are also approximately in phase quadrature, but associated only with similar eigenvalues, making the identification of the oscillatory pairs less straightforward than in DAHD and HCD. Further, a mixing of frequencies can occur in the M-SSA modes. In Fig. 12(e) this is most obvious in the two panels on the right. A structured varimax rotation as proposed by [62], however, allows us to visibly reduce the mixing of frequencies within the M-SSA modes [Fig. 12(f)].

Unlike for the traveling wave in Fig. 1, the leading modes from HCD and DAHD look similar for the EEG data. The difference between HCD and DAHD modes for a given data set appears to be related to the asymmetry of cross-correlations. The stronger the asymmetries in cross-correlations, the stronger the effect that the DAHD grand matrix symmetrization has on the shape of the DAHD modes. For the EEG recordings, the cross-correlations exhibit less asymmetry than in the case of the traveling wave. In other words, the dominant oscillatory components in the EEG recordings that are shared by distinct channels seem to appear at zero or near-zero lag, as seen in the phase spectra.

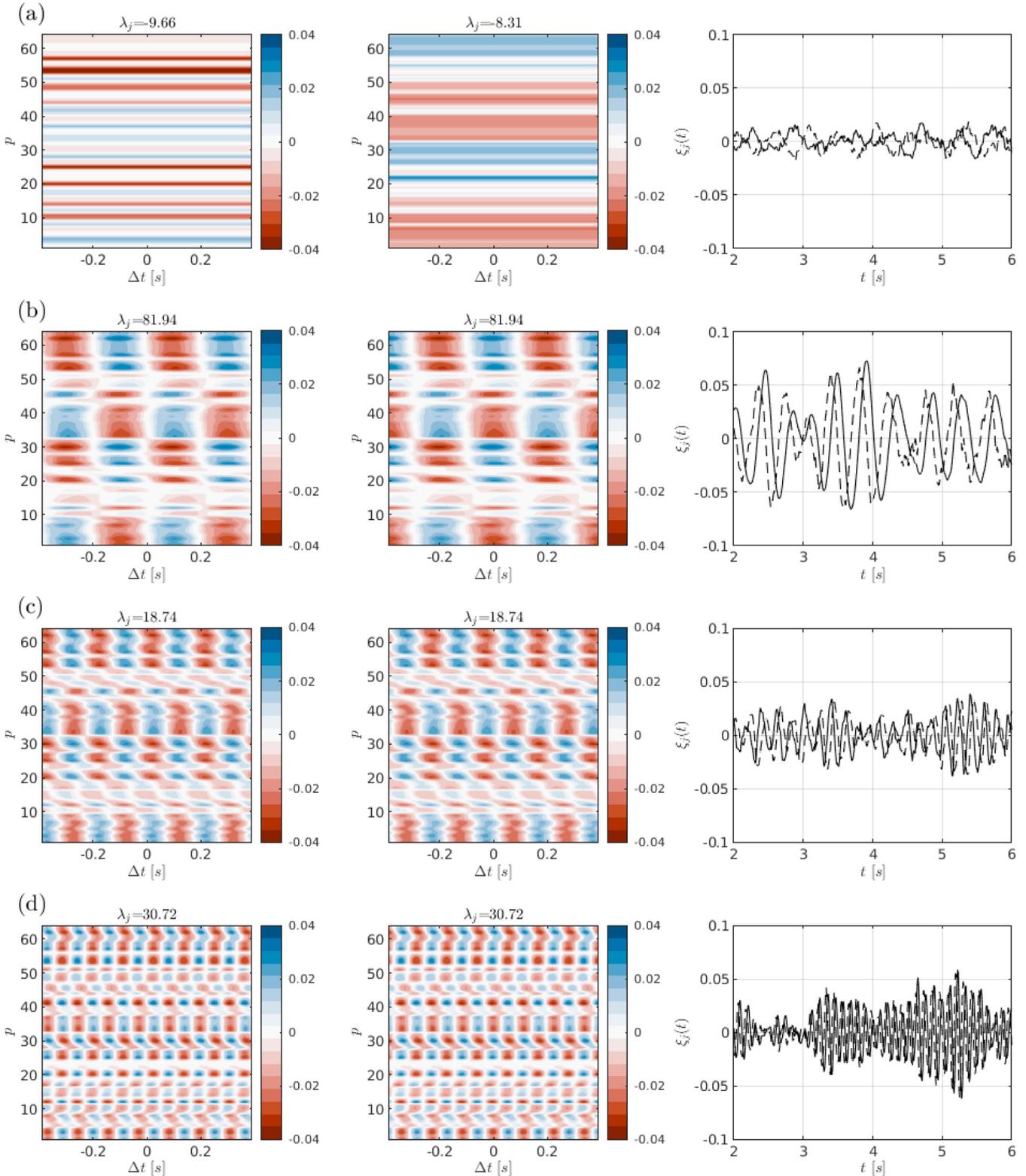


FIG. 10. HCD of a 64-channel EEG recording. Shown is the leading mode pair (i.e., the pair of modes associated with the largest eigenvalues  $\lambda$ ) and the corresponding coefficient time series  $\xi_j(t)$  (right) at four different frequencies: (a)  $f_1 = 0$  Hz, (b)  $f_3 = 2.56$  Hz, (c)  $f_5 = 5.12$  Hz, and (d)  $f_8 = 8.97$  Hz.

**V. DISCUSSION**

A comparison of the main features of HCD, DAHD, and M-SSA is summarized in Table II. Central to all three spec-

tral decomposition techniques is an eigendecomposition of a grand matrix containing lagged cross-correlations (or covariances). The differences between the decompositions arise from the different ways in which the cross-correlations are

