Rotated Spectral Principal Component Analysis (rsPCA) for Identifying Dynamical Modes of Variability in Climate Systems

CLÉMENT GUILLOTEAU,[a] ANTONIOS MAMALAKIS,[a] LAWRENCE VULIS,[a] PHONG V. V. LE,[a,b] TRYPHON T. GEORGIOU,[c] AND EFI FOUFOULA-GEORGIOU[a,d]

[a] *Department of Civil and Environmental Engineering, University of California Irvine, Irvine, California*
[b] *Faculty of Hydrology Meteorology and Oceanography, Vietnam National University, Hanoi, Vietnam*
[c] *Department of Mechanical and Aerospace Engineering, University of California Irvine, Irvine, California*
[d] *Department of Earth Systems Science, University of California Irvine, Irvine, California*

(Manuscript received 15 April 2020, in final form 28 September 2020)

ABSTRACT: Spectral PCA (sPCA), in contrast to classical PCA, offers the advantage of identifying organized spatiotemporal patterns within specific frequency bands and extracting dynamical modes. However, the unavoidable trade-off between frequency resolution and robustness of the PCs leads to high sensitivity to noise and overfitting, which limits the interpretation of the sPCA results. We propose herein a simple nonparametric implementation of sPCA using the continuous analytic Morlet wavelet as a robust estimator of the cross-spectral matrices with good frequency resolution. To improve the interpretability of the results, especially when several modes of similar amplitude exist within the same frequency band, we propose a rotation of the complex-valued eigenvectors to optimize their spatial regularity (smoothness). The developed method, called rotated spectral PCA (rsPCA), is tested on synthetic data simulating propagating waves and shows impressive performance even with high levels of noise in the data. Applied to global historical geopotential height (GPH) and sea surface temperature (SST) daily time series, the method accurately captures patterns of atmospheric Rossby waves at high frequencies (3–60-day periods) in both GPH and SST and El Niño–Southern Oscillation (ENSO) at low frequencies (2–7-yr periodicity) in SST. At high frequencies the rsPCA successfully unmixes the identified waves, revealing spatially coherent patterns with robust propagation dynamics.

KEYWORDS: Dynamics; Sea surface temperature; Pressure; Empirical orthogonal functions; Pattern detection; Spectral analysis/models/distribution

1. Introduction

Identifying spatiotemporal relations and defining coherent spatiotemporal features, typically referred to as "modes" of variability, is a proficient way to characterize complex systems. These modes, whether they are trends or cyclic patterns, are regular features whose evolution is expected to be more predictable than the general background variability and whose impact on other parts of the system is often significant. Earth's climate system is a complex system of particular interest, exhibiting a plethora of modes caused by different physical processes (e.g., solar forcing, oceanic/atmospheric circulations, land–atmosphere interactions, etc.), and imprinting themselves at various spatial and temporal scales. The accurate identification and modeling of the modes of the climate system is necessary for many key problems in geosciences, such as weather/climate prediction, attribution of extreme events and hazards, and assessment of climate change impacts.

The comprehensive knowledge of the climatic "state" of the Earth at any instant requires measured or estimated numerical values at every location of the globe for dozens of variables, at a sufficiently high spatiotemporal resolution to encompass all relevant scales of variability. For example, the NCEP–NCAR reanalysis (Kalnay et al. 1996) comprises more than 20 variables every 6 h from 1948 to the present globally with 17 pressure levels on a 2.5° × 2.5° grid (more than 1 billion entries per variable and per level). To make such large datasets interpretable and usable for diagnostic and prediction purposes, it is often necessary to find a proper dimensionality reduction scheme to reduce the number of variables of this complex system while minimizing loss of information; in other words, to reduce the system to a manageable number of dynamical modes.

A wide range of dimensionality reduction methods exists. Principal component analysis (PCA), also commonly referred to as empirical orthogonal function (EOF) analysis, is a nonparametric method widely used in climate science (Lorenz 1956; Wallace and Gutzler 1981; Jolliffe 1986; Keiner and Yan 1997; Ghil et al. 2002; Hsieh 2004; Hannachi et al. 2007; Navarra and Simoncini 2010). It consists in performing a singular value decomposition (SVD) of a multivariate system. In practice, this is generally achieved by computing the

Denotes content that is immediately available upon publication as open access.

Corresponding author: Clément Guilloteau, cguillot@uci.edu

DOI: 10.1175/JCLI-D-20-0266.1

empirical covariance matrix of the variables and extracting its eigenvectors. Its simplicity of implementation as well as the fact that this empirical method does not require any a priori assumption or model/parameter selection have contributed to its popularity and success. Nevertheless, because when it is used to identify relations across several time series it handles each time step as an independent realization of the variables, the classical PCA cannot identify dynamical relations in the data. This and other limitations as further discussed below have led to many variations or extensions of the classical PCA. For example, if the variables are given under the form of regularly sampled time series, one can compute the lagged covariances or the Fourier cross-spectra of the variables rather than computing only the (zero-lag) covariances, and then construct the empirical lagged-covariance matrices or the cross-spectral matrices of the system. This gives rise to the lagged PCA and the spectral PCA as "natural" extensions of the classical PCA.

The lagged PCA allows one to better extract dynamical modes when different variables or areas of the studied domain have delayed linear responses to the same signal with different delays. The spectral PCA (sPCA), through the phase (complex argument) information in the complex cross-spectral coefficients, also allows one to handle lagged correlations. Additionally, it offers the possibility to look for modes in specific frequency bands and is particularly potent at extracting wave-type modes and handling propagation effects (nonstationary waves). Many other methods rely on the use of the classical PCA in association with Fourier spectral analysis and frequency filtering (e.g., Kidson 1999; Power et al. 1999; Kessler 2001; Wheeler and Hendon 2004; Roundy and Schreck 2009; Chen and Wallace 2016; Chen et al. 2017; Wills et al. 2018). For the spectral PCA, the linear relations between the variables leading to the spectral PCs are defined directly in the Fourier frequency domain. While sPCA (also known as frequency domain EOF) was introduced and theorized in the early 1970s (Wallace and Dickinson 1972; Wallace 1972) and later redefined and implemented under various forms (Horel 1984; Hasselmann 1988; Johnson and McPhaden 1993; Mann and Park 1994, 1999; Ghil et al. 2002; Thornhill et al. 2002; Mann et al. 2020), it has not become a standard method in atmospheric and climate science. This may be due to the fact that, in spite of being relatively straightforward in theory, its implementation requires choosing an appropriate method for computing robust cross-spectral coefficients from finite-length time series and avoiding overfitting. Additionally, as already pointed out by Wallace (1971) and Wallace and Dickinson (1972), the interpretability of the extracted principal components (PCs) may be difficult when several dynamical modes are mixed, which is likely to be the case when several PCs of similar amplitude are found within the same spectral band.

To overcome the above difficulties, we herein propose a wavelet-based implementation of the sPCA, which relies on the complex Morlet wavelet for the estimation of the cross-spectral matrices. The continuous wavelet transform is nowadays a standard and popular tool for spectral and cross-spectral analysis (Hudgins et al. 1993; Perrier et al. 1995; Kumar and Foufoula-Georgiou 1997; Jiang and Mahadevan 2011; Banskota et al. 2017). The Morlet analytic wavelet in particular allows robust estimation of power spectra and cross-spectra in the frequency domain, even from relatively short time series, while allowing for reasonably good frequency localization (Kirby 2005; Cottis et al. 2016). In addition, we propose a rotation of the eigenvectors resulting from the wavelet-based sPCA, using the spatial regularity (smoothness) of the magnitude and phase of the rotated vectors, quantified through the L1 norm (sum of absolute values) of the spatial Laplacian, as an optimality criterion to select the "best" rotation. The combination of these two innovations (robust spectral estimation through the Morlet wavelet and spatial regularization of the rotated eigenvectors) improves the interpretability and reduces the sensitivity of the extracted modes to noise and sampling variability.

The article is organized as follows. Section 2 starts with the classical PCA, introduces the implementation of sPCA through the complex Morlet wavelet transform, and describes the methodology for the rotation of the eigenvectors (rsPCA). In section 3, the proposed methodology is tested on synthetic data, namely numerically generated nonstationary waves propagating in a 2D plane, plus a colored random noise. Section 4 presents the results of applying the rsPCA method to historical global fields of daily geopotential height (GPH) and sea surface temperature (SST) with focus on both multiannual and subannual modes of variability. Discussion and conclusions are presented in section 5.

2. Methodology

Let us consider a dataset made of L observations of N centered variables (e.g., zero-mean time series of length L associated with N spatial locations). This dataset corresponds to the $N \times L$ data matrix $\mathbf{X}$:

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,L} \\ \vdots & \ddots & \vdots \\ x_{N,1} & \cdots & x_{N,L} \end{bmatrix}. \tag{1}$$

In what follows, the notation $\mathbf{x}_n = (x_{n,1}, \ldots, x_{n,L})$, designates the nth row of the matrix, that is, the vector of observations of the nth variable (e.g., the time series at location n). All notations are defined in appendix A.

a. Classical PCA

The empirical sample covariance matrix $\tilde{\mathbf{C}}$ of the $\mathbf{X}$ dataset is an $N \times N$ matrix defined as

$$\tilde{\mathbf{C}} = \frac{1}{L-1}\mathbf{X}\mathbf{X}', \tag{2}$$

with $\mathbf{X}'$ denoting the transpose conjugate matrix of $\mathbf{X}$.

The principal component analysis is performed by extracting the eigenvectors $\mathbf{u}_n$ and associated eigenvalues λ_n^2 of the covariance matrix, which are the solutions of the system:

$$\tilde{\mathbf{C}}\mathbf{u}_n = \lambda_n^2 \mathbf{u}_n. \tag{3}$$

By definition, if $\mathbf{u}_n$ is a solution of Eq. (3), so is $\alpha\mathbf{u}_n$ for any real number α. We therefore choose to impose a unit norm (L2 Euclidian norm) for all eigenvectors. We note that any unit-norm $\mathbf{u}_n$ can always be arbitrarily multiplied by -1. The eigenvalues λ_n^2 are ranked in decreasing order such as $\lambda_n^2 \geq \lambda_{n+1}^2$. The principal component time series associated with the eigenvalue λ_n^2 and the eigenvector $\mathbf{u}_n$ is obtained as

$$\boldsymbol{\kappa}_n = \frac{\lambda_n}{N}\mathbf{X}'\mathbf{u}_n. \tag{4}$$

The rank of a system is the number of linearly independent variables composing it. The empirical rank of a system is taken as the rank of its covariance matrix, that is, the number of linearly independent columns of the $\hat{\mathbf{C}}$ matrix, which is equal to N minus the number of linearly independent eigenvectors associated with the zero eigenvalue. A system of rank $R < N$ can be reduced to R variables without loss of information. In practice, the covariance matrix derived from the data is generally found to be full rank (of rank N), simply because the relations between the variables of the system are not perfectly linear or because of measurement noise. However, PCs associated with small eigenvalues may be neglected with minimum loss of information. Therefore, the decay rate of the eigenvalues can be used as a measure of the "reducibility" of a system (i.e., how easily a high-dimensional system can be compressed into a low number of modes without important loss of information).

PCA is a powerful tool for extracting linear modes of variability of complex systems and has seen many applications in atmospheric sciences and beyond. However, it has one main limitation in that it can only identify linear "synchronous" relationships between the variables of a system. In particular, when several variables of the system have a delayed linear response to a given signal and if the time delay is not identical for all the variables, the PCA will generally fail to identify this dynamic relationship or will require several PCs to compressively capture it.

b. *Spectral PCA via the Morlet wavelet*

The spectral PCA (sPCA) relies on the computation of the cross-spectral matrix $\mathbf{S}_k$ between the time series $\mathbf{x}_n$ at locations $n = 1, \ldots, N$ and extraction of its eigenvectors in various frequency bands $b_k(f)$. In the frequency band $b_k(f)$, the complex $N \times N$ matrix $\mathbf{S}_k$ is defined as

$$\mathbf{S}_k = \frac{1}{\int_{-\infty}^{+\infty} b_k(f)df}\begin{bmatrix} \int_{-\infty}^{+\infty} s_{\mathbf{x}_1,\mathbf{x}_1}(f)b_k(f)df & \cdots & \int_{-\infty}^{+\infty} s_{\mathbf{x}_1,\mathbf{x}_N}(f)b_k(f)df \\ \vdots & \ddots & \vdots \\ \int_{-\infty}^{+\infty} s_{\mathbf{x}_N,\mathbf{x}_1}(f)b_k(f)df & \cdots & \int_{-\infty}^{+\infty} s_{\mathbf{x}_N,\mathbf{x}_N}(f)b_k(f)df \end{bmatrix}, \tag{5}$$

with $s_{\mathbf{x}_n,\mathbf{x}_m}(f)$ being the Fourier cross-spectrum between the two time series $\mathbf{x}_n$ and $\mathbf{x}_m$ at locations n and m, respectively, and $b_k(f)$ a bandpass transfer function centered on frequency f_k. The matrix $\mathbf{S}_k$ is a complex Hermitian matrix with real diagonal coefficients. The index k in $\{1, 2, \ldots, K\}$ is the frequency band index.

The sPCA is a direct extension of classical PCA in the Fourier domain. In the fields of signal processing and systems control theory, this approach, which consists in identifying empirical linear dynamical relations between the variables of the system, is more commonly referred to as linear dynamical systems identification (Picci and Pinzoni 1986; Georgiou and Lindquist 2019). While the essence of the approach always remains the same, it may be implemented in different ways. The various implementations of the method essentially differ in the way the cross-spectral matrix $\mathbf{S}_k$ is computed from the data and in the definition of the frequency bands $\{b_k(f)\}$.

A direct consequence of the greater flexibility (more degrees of freedom) of the sPCA method compared to the classical PCA is that the sPCA is prone to overfitting when the cross-spectral coefficients are not robustly estimated. For example, from a set of L observations (with L finite) of N variables, if no regularity is imposed on the cross-spectra, and if the number K of independent frequency bands of the empirical cross-spectra is greater than or equal to L, one can always fit a "perfect" empirical dynamical relation (i.e., a transfer function) between any two variables $\mathbf{x}_n$ and $\mathbf{x}_m$ and thus obtain a rank-one cross-spectral matrix for each frequency band (see appendix B). Therefore, there is a necessary trade-off between the spectral resolution (i.e., the number and width of frequency bands) and the robustness of the cross-spectral matrix estimation. The Bartlett and Welch periodogram methods (Bartlett 1950; Welch 1967; Proakis 2001) are the most frequently used methods for estimating robust (with low sampling variance) Fourier power spectra and cross-spectra from finite-length time series. They consist in splitting the time series into H segments and then performing a discrete Fourier transform (DFT) for each segment. The Welch method uses overlapping segments while the Bartlett method uses nonoverlapping segments. A robust estimation of the cross-spectral coefficients in each frequency band is obtained by averaging the complex cross-spectral coefficients obtained for each segment. As a trade-off for the greater robustness of the computed cross-spectral coefficients with reduced sample variance due to averaging, the Welch and Bartlett periodograms have reduced spectral resolutions since each one of the H segments corresponds to a shorter time series (see appendix B).

While the Welch and Bartlett methods for computing periodograms are classified as nonparametric (as they do not rely on a parametric spectrum model), the user still needs to define the number of segments for the Bartlett method, plus the overlapping fraction for the Welch periodogram. The commonly used "modified" version of the Welsh periodogram also requires selecting a windowing function to be applied to each segment. All these methodological choices affect the computed cross-spectral matrix, in particular its rank and the decay rate of its eigenvalues (see appendix B).

Other spectral estimation methods have been used in climate science to perform spectral PCAs. In Mann and Park (1994, 1999) Slepian tapers are used as weighting functions to compute cross-spectral quantities. The Slepian tapers are designed to minimize spectral leakage, that is, to compute cross-spectral coefficients over narrow frequency bands. The Slepian tapers method is adapted if, in the frequency band of interest, most of the variability of the system can be explained by a small number of PCs. Indeed, with this method, the rank of the empirical cross-spectral matrix is at most

equal to the number of orthogonal tapers. In Mann and Park (1994) and Mann et al. (2020), the method is implemented with 3–6 orthogonal tapers. In the cases when the system is too complex to be reduced to a few PCs, using the Slepian tapers method will inevitably lead to overfitting. The "multiwavelet" method (Lilly and Park 1995; Park and Mann 2000) corresponds to a localized in time version of the Slepian taper method. The principal oscillation patterns (POPs) introduced by Hasselmann (1988) can be seen as a parametric version of sPCA where the PCs are autoregressive moving average (ARMA) processes.