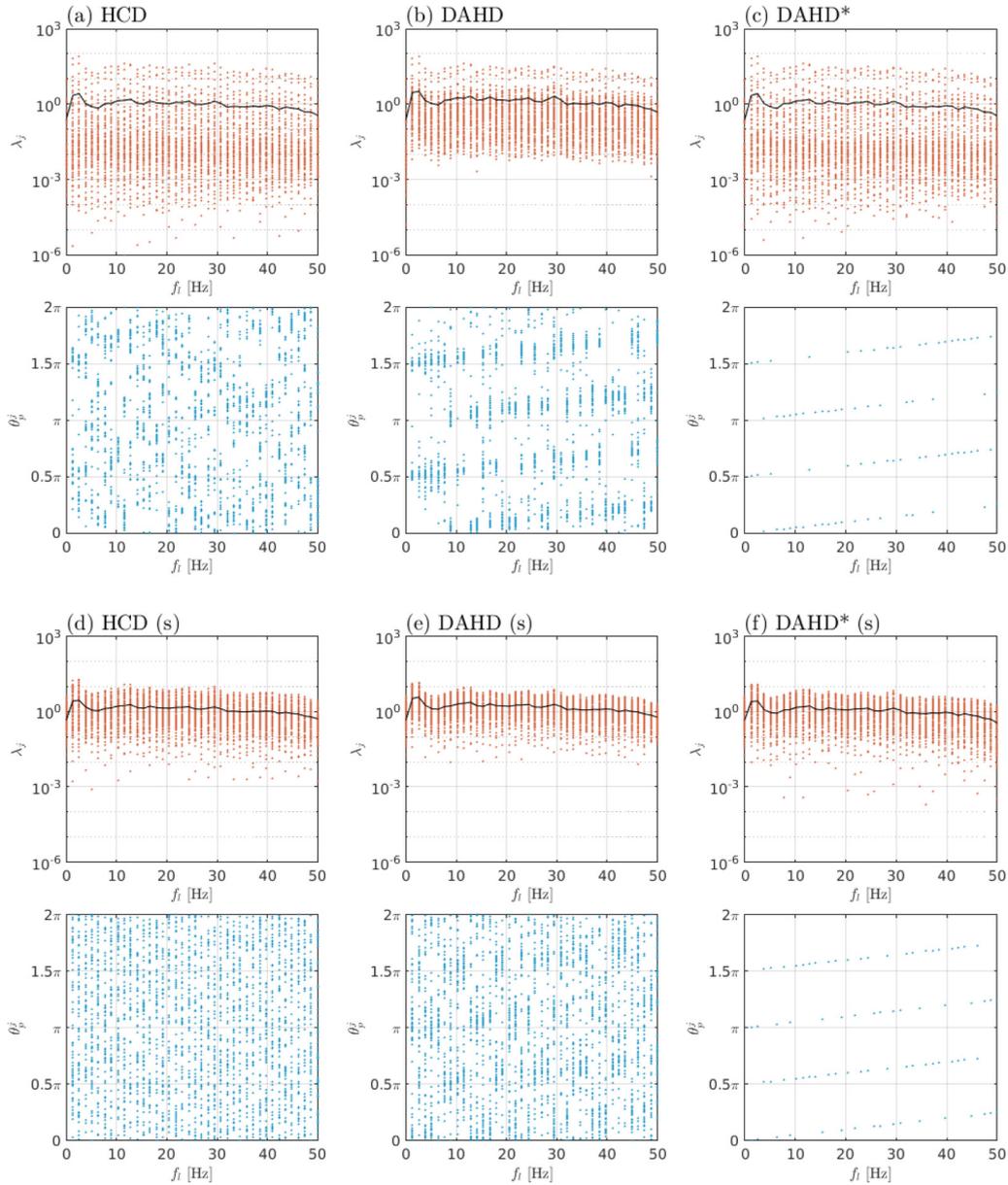


FIG. 11. Multidimensional power spectra (top) and leading mode phase spectra (bottom) from the spectral decomposition of the 64-channel EEG recording (a)–(c) and surrogate EEG (d)–(f): (a), (d) show the HCD spectra from the EEG recording, (b), (e) show the DAHD spectra, and (c), (f) show the spectra from the modified version of the DAHD method where the grand matrix is symmetrized by Eq. (18). The back line in the upper panels shows the average over all eigenvalues at the respective frequency.

arranged within the matrix. In M-SSA, the grand matrix is either obtained as the correlation matrix of the lag-embedded multivariate series or as a block matrix of Toeplitz blocks. The DAHD grand matrix is built from left-circulant blocks (i.e., circulant Hankel matrices). The HCD grand matrix is built from right-circulant blocks (i.e., circulant Toeplitz matrices).

The eigenvectors of the grand matrices can be rearranged into spatiotemporal modes, and for time series containing some sort of oscillatory behavior a subset of the modes exhibits an oscillatory pattern along the temporal dimension. The oscillatory modes come in pairs that can be identified via the corresponding eigenvalues. In the HCD method, the eigenvalues of an oscillatory pair have the same value and

sign. In the DAHD method, the eigenvalues of an oscillatory pair have the same value but opposite signs. In M-SSA, the mode pairing is only approximate. The modes associated with an eigenvalue pair exhibit the same (for HCD and DAHD) or a similar (for M-SSA) spatiotemporal pattern solely shifted by  $\pi/2$  in the temporal direction.

Along the temporal dimension, HCD and DAHD modes are pure sinusoids with the Fourier frequencies on the embedding interval. The data-adaptive nature of HCD and DAHD manifests in the amplitudes and phases of the sinusoids, while in M-SSA amplitudes, phases and frequencies are data-adaptive. In that sense, one may view M-SSA as the most data-adaptive. The advantage of HCD and DAHD is that each mode can be uniquely assigned to a particular frequency,