As a simple alternative to the spectral computation methods mentioned above, we propose to apply the sPCA by relying on the complex analytic Morlet wavelet to estimate the cross-spectral matrix. The continuous wavelet transform is obtained by convolving the analyzed signal with a basis of wavelet functions which are all dilated and translated versions of the same "mother" wavelet function. The Morlet mother wavelet (Morlet et al. 1982; Addison 2017) is defined as

$$\Psi(t) = \pi^{-1/4}(e^{i2\pi f_0 t} - e^{[-(2\pi f_0)^2/2]})e^{-t^2/2}, \tag{6}$$

which, if $2\pi f_0 > 5$, can be approximated as

$$\Psi(t) = \pi^{-1/4} e^{i2\pi f_0 t} e^{-t^2/2}. \tag{7}$$

The Morlet wavelet is therefore the complex exponential function $e^{i2\pi f_0 t} = \cos(2\pi f_0 t) + i\sin(2\pi f_0 t)$ modulated by a Gaussian envelope. The continuous wavelet transform $w_x(\nu, t)$ of the signal $x(t)$ is defined as

$$w_x(\nu, t) = \frac{1}{\sqrt{\nu}} \int_{-\infty}^{+\infty} x(u) \Psi'\left(\frac{u - t}{\nu}\right) du, \tag{8}$$

with ν being the scale parameter in the wavelet time-scale domain; $\Psi'(u)$ designates the complex conjugate of $\Psi(u)$.

At scale ν_k, $\mathbf{W}_k$ is the $N \times L$ matrix of wavelet coefficients derived from the data matrix $\mathbf{X}$:

$$\mathbf{W}_k = \begin{bmatrix} w_{x_1}(\nu_k, t_1) & \cdots & w_{x_1}(\nu_k, t_L) \\ \vdots & \ddots & \vdots \\ w_{x_N}(\nu_k, t_1) & \cdots & w_{x_N}(\nu_k, t_L) \end{bmatrix}. \tag{9}$$

Each scale ν_k in the wavelet scale domain corresponds to a frequency band b_k in the Fourier frequency domain with the central frequency of b_k being $f_k = f_0/\nu_k$ (see appendix B). Because of the correspondence between the Morlet wavelet transform and the Fourier transform [Eqs. (7) and (8) and appendix C], the empirical sample cross-spectral matrix for the frequency band b_k can be computed as

$$\widetilde{\mathbf{S}}_k = \frac{1}{L-1} \mathbf{W}_k \mathbf{W}_k'. \tag{10}$$

The Morlet wavelet method to compute the cross-spectral matrix can be seen as similar to using a modified Welch periodogram with a Gaussian window whose length varies inversely proportional to the frequency. Our proposed method can also be related to the "multiwavelet" method (Lilly and Park 1995; Park and Mann 2000), except that in our case, robustness is obtained through temporal integration rather than using multiple wavelets with mutually exclusive frequency support.

Similarly to the classical real-value PCA case, the complex eigenvectors $\mathbf{u}_{n,k}$ of the $\widetilde{\mathbf{S}}_k$ matrix are the solutions of the equation:

$$\widetilde{\mathbf{S}}_k \mathbf{u}_{n,k} = \lambda_{n,k}^2 \mathbf{u}_{n,k}. \tag{11}$$

The eigenvalues $\lambda_{n,k}^2$ of $\widetilde{\mathbf{S}}_k$ are real because $\widetilde{\mathbf{S}}_k$ is by construction a Hermitian matrix. As for the real case, we impose unit L2 norm for each eigenvector. The PC series of complex wavelet coefficients associated with the eigenvector $\mathbf{u}_{n,k}$ at the scale ν_k is

$$\boldsymbol{\kappa}_{n,k} = \frac{\lambda_{n,k}}{N} \mathbf{W}_k' \mathbf{u}_{n,k}. \tag{12}$$

The wavelet PC $\boldsymbol{\kappa}_{n,k}$ is a complex signal defined in the wavelet space; the real-value time series corresponding to the nth wavelet PC at scale ν_k can be reconstructed through an inverse wavelet transform. We can also combine several wavelet PCs at various scales and then apply the inverse wavelet transform to reconstruct a time series that encapsulates the variability of the original signal within any desired frequency band (range of scales). The reconstructed signal corresponds to a linear combination of the bandpass filtered and phase-synchronized original time series. One shall note that the phase shift between any two complex scalar elements of the $\mathbf{u}_{n,k}$ vector is interpretable as the phase shift at frequency f_k between the two time series at the corresponding locations. However, the argument of the scalar elements of $\mathbf{u}_{n,k}$ is arbitrary and not physically meaningful (similarly to the sign for real-valued eigenvectors). Indeed, if $\mathbf{u}_{n,k}$ is a solution of Eq. (11), so is $e^{i\omega}\mathbf{u}_{n,k}$ for any value of ω. However, multiplying $\mathbf{u}_{n,k}$ by $e^{i\omega}$ introduces a phase shift in the reconstructed PC. For easier interpretation of the timing of the reconstructed PC at frequency f_k, we choose the reconstructed PC which is aligned in phase with the highest contributing time series at that frequency (i.e., we impose a zero argument for the scalar element of $\mathbf{u}_{n,k}$ having the highest absolute value). When reconstructing a real time series from several wavelet PCs, it is important to control the relative phase of the combined wavelet PCs. Here we chose to impose for each PC a zero phase shift relatively to the time series having the highest total contribution across all the combined wavelet PCs. Also note that the inverse wavelet transform with the Morlet wavelet is not an exact reconstruction; however, in practice we find that the time series are well enough reconstructed to be interpretable.

The only parameter that has to be selected by the user for the implementation of the Morlet wavelet sPCA is the central frequency f_0 of the mother wavelet. It is generally chosen between 0.8 and 1; the value $f_0 = 0.849\ \{= [2\ln(2)]^{-1/2}\}$ can be chosen such that the magnitude of the second highest peak of the wavelet is half the magnitude of the highest peak (central peak). In practice, values between 0.8 and 1 will produce similar spectra and cross-spectra. Taking a value of f_0 higher than 1 will lead to a narrower spectral bandwidth (higher-frequency resolution). It will also lead to poorer localization of the wavelet coefficients in the time domain (lower time resolution), which in practice translates into wavelet coefficients more correlated in time (i.e., fewer *independent* samples for estimating the cross-spectral coefficients), meaning less robust empirical cross-spectra and

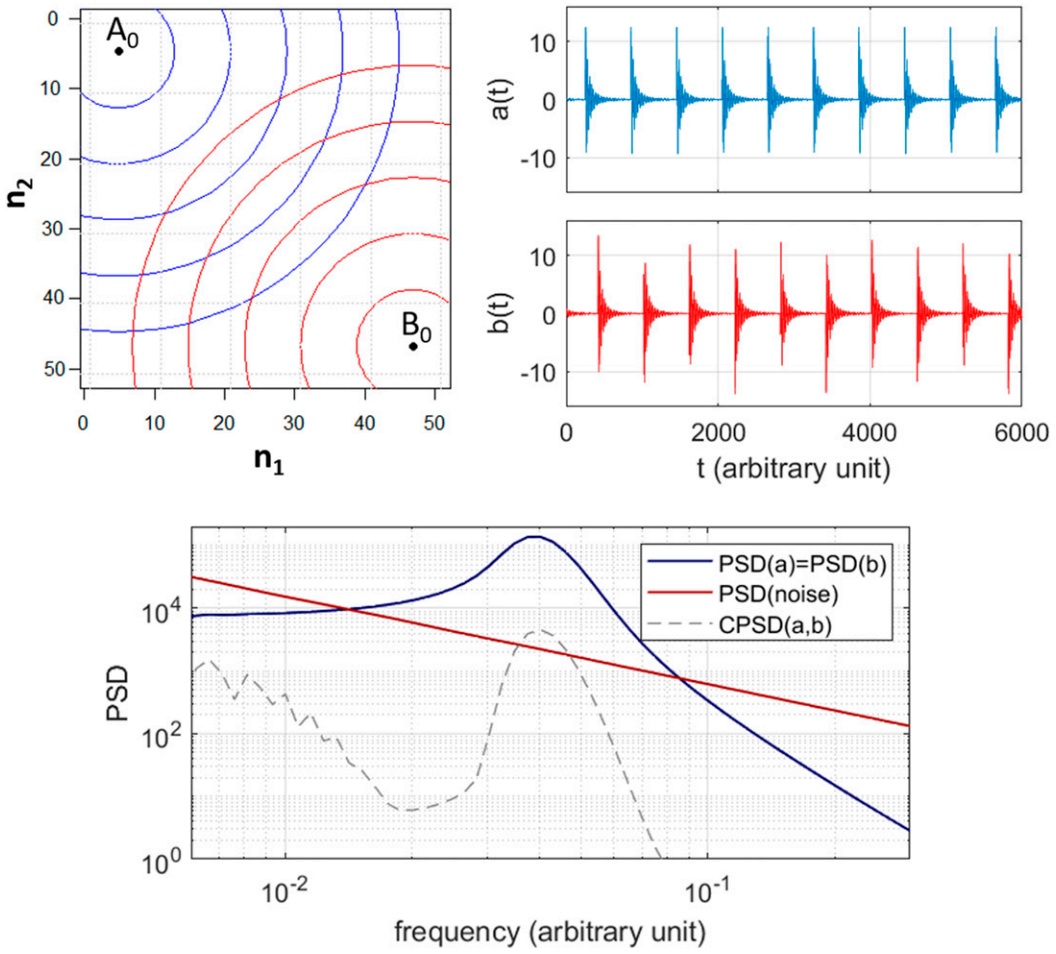

FIG. 1. Wave propagation example. (top left) Representation of the simulated 51 × 51 spatial grid domain with the origins A_0 and B_0 of the two waveforms $a(t)$ and $b(t)$ marked. (top right) Synthetic waveforms $a(t)$ and $b(t)$. (bottom) Temporal power spectral density and cross-spectral density of the waves $a(t)$ and $b(t)$ and of the noise $\mu_{n_1,n_2}(t)$. Power spectral densities (PSD) and cross-spectral densities (CPSD) are estimated using the Morlet wavelet.

higher risk of overfitting. As for the periodogram methods, the trade-off between frequency resolution and time resolution/number of independent samples is related to the Heisenberg–Gabor limit (Gabor 1946; Addison 2017). One of the interesting properties of the Morlet wavelet is that it actually reaches the Heisenberg–Gabor limit, therefore allowing the best possible time resolution for a given frequency resolution (i.e., minimal temporal correlation of the wavelet coefficients and maximal effective sample size). Note that, because the wavelets shrink as scale gets finer and the frequency gets higher [Eq. (8)], the effective sample size increases with frequency, allowing us to potentially extract robust coherent high-frequency modes even if their amplitude is low relatively to the noise. The counterpart of this is reduced frequency resolution at high frequencies.

The Morlet wavelet, being a differentiable wavelet (Vetterli and Kovacevic 1995; Addison 2017), is in theory blind to linear trends; that is, the wavelet transform of a linear function of time is zero everywhere. However, the wavelet coefficients will be locally affected by the ''transition points'' in the time series, which occur when different trends affect different portions of the analyzed signal. The spectral PCA is designed specifically to extract periodic modes of variability and is not a recommended tool for trend analysis. However, wavelets can be used to efficiently remove periodic signals from time series (Kumar and Foufoula-Georgiou 1997; Addison 2017) to then perform a trend analysis on the residual signal (e.g., through a classical PCA). Note that the wavelet coefficients are affected by edge effects at the beginning and at the end of a time series; these can be avoided by considering only coefficients outside of the cone of influence or by performing appropriate padding of the time series before the wavelet transform (see appendix D).

c. *Rotation of eigenvectors with spatial regularization for physical interpretability*

For the PCA and sPCA, the dimensionality reduction is obtained by projecting the data $\mathbf{X}$ (respectively the wavelet coefficients $\mathbf{W}_k$ at frequency f_k) onto the subspace generated by the first P eigenvectors (with $P < N$). The set of eigenvectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_P\}$ (respectively $\{\mathbf{u}_{1,k}, \mathbf{u}_{2,k}, \ldots, \mathbf{u}_{P,k}\}$) is an orthogonal basis of this subspace. One can generate other

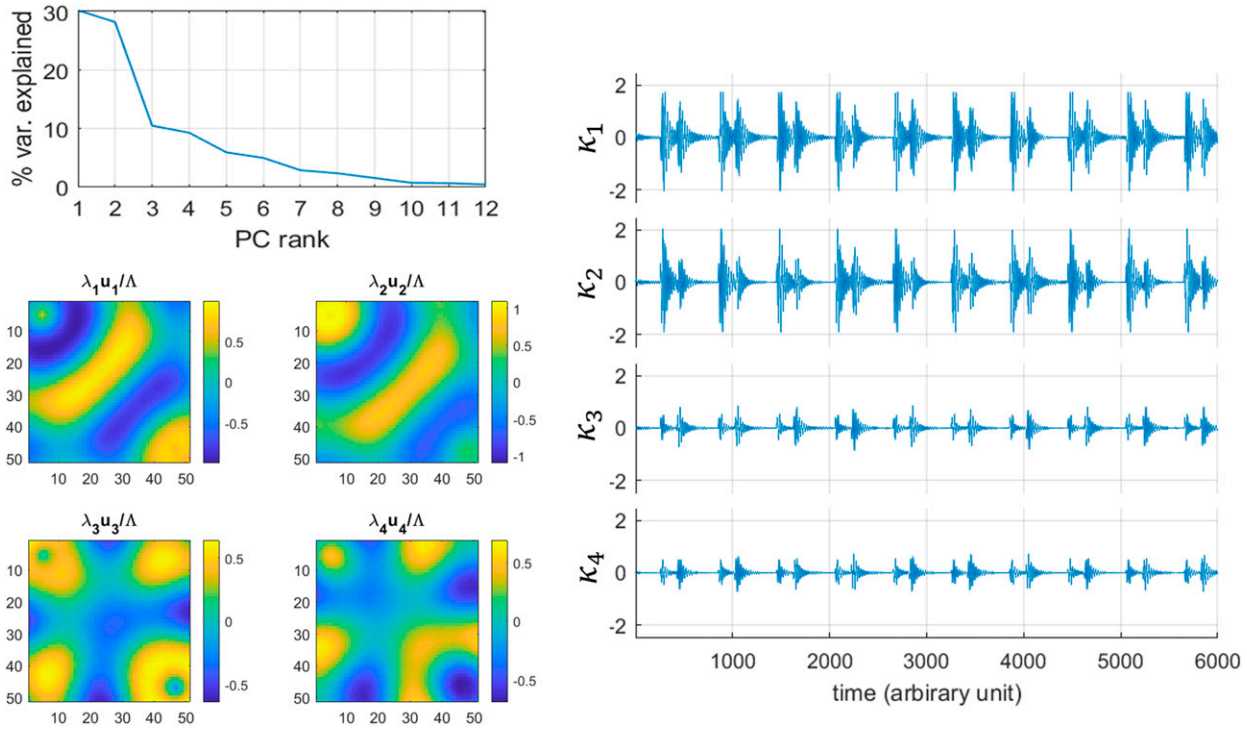

FIG. 2. Classical PCA applied to the synthetic wave propagation example, setup 1 (waves with different amplitudes), without noise. (top left) Fraction of the variance of the synthetic system explained by the first 12 PCs for the classical PCA. (bottom left) Spatial representation of the eigenvectors associated with the first four PCs, the unit-norm eigenvectors are shown with a scaling factor λ_n/Λ, with $\Lambda = \sqrt{\sum_{n=1}^{N} \lambda_n^2/N}$. (right) First four PC time series.

orthogonal bases of the same subspace by performing a rotation of the initial basis (i.e., by recombining linearly the eigenvectors while preserving orthogonality). While the information retained by the projection of the data onto the subspace would remain the same under any rotation of the basis, the physical interpretability of the basis vectors may be improved by finding the right rotation, which allows "unmixing" (i.e., distinguishing) physical modes of variability (Richman 1986). In particular, in the case when the PCA identifies two or more PCs of equal importance (i.e., associated with equal eigenvalues) all rotated bases are equally valid solutions to the SVD problem, and therefore without performing regularization by rotation one cannot find a unique solution of the PCA. In practice when the eigenvalues are relatively close to each other, physical modes of variability are more likely to be mixed. This fact had already been pointed out by Wallace (1971) and Wallace and Dickinson (1972), among others, as a limitation of the sPCA technique. While this also affects the classical PCA [see, e.g., the "rule of thumb" in North et al. (1982)], it is more likely to be troublesome for the sPCA as the higher number of degrees of freedom compared to the PCA makes it more likely for two separate physical modes to have a nonzero spectral coherence in a specific frequency band (than two separate physical modes having a nonzero correlation for the classical PCA).