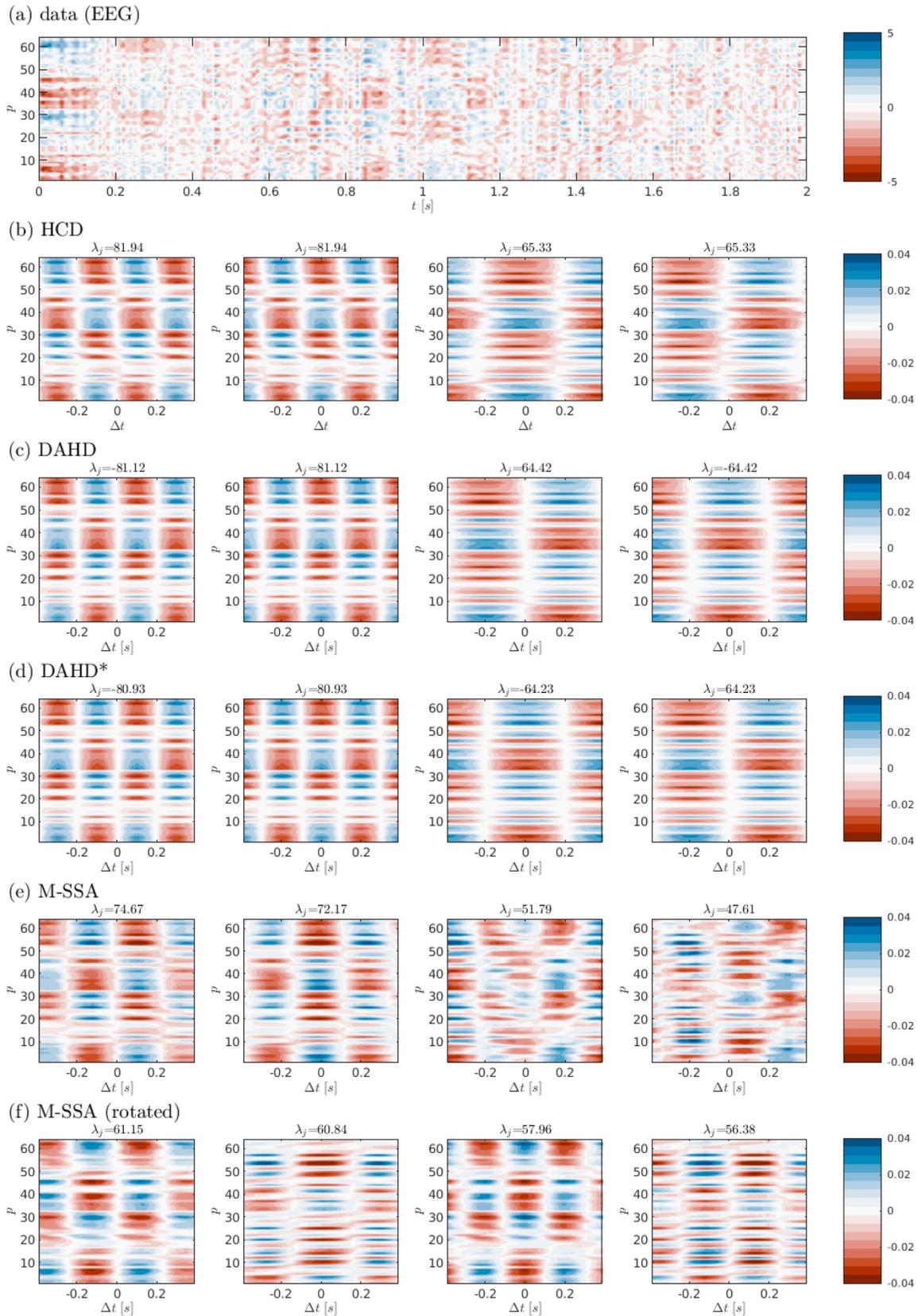


FIG. 12. Excerpt of a 64-channel EEG recording (a) and the leading modes of its spectral decomposition. Shown are the two leading mode pairs across all frequencies (i.e., the pair of modes associated with the largest absolute eigenvalues  $|\lambda|$ ) from the different decomposition techniques: (b) HCD, (c) DAHD, (d) DAHD\*, (e) M-SSA, and (f) varimax rotated M-SSA. DAHD\* refers to a modified version of the DAHD method where the grand matrix is symmetrized by Eq. (18). For panel (f) the structured varimax rotation of [62] has been applied to the leading 50 M-SSA modes.

TABLE II. Comparison of the main characteristics of the lagged cross-correlation-based spectral decomposition techniques: multivariate singular spectrum analysis (M-SSA [50]), data-adaptive harmonic decomposition (DAHD [43]), and harmonic cross-correlation decomposition (HCD [this paper]).

	M-SSA	DAHD	HCD
grand matrix	correlation matrix or Toeplitz blocks	left-circulant (Hankel) blocks	(right-)circulant (Toeplitz) blocks
oscillatory modes	data adaptive (mixtures possible <sup>a</sup> )	sinusoids	sinusoids
mode frequencies	data adaptive (mixtures possible <sup>a</sup> )	Fourier frequencies	Fourier frequencies
mode pairing	approximate (similar in eigenvalue)	exact (opposite sign)	exact (same sign)
multivariate power and phase spectra	no	yes	yes
data adaptive	phase, amplitude and frequency/shape	phase and amplitude	phase and amplitude
invariance under channel permutation	yes	no	yes

<sup>a</sup>Mixtures can be reduced by a structured varimax rotation of the M-SSA modes [62].

which, along with their particular ordering ( $d$  mode pairs at each Fourier frequency and  $d$  nonoscillatory modes at  $f = 0$ ), gives rise to multivariate power and phase spectra. This power and phase spectra can be useful tools for exploratory analysis of multivariate time-series data.

The DAHD spectra, however, do not uniquely characterize correlation and phase relations in multivariate time-series data as the DAHD spectra are not invariant under channel permutation. That is, a reordering of the channels within the DAHD grand matrix will typically give different DAHD modes and spectra. This is a side effect of the explicit symmetrization required to ensure that the grand matrix from Hankel blocks is symmetric. When constructing a grand matrix from Toeplitz blocks, no explicit symmetrization is needed. The HCD is accordingly invariant under channel permutation. HCD power and phase spectra uniquely characterize correlation and phase relations in multivariate time-series data in an intuitive way.