Rotated PCA methods aim at finding a rotation of the eigenvectors (or rotation of the PCs) maximizing a given criterion (or minimizing a cost function). The varimax method (Kaiser 1958) is for example the most widely used form of rotated PCA. It consists in finding the rotation that maximizes the sum of the squared correlations between the original variables and the rotated PCs. While rotated PCAs have often been used in climate science (Mestas-Nuñez and Enfield 1999; Lian and Chen 2012; Chen et al. 2017), they have rarely been applied in a complex domain such as the Fourier domain, and, in these rare examples (Wallace and Dickinson 1972; Bloomfield and Davis 1994; Bueso et al. 2020), only using the varimax or promax criteria, which ignore the phase information and the spatial structure of the eigenvectors.

When each variable of the system can be attributed to a spatial location, the rotated eigenvectors can be mapped as a D-dimensional field whose structure and regularity can be analyzed and used as a criterion to find interpretable and meaningful rotated PCs. We propose here to evaluate the spatial regularity (smoothness) of the rotated eigenvectors through the Laplacian operator (sum of second partial derivatives). For the sPCA, the spatial structure of the phase (complex argument) of the eigenvectors is particularly informative, especially when one seeks to identify dynamical propagating modes. Indeed, for a propagating wave, the phase is expected to be a linear (or at least locally linear) function of space.

Let us consider $\mathbf{u}_{1,k}$ and $\mathbf{u}_{2,k}$ two unit-norm eigenvectors of the cross-spectral matrix $\widetilde{\mathbf{S}}_k$ associated with the $\lambda_{1,k}^2$ and $\lambda_{2,k}^2$ eigenvalues. We define

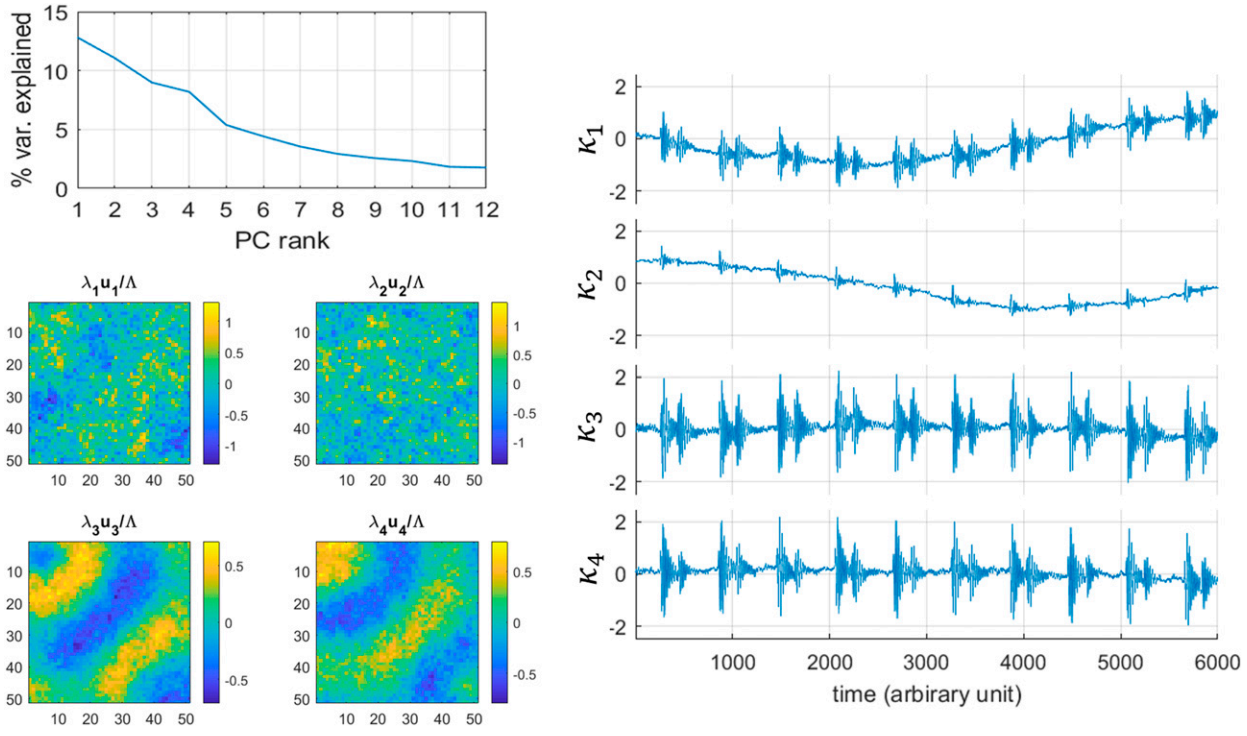

FIG. 3. As in Fig. 2, with noise added to the system. The noise accounts here for 70% of the system's variance.

$$\mathbf{u}^{\mathrm{R1}}_{\theta,\varphi,k} = \cos(\theta)\mathbf{u}_{1,k} + \sin(\theta)e^{i\varphi}\mathbf{u}_{2,k}, \tag{13}$$

and

$$\mathbf{u}^{\mathrm{R2}}_{\theta,\varphi,k} = -\sin(\theta)\mathbf{u}_{1,k} + \cos(\theta)e^{i\varphi}\mathbf{u}_{2,k}. \tag{14}$$

The vectors $\mathbf{u}^{\mathrm{R1}}_{\theta,\varphi,k}$ and $\mathbf{u}^{\mathrm{R2}}_{\theta,\varphi,k}$ are by definition unit-norm orthogonal vectors of $\mathbb{C}^N$. In the case when all N time series forming the $\mathbf{X}$ dataset are associated with a spatial location (in a D-dimensional space), the complex vectors $\mathbf{u}^{\mathrm{R1}}_{\theta,\varphi,k}$and $\mathbf{u}^{\mathrm{R2}}_{\boldsymbol{\theta},\boldsymbol{\varphi},\mathbf{k}}$ can be represented as complex functions of space $u^{\mathrm{R1}}_{\theta,\varphi,k}(\mathbf{y})$ and $u^{\mathrm{R2}}_{\theta,\varphi,k}(\mathbf{y})$; $\mathbf{y}$ is a D-dimensional coordinate vector. We can then define an objective function:

$$J(\theta,\varphi) = \int_{\mathbb{R}^D} [|\Delta u^{\mathrm{R1}}_{\theta,\varphi,k}(\mathbf{y})| + |\Delta u^{\mathrm{R2}}_{\theta,\varphi,k}(\mathbf{y})|]d\mathbf{y}, \tag{15}$$

where Δ denotes the Laplacian operator in D dimensions. To identify smooth rotated eigenvectors (i.e., with both phase and magnitude being locally regular in space), we look for the rotation parameters minimizing the objective function $J(\theta, \varphi)$:

$$(\theta_{\mathrm{opt}}, \varphi_{\mathrm{opt}}) = \mathrm{argmin}[J(\theta,\varphi)]. \tag{16}$$

The optimally rotated vectors at frequency f_k are therefore

$$\mathbf{u}^{\mathrm{opt1}}_k = \cos(\theta_{\mathrm{opt}})\mathbf{u}_{1,k} + \sin(\theta_{\mathrm{opt}})e^{i\varphi_{\mathrm{opt}}}\mathbf{u}_{2,k} \tag{17}$$

and

$$\mathbf{u}^{\mathrm{opt2}}_k = -\sin(\theta_{\mathrm{opt}})\mathbf{u}_{1,k} + \cos(\theta_{\mathrm{opt}})e^{i\varphi_{\mathrm{opt}}}\mathbf{u}_{2,k}. \tag{18}$$

In practice, with a finite number N of time series corresponding to discrete locations $\mathbf{y}_n$, $J(\theta, \varphi)$ is approximated as

$$\tilde{J}(\theta,\varphi) = \sum_{n=1}^{N} |\tilde{\Delta} u^{\mathrm{R1}}_{\theta,\varphi,k}(\mathbf{y}_n)| + |\tilde{\Delta} u^{\mathrm{R2}}_{\theta,\varphi,k}(\mathbf{y}_n)|, \tag{19}$$

where $\tilde{\Delta}$ is a discrete Laplacian operator (see appendix E). In fact, the minimized quantity is the L1 norm of the spatial Laplacian of the mapped eigenvector. We therefore term the proposed rotation criterion spatial Laplacian regularization (SLR). Rotations with more than two eigenvectors can be implemented by performing iterative pairwise rotations with all pairs of vectors until convergence to a stable solution.