The eigenvalue spectrum of the zero-lag correlation matrix has gained a lot of attention in the literature, for instance in the analysis of EEG data, as it can provide information on correlation strengths and correlation structure within multivariate time-series data [56,57,61,63–66]. The HCD power spectra may thus offer a tool to obtain frequency-specific insight into correlation structure and strengths not restricted to correlations at zero-lag and without filtering to a particular frequency band beforehand.

The HCD phase spectra depict predominant phase relations between time series sampled at distinct channels, and they can thereby be useful in identifying how oscillatory signals propagate through a system. The amplitudes of the mode snippets associated with the channels is linked to the magnitude of the respective oscillatory pattern at the respective channels. Especially in cases in which the relative amplitudes of the channel time series are considered important, as might be the case for measurement data, it may be more suitable to perform the HCD based on covariances rather than correlations, or in other words to only centralize the data to zero mean, but not to normalize it to unit variance before applying Eq. (1). A varimax rotation as commonly applied to PCA to simplify the leading EOFs by enhancing sparseness may further allow us to simplify and improve the interpretability of the HCD modes. The varimax rotation used in PCA is not directly transferable to spatiotemporal modes. However, Ref. [62] introduced a modified varimax rotation (for M-SSA) that increases sparsity

along the spatial dimension of the leading  $n$  modes while preserving the oscillation along the temporal dimension.

As outlined in the Introduction, Ref. [49] recently introduced a Hermitian frequency domain formulation of DAHD. The Hermitian DAHD formulation systematically differs from the original DAHD method in that it does not include an explicit symmetrization step. See the Supplemental Material [52] for an intercomparison of the frequency domain formulation of the original DAHD (equivalent to time domain DAHD) and the Hermitian DAHD formulation (equivalent to HCD) by means of an example data set studied in [49]. Further recent studies have introduced spectral empirical orthogonal function (SEOF) analysis [11] and spectral proper orthogonal decomposition (SPOD) [67] as spectral decomposition techniques for multivariate time series. Both decompositions are derived in the frequency domain. An intercomparison of the above techniques with DAHD and HCD is a future challenge.

DAHD has been developed as part of an inverse modeling framework in which the paired oscillatory DAHD coefficient series are identified with the real and imaginary parts of Stuart-Landau oscillators [43]. Real-world applications of that modeling framework include stochastic modeling of Arctic sea ice extent [46], solar wind-magnetosphere coupling [44], and ocean gyres [45]. Future work may address a similar modeling approach based on the HCD method or the equivalent Hermitian DAHD formulation of [49]. Especially when aiming to obtain insights into the system dynamics from data-driven modeling, the independence of the channel ordering may be advantageous. The more efficient representation of the data in the HCD mode space that we have observed for the traveling wave (see also the Rossby waves example discussed in the Supplemental Material [52]) may further allow one to reduce the dimension, i.e., to use fewer oscillators, in modeling. It should, however, be noted that the linear relationship between eigenvalues and explained variance from M-SSA appears not to be directly transferable to the HCD method.

#### ACKNOWLEDGMENTS

We would like to thank George Stothart, Nina Kazanina, and Luke Tait for providing the EEG used in this study. We gratefully acknowledge the financial support of the EPSRC Centre for Predictive Modeling in Healthcare, via Grant No. EP/N014391/1.

- [1] G. Buzsáki and B. O. Watson, Brain rhythms and neural syntax: Implications for efficient coding of cognitive content and neuropsychiatric disease, *Dialogues Clin. Neurosci.* **14**, 345 (2012).
- [2] E. Başar, C. Schmiedt-Fehr, B. Mathes, B. Femir, D. D. Emek-Savaş, E. Tülay, D. Tan, A. Düzgün, B. Güntekin, A. Özerdem *et al.*, What does the broken brain say to the neuroscientist? Oscillations and connectivity in schizophrenia, Alzheimer's disease, and bipolar disorder, *Int. J. Psychophysiol.* **103**, 135 (2016).
- [3] H. Schmidt, W. Woldman, M. Goodfellow, F. A. Chowdhury, M. Koutroumanidis, S. Jewell, M. P. Richardson, and J. R. Terry, A computational biomarker of idiopathic generalized epilepsy from resting state EEG, *Epilepsia* **57**, e200 (2016).
- [4] R. Cassani, M. Estarellas, R. San-Martin, F. J. Fraga, and T. H. Falk, Systematic review on resting-state EEG for Alzheimer's disease diagnosis and progression assessment, *Disease Markers* **2018**, 5174815 (2018).
- [5] V. J. Geraedts, L. I. Boon, J. Marinus, A. A. Gouw, J. J. van Hilten, C. J. Stam, M. R. Tannemaat, and M. F. Contarino, Clinical correlates of quantitative EEG in Parkinson disease: A systematic review, *Neurology* **91**, 871 (2018).
- [6] L. Tait, F. Tamagnini, G. Stothart, E. Barvas, C. Monaldini, R. Frusciante, M. Volpini, S. Guttmann, E. Coulthard, J. T. Brown *et al.*, EEG microstate complexity for aiding early diagnosis of Alzheimer's disease, *Sci. Rep.* **10**, 17627 (2020).
- [7] E. M. Rasmusson and J. M. Wallace, Meteorological aspects of the El Niño/Southern Oscillation, *Science* **222**, 1195 (1983).
- [8] J. W. Hurrell, Y. Kushnir, G. Otttersen, and M. Visbeck, An overview of the North Atlantic oscillation, *Geophys. Monograph-Am. Geophys. Union* **134**, 1 (2003).
- [9] S. Brönnimann, Impact of El Niño–Southern Oscillation on European climate, *Rev. Geophys.* **45**, RG3003 (2007).
- [10] A. Pepelyshev and A. Zhigljavsky, SSA analysis and forecasting of records for Earth temperature and ice extents, *Stat. Its Interface* **10**, 151 (2017).
- [11] O. T. Schmidt, G. Mengaldo, G. Balsamo, and N. P. Wedi, Spectral empirical orthogonal function analysis of weather and climate data, *Mon. Weather Rev.* **147**, 2979 (2019).
- [12] F. Takens, Detecting strange attractors in turbulence, in *Dynamical Systems and Turbulence, Warwick, 1980* (Springer, Berlin, Heidelberg, 1981), pp. 366–381.
- [13] D. S. Broomhead and G. P. King, Extracting qualitative dynamics from experimental data, *Physica D* **20**, 217 (1986).
- [14] K. Fraedrich, Estimating the dimensions of weather and climate attractors, *J. Atmos. Sci.* **43**, 419 (1986).
- [15] J. B. Elsner and A. A. Tsonis, *Singular Spectrum Analysis: A New Tool in Time Series Analysis* (Plenum, New York, NY, 2013).
- [16] N. Golyandina and A. Zhigljavsky, *Singular Spectrum Analysis for Time Series* (Springer Science & Business Media, Berlin, 2013).
- [17] N. Golyandina and A. Korobeynikov, Basic singular spectrum analysis and forecasting with R, *Comput. Stat. Data Anal.* **71**, 934 (2014).
- [18] H. Hassani, S. Heravi, and A. Zhigljavsky, Forecasting European industrial production with singular spectrum analysis, *Int. J. Forecasting* **25**, 103 (2009).