3. Demonstration of the rotated spectral PCA (rsPCA) in a synthetic example

We demonstrate the possibilities of the proposed methodology by applying it to a synthetic data generated by a simple numerical model simulating two waves propagating in opposite directions on a 2D plane plus a colored random noise. The magnitudes, phase, and propagation speed of the waves are controlled, as is the noise level of the system. The simulated system corresponds to $N = 2601$ time series of length $L = 6000$ samples (with arbitrary time unit), spatially distributed on a 51×51 spatial grid (Fig. 1). At the (n_1, n_2) location (with n_1 and n_2 in $\{1, 2, \ldots, 51\}$), the time series $x_{n_1,n_2}(t)$ is

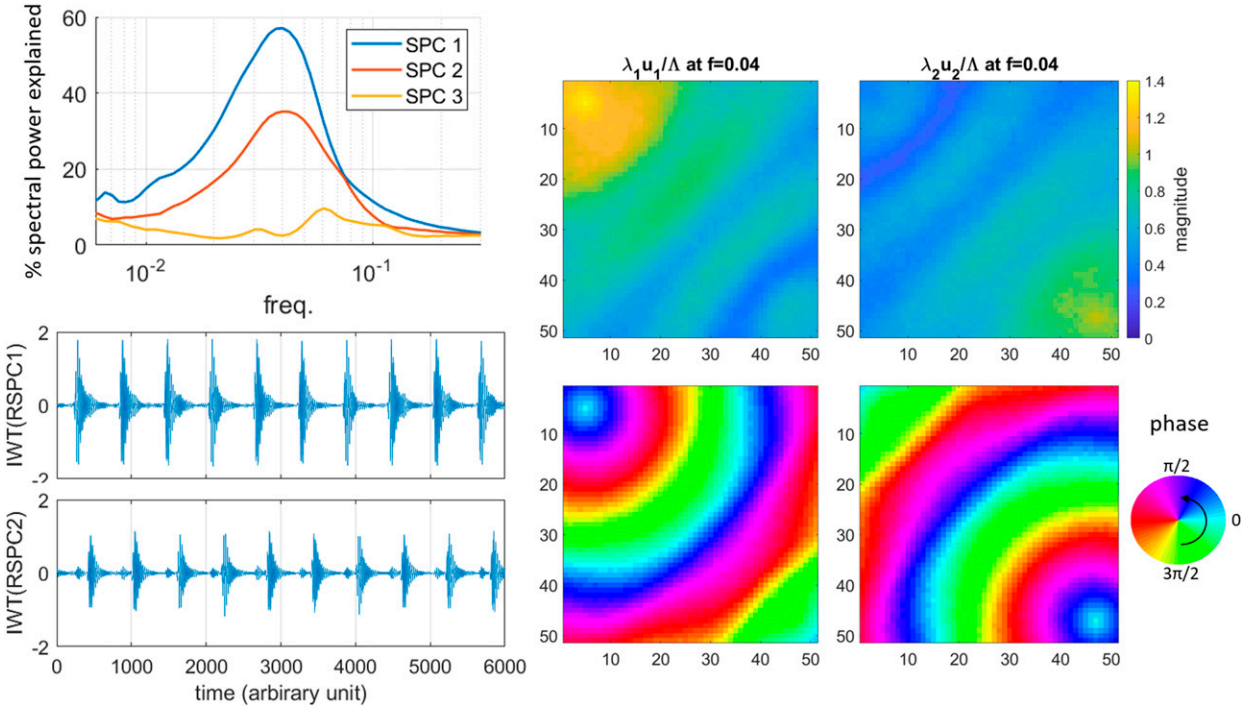

FIG. 4. Spectral PCA applied to the synthetic wave propagation example for setup 1. For this setup, $\gamma_1 = 0.8$ so the waves a and b have different amplitude. The noise accounts for 70% of the variance of the system. (top left) Fraction of the spectral power of the synthetic system explained by the first three PCs of the sPCA as a function of the frequency. (right) Spatial representation of the eigenvectors (magnitude and phase) associated with the first two PCs at frequency $f_k = 0.04$; the unit-norm eigenvectors are shown with a scaling factor $\lambda_{n,k}/\Lambda$, with $\Lambda = \sqrt{\sum_{n=1}^{N}\lambda_{n,k}^2/N}$. The circular arrow in the color scale of the phase indicates the direction of propagation of the extracted waves. (bottom left) Reconstructed signal through the inverse wavelet transform of the first (s^{-1}) wavelet PCs at all frequencies between 0.025 and 0.065.

$$x_{n_1,n_2}(t) = \frac{1}{d_0 + d_a} a\left(t + \frac{d_a}{c}\right) + \frac{1}{d_0 + d_b}\gamma_1 b\left(t + \frac{d_b}{c}\right) + \gamma_2 \mu_{n_1,n_2}(t), \tag{20}$$

with

$$d_a = \sqrt{(n_1 - 5)^2 + (n_2 - 5)^2}, \tag{21}$$

$$d_b = \sqrt{(n_1 - 47)^2 + (n_2 - 47)^2}. \tag{22}$$

The functions $a(t)$ and $b(t)$, shown in Fig. 1, are two series of nondispersive wave packets oscillating at the same frequency. The $a(t)$ signal originates from the (5,5) grid point (point A_0) and propagates linearly through the domain. Similarly, the $b(t)$ signal originates from the (47,47) grid point (point B_0) and propagates linearly through the domain. The parameter c controls the propagation speed of the $a(t)$ and $b(t)$ signals. The parameter d_0 is a constant ensuring that the amplitude of the waves remain finite at their origin. The parameters γ_1 and γ_2 control the relative amplitude of the two waves and of the noise $\mu_{n_1,n_2}(t)$. The noise $\mu_{n_1,n_2}(t)$ is a colored noise with both its temporal and spatial power spectral density (PSD) proportional to $f^{-1.4}$, where f is the Fourier temporal/spatial frequency (see Fig. 1 for the temporal PSD). The instantaneous phase shift between $a(t)$ and $b(t)$ is randomized (by adding a random phase shift in each individual wave packet of the $b(t)$ signal) to minimize the spectral coherence between the two waves. Two setups are presented below. In setup 1 the two waves have different amplitudes ($\gamma_1 = 0.8$). In setup 2, the two waves have the same amplitude ($\gamma_1 = 1$). For both setups, $d_0 = 30$ and $c = 1.2$.

The classical PCA is applied to identify the principal modes of variation of the 2601 time series of the synthetic system for setup 1, first without noise ($\gamma_2 = 0$) and then with a noise level adjusted such as the variance of the noise accounts for 70% of the total variance of the system. The result of the PCA without noise is shown in Fig. 2. Note that four PCs are needed to capture 80% of the variance and no less than nine PCs are needed to capture 95% of the variance. Additionally, none of the first four eigenvectors can be related specifically to one of the two waves but all four are rather a combination of both $a(t)$ and $b(t)$. This illustrates the poor ability of the classical PCA to simply capture dynamical linear relations between the variables when propagation effects are involved. Figure 3 shows the result of the PCA when noise is added to the system. One can see that all first four eigenvectors are strongly affected by the noise. The first two PCs essentially capture low-frequency components of the random noise, mixed with the $a(t)$ and $b(t)$ signals.