- [19] A. Pepelyshev and A. Zhigljavsky, Assessing the stability of long-horizon SSA forecasting, *Statistics Its Interface* **3**, 321 (2010).
- [20] F. Alonso, J. Del Castillo, and P. Pintado, Application of singular spectrum analysis to the smoothing of raw kinematic signals, *J. Biomech.* **38**, 1085 (2005).
- [21] T. Harris and H. Yuan, Filtering and frequency interpretations of singular spectrum analysis, *Physica D* **239**, 1958 (2010).
- [22] D. Kondrashov and M. Ghil, Spatio-temporal filling of missing points in geophysical data sets, *Nonlin. Proc. Geophys.* **13**, 151 (2006).
- [23] D. Kondrashov, Y. Shprits, and M. Ghil, Gap filling of solar wind data by singular spectrum analysis, *Geophys. Res. Lett.* **37**, L15101 (2010).
- [24] R. Rekapalli, R. Tiwari, M. K. Sen, and N. Vedanti, 3D seismic data de-noising and reconstruction using multichannel time slice singular spectrum analysis, *J. Appl. Geophys.* **140**, 145 (2017).
- [25] H. R. Ghafarian Malamiri, I. Roustaa, H. Olafsson, H. Zare, and H. Zhang, Gap-filling of MODIS time series land surface temperature (LST) products using singular spectrum analysis (SSA), *Atmosphere* **9**, 334 (2018).
- [26] S. Aydın, H. M. Saraoğlu, and S. Kara, Singular spectrum analysis of sleep EEG in insomnia, *J. Med. Syst.* **35**, 457 (2011).
- [27] S. Kouchaki, S. Sanei, E. L. Arbon, and D.-J. Dijk, Tensor based singular spectrum analysis for automatic scoring of sleep EEG, *IEEE Trans. Neural Syst. Rehabilitation Eng.* **23**, 1 (2014).
- [28] S. M. Mohammadi, S. Enshaeifar, M. Ghavami, and S. Sanei, Classification of awake, REM, and NREM from EEG via singular spectrum analysis, in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE, Piscataway, NJ, 2015), pp. 4769–4772.
- [29] S. M. Mohammadi, S. Kouchaki, M. Ghavami, and S. Sanei, Improving time–frequency domain sleep EEG classification via singular spectrum analysis, *J. Neurosci. Methods* **273**, 96 (2016).
- [30] S. Sanei, M. Ghodsi, and H. Hassani, An adaptive singular spectrum analysis approach to murmur detection from heart sounds, *Med. Eng. Phys.* **33**, 362 (2011).
- [31] J. Harmouche, D. Fourer, F. Auger, P. Borgnat, and P. Flandrin, The sliding singular spectrum analysis: A data-driven nonstationary signal decomposition tool, *IEEE Trans. Signal Process.* **66**, 251 (2017).
- [32] P. Celka and P. Colditz, A computer-aided detection of EEG seizures in infants: A singular-spectrum approach and performance comparison, *IEEE Trans. Biomed. Eng.* **49**, 455 (2002).
- [33] T. Siddharth, R. K. Tripathy, and R. B. Pachori, Discrimination of focal and non-focal seizures from EEG signals using sliding mode singular spectrum analysis, *IEEE Sens. J.* **19**, 12286 (2019).
- [34] B. J. Weder, K. Schindler, T. J. Loher, R. Wiest, M. Wissmeyer, P. Ritter, K. Lovblad, F. Donati, and J. Missimer, Brain areas involved in medial temporal lobe seizures: a principal component analysis of ictal SPECT data, *Human Brain Mapping* **27**, 520 (2006).
- [35] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection, *IEEE Trans. Biomed. Eng.* **55**, 512 (2008).
- [36] F. Artoni, A. Delorme, and S. Makeig, Applying dimension reduction to EEG data by principal component analysis reduces

- the quality of its subsequent independent component decomposition, *NeuroImage* **175**, 176 (2018).
- [37] B. C. Weare and R. Newell, Empirical orthogonal analysis of Atlantic Ocean surface temperatures, *Q. J. R. Meteorol. Soc.* **103**, 467 (1977).
- [38] A. H. Monahan and A. Dai, The spatial and temporal structure of ENSO nonlinearity, *J. Clim.* **17**, 3026 (2004).
- [39] A. Hannachi, I. Jolliffe, and D. Stephenson, Empirical orthogonal functions and related techniques in atmospheric science: A review, *Int. J. Climatol.* **27**, 1119 (2007).
- [40] M. Ghil, M. Allen, M. Dettinger, K. Ide, D. Kondrashov, M. Mann, A. W. Robertson, A. Saunders, Y. Tian, F. Varadi *et al.*, Advanced spectral methods for climatic time series, *Rev. Geophys.* **40**, 3 (2002).
- [41] N. Golyandina, Particularities and commonalities of singular spectrum analysis as a method of time series analysis and signal processing, *Wiley Interdisciplinary Rev.: Comput. Stat.* **12**, e1487 (2020).
- [42] B. C. Weare and J. S. Nasstrom, Examples of extended empirical orthogonal function analyses, *Mon. Weather Rev.* **110**, 481 (1982).
- [43] M. D. Chekroun and D. Kondrashov, Data-adaptive harmonic spectra and multilayer Stuart-Landau models, *Chaos* **27**, 093110 (2017).
- [44] D. Kondrashov and M. D. Chekroun, Data-adaptive harmonic analysis and modeling of solar wind-magnetosphere coupling, *J. Atmos. Sol. Terr. Phys.* **177**, 179 (2018).
- [45] D. Kondrashov, M. Chekroun, and P. Berloff, Multiscale Stuart-Landau emulators: Application to wind-driven ocean gyres, *Fluids* **3**, 21 (2018).
- [46] D. Kondrashov, M. D. Chekroun, and M. Ghil, Data-adaptive harmonic decomposition and prediction of Arctic sea ice extent, *Dyn. Stat. Climate Syst.* **3**, dzy001 (2018).
- [47] E. Ryzhov, D. Kondrashov, N. Agarwal, and P. Berloff, On data-driven augmentation of low-resolution ocean model dynamics, *Ocean Model.* **142**, 101464 (2019).
- [48] E. Ryzhov, D. Kondrashov, N. Agarwal, J. McWilliams, and P. Berloff, On data-driven induction of the low-frequency variability in a coarse-resolution ocean model, *Ocean Model.* **153**, 101664 (2020).
- [49] D. Kondrashov, E. Ryzhov, and P. Berloff, Data-adaptive harmonic analysis of oceanic waves and turbulent flows, *Chaos* **30**, 061105 (2020).
- [50] R. Vautard, P. Yiou, and M. Ghil, Singular-spectrum analysis: A toolkit for short, noisy chaotic signals, *Physica D* **58**, 95 (1992).
- [51] C. Baesens, J. Guckenheimer, S. Kim, and R. S. MacKay, Three coupled oscillators: Mode-locking, global bifurcations and toroidal chaos, *Physica D* **49**, 387 (1991).
- [52] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevE.103.062213> for an intercomparison of the frequency domain formulation of the original DAHD method and the Hermitian DAHD formulation.