The spectral PCA is then applied to the system for setups 1 and 2 described previously and with a noise accounting in both cases for 70% of the system's variance. The results for

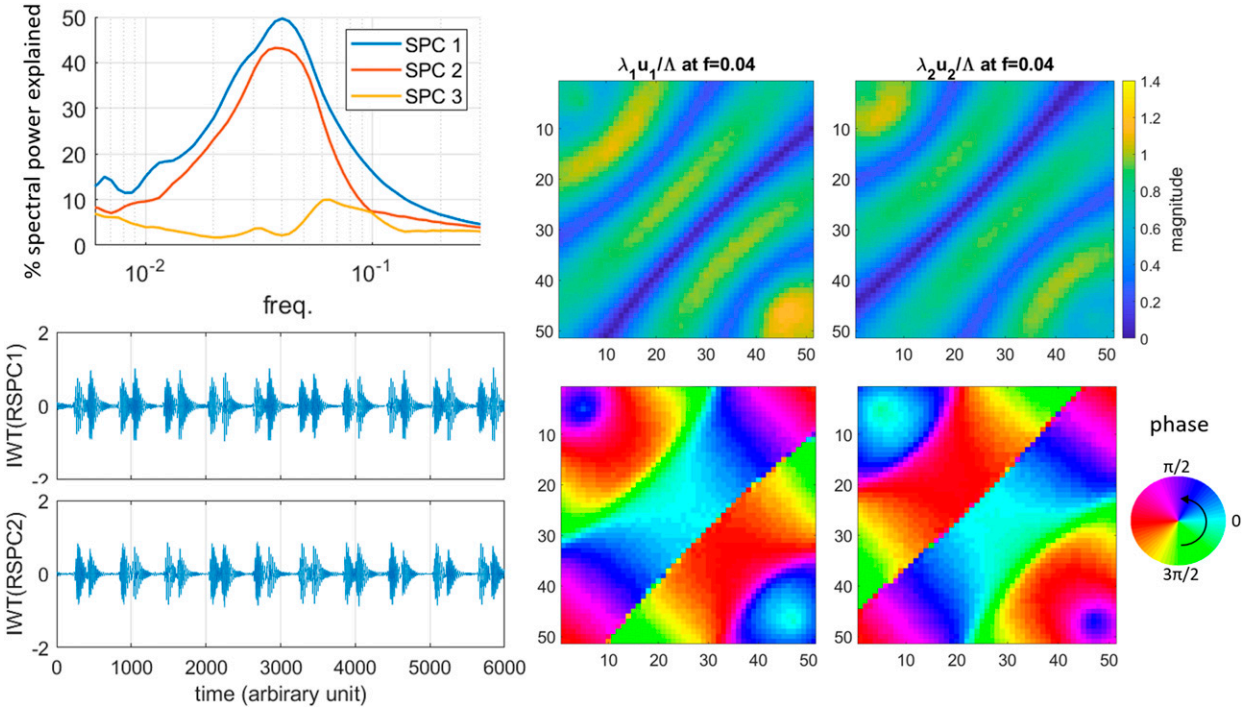

FIG. 5. As in Fig. 4, but for setup 2. For this setup, $\gamma_1 = 1$ so the waves a and b have same amplitude. The noise accounts for 70% of the variance of the system. It is observed that in this case the first two PCs are hardly interpretable since they are both reflecting mixtures of the two waves.

setup 1 are shown in Fig. 4. For the frequency band centered at the frequency $f_k = 0.04$ the first two PCs explain respectively 56% and 35% of the spectral power. The phase and magnitude structure of the two corresponding eigenvectors accurately describe the propagation of $a(t)$ and $b(t)$ with good separation of the two waves. The counterclockwise progression of the phase corresponds to a forward delay in time. The signal reconstructed through inverse wavelet transform from the first wavelet PCs at frequencies 0.025–0.065 corresponds mostly to the $a(t)$ signal while the signal reconstructed from the second PCs within the same frequency range corresponds mostly to the $b(t)$ signal. In spite of the relatively good separation of the two waves some interference patterns are visible in the spatial structure of the eigenvectors (in both the phase and the magnitude). These arise because the two waves are not perfectly independent and can be locally coherent over some periods of time. Actually, the spectral coherence between the waves $a(t)$ and $b(t)$ at the peak frequency is 0.1 (see Fig. 1). In practice, for a real case, nonsignificant coherence between two dynamical modes, arising from numerical approximations and limited number of samples, can have the same effect.

In setup 2, the waves a and b have same total energy. The results of the sPCA for this setup are shown in Fig. 5. In this case, sPCA is able to determine that two modes of equal

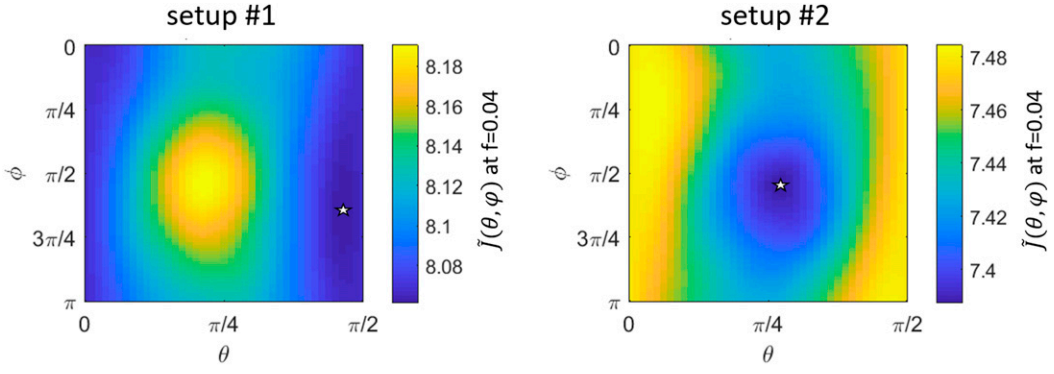

FIG. 6. Spatial Laplacian regularization (SLR) objective function $\tilde{J}(\theta, \varphi)$ as a function of the rotation parameters θ and φ for the rotation of the first two eigenvectors of the sPCA $\mathbf{u}_{1,k}$ and $\mathbf{u}_{2,k}$ at frequency 0.04 for the two setups of the synthetic wave propagation example. In both panels, the white star indicates the minimum of $\tilde{J}(\theta, \varphi)$.

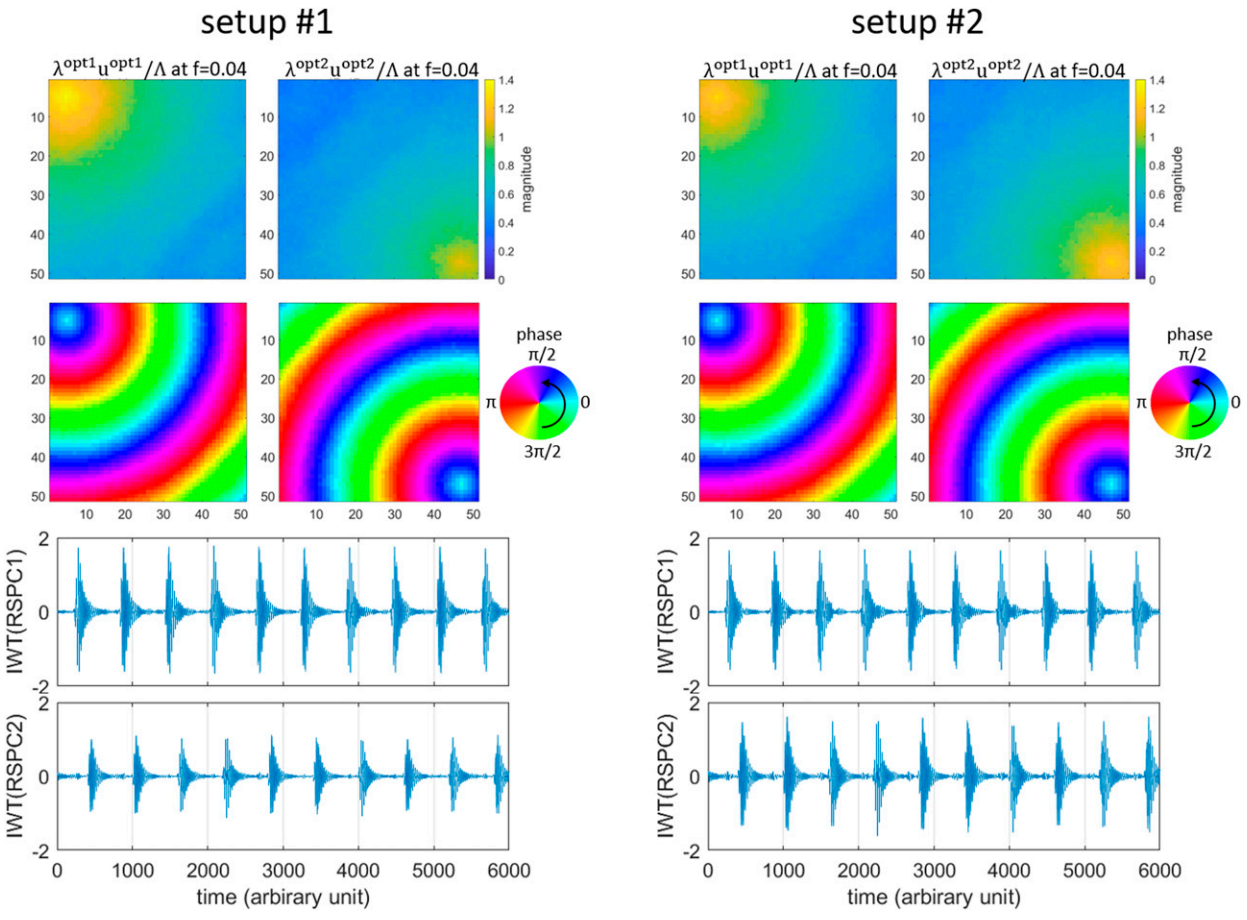

FIG. 7. Result of the proposed rotation procedure of the first two eigenvectors of the sPCA $\mathbf{u}_{1,k}$ and $\mathbf{u}_{2,k}$ at frequency 0.04 applied to the synthetic wave propagation example for the setups 1 (waves with different amplitude) and 2 (waves with same amplitude). (top) Optimally rotated vectors $\mathbf{u}_k^{\text{opt1}}$ and $\mathbf{u}_k^{\text{opt2}}$. The unit-norm vectors are shown with a scaling factor $\lambda_k^{\text{opt}n}/\Lambda$, with $\Lambda = \sqrt{\sum_{n=1}^{N}\lambda_{n,k}^2/N}$. (bottom) Reconstructed signal through the inverse wavelet transform of the first (s^{-1}) optimally rotated wavelet PCs at all frequencies between 0.025 and 0.065.

importance, explaining together more than 90% of the spectral power, exist around the frequency 0.04. However, the first two eigenvectors are not easily interpretable since both of them are reflecting a complex combination of the waves *a* and *b*. As mentioned before, with the two waves having equal amplitude, any two orthogonal unit-norm vectors being linear combinations of the two eigenvectors corresponding to the signals *a* and *b* can be solutions of the SVD problem. In practice, numerical approximations and small fluctuations caused by the noise will lead to a random "optimal" decomposition. The result of the sPCA is therefore extremely sensitive to noise and is also unstable across nearby frequency bands in that case.

We then perform the rotation of the eigenvectors $\mathbf{u}_{1,k}$ and $\mathbf{u}_{2,k}$ at frequency 0.04 with the SLR criterion as described in section 2d for the two setups presented above. Figure 6 shows the value of the criterion $\tilde{J}(\theta, \varphi)$ as a function of the rotation parameters θ and φ. For both setups, the function $\tilde{J}(\theta, \varphi)$ shows a unique minimum on the $[0, \pi/2] \times [0, \pi]$ hypercube. One shall note that eight symmetrical solutions always exist on $[0, 2\pi] \times [0, 2\pi]$, which correspond to $\mathbf{u}_k^{\text{opt1}}$ and $\mathbf{u}_k^{\text{opt2}}$ being switched and/or multiplied by -1. For the setup 1, the pair $(\theta_{\text{opt}}, \varphi_{\text{opt}})$ minimizing $\tilde{J}(\theta, \varphi)$ is $(0.93\pi/2, 1.32\pi/2)$. The values $\cos(\theta_{\text{opt}}) = 0.11$ and $\sin(\theta_{\text{opt}}) = 0.99$ reveal that the initial solutions identified by the sPCA were relatively close to the optimally rotated solutions. For setup 2, $(\theta_{\text{opt}}, \varphi_{\text{opt}}) = (0.55\pi/2, 1.1\pi/2)$ with $\cos(\theta_{\text{opt}}) = 0.65$ and $\sin(\theta_{\text{opt}}) = 0.76$ revealing that the eigenvectors $\mathbf{u}_{1,k}$ and $\mathbf{u}_{2,k}$ identified by the sPCA were equipoised mixtures of the optimally rotated vectors $\mathbf{u}_k^{\text{opt1}}$ and $\mathbf{u}_k^{\text{opt2}}$.

The rotation converged to identical vectors $\mathbf{u}_k^{\text{opt1}}$ and $\mathbf{u}_k^{\text{opt2}}$ leading to identical PCs (up to the difference in the magnitude of the second PC) in the two cases (Fig. 7). One can see that, after rotation, the two waves *a* and *b* are effectively separated. The separation of the two waves is improved even for setup 1, which for the initial decomposition found by the sPCA was relatively good in identifying separately the waves *a* and *b*. We can estimate the propagation speed from the phase information from the spatial structure of the phase of the eigenvectors: a phase shift of 2π at $f_k = 0.04$ corresponds to 30 grid increments, and we therefore estimate $\tilde{c} = 30 \times 0.04 = 1.2$, which is an accurate estimate of the true propagation speed.

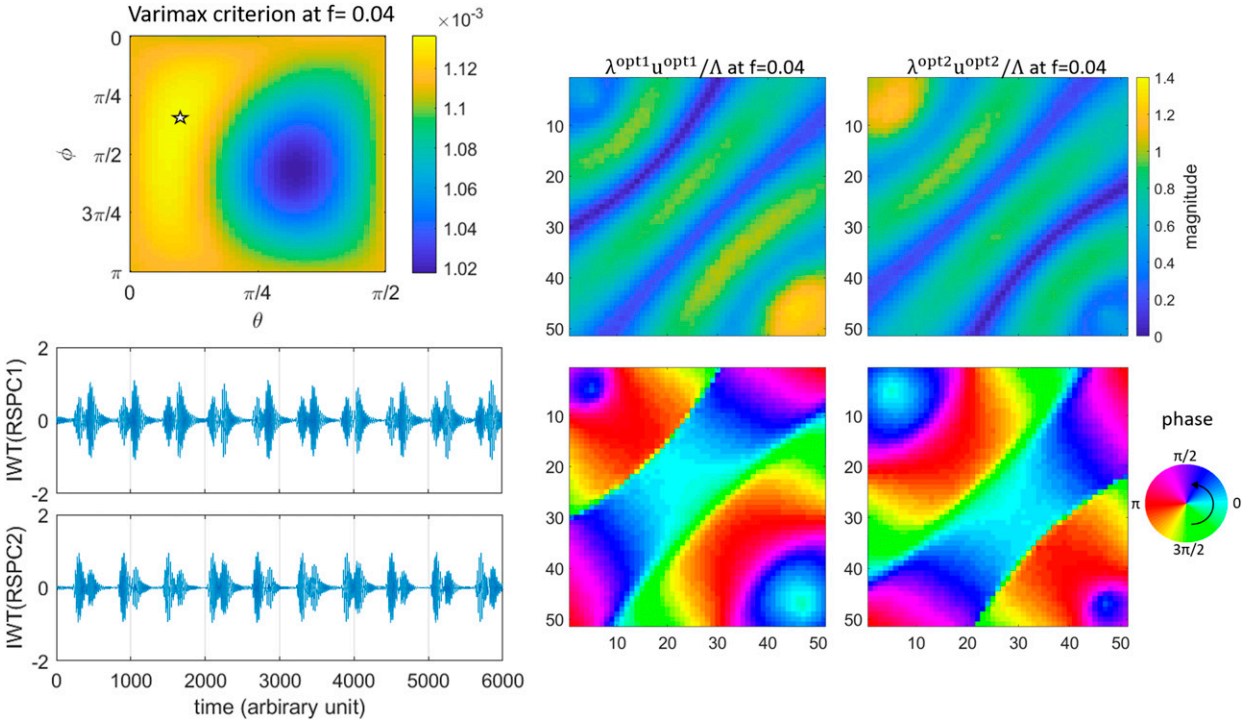

FIG. 8. Result of the varimax rotation of the first two eigenvectors of the sPCA $\mathbf{u}_{1,k}$ and $\mathbf{u}_{2,k}$ at frequency 0.04 for the setup 2 (waves