- [53] L. Tait, G. Stothart, E. Coulthard, J. T. Brown, N. Kazanina, and M. Goodfellow, Network substrates of cognitive impairment in Alzheimer's disease, *Clin. Neurophys.* **130**, 1581 (2019).
- [54] T. Schreiber and A. Schmitz, Improved Surrogate Data for Nonlinearity Tests, *Phys. Rev. Lett.* **77**, 635 (1996).
- [55] G. U. Yule, Why do we sometimes get nonsense-correlations between time-series?—a study in sampling and the nature of time-series, *J. R. Stat. Soc.* **89**, 1 (1926).
- [56] M. Müller, G. Baier, C. Rummel, and K. Schindler, Estimating the strength of genuine and random correlations in non-stationary multivariate time series, *Europhys. Lett.* **84**, 10009 (2008).
- [57] C. Rummel, M. Müller, G. Baier, F. Amor, and K. Schindler, Analyzing spatio-temporal patterns of genuine cross-correlations, *J. Neurosci. Methods* **191**, 94 (2010).
- [58] L. Laloux, P. Cizeau, M. Potters, and J.-P. Bouchaud, Random matrix theory and financial correlations, *Int. J. Theor. Appl. Finance* **3**, 391 (2000).
- [59] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, T. Guhr, and H. E. Stanley, Random matrix approach to cross correlations in financial data, *Phys. Rev. E* **65**, 066126 (2002).
- [60] A. Utsugi, K. Ino, and M. Oshikawa, Random matrix theory analysis of cross correlations in financial markets, *Phys. Rev. E* **70**, 026110 (2004).
- [61] M. Müller, G. Baier, A. Galka, U. Stephani, and H. Muhle, Detection and characterization of changes of the correlation structure in multivariate time series, *Phys. Rev. E* **71**, 046116 (2005).
- [62] A. Groth and M. Ghil, Multivariate singular spectrum analysis and the road to phase synchronization, *Phys. Rev. E* **84**, 036206 (2011).
- [63] M. Müller, K. Wegner, U. Kummer, and G. Baier, Quantification of cross correlations in complex spatiotemporal systems, *Phys. Rev. E* **73**, 046106 (2006).
- [64] C. Rummel, G. Baier, and M. Müller, Automated detection of time-dependent cross-correlation clusters in nonstationary time series, *Europhys. Lett.* **80**, 68004 (2007).
- [65] K. Schindler, H. Leung, C. E. Elger, and K. Lehnertz, Assessing seizure dynamics by analysing the correlation structure of multichannel intracranial EEG, *Brain* **130**, 65 (2007).
- [66] C. Rummel, M. Goodfellow, H. Gast, M. Hauf, F. Amor, A. Stibal, L. Mariani, R. Wiest, and K. Schindler, A systems-level approach to human epileptic seizures, *Neuroinformatics* **11**, 159 (2013).
- [67] A. Towne, O. T. Schmidt, and T. Colonius, Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis, *J. Fluid Mech.* **847**, 821 (2018).
- [68] H. Abdi and L. J. Williams, Principal component analysis, *Wiley Interdisciplinary Rev.: Comput. Stat.* **2**, 433 (2010).
- [69] D. S. Wilks, *Statistical Methods in the Atmospheric Sciences* (Academic Press, Amsterdam, 2011), Vol. 100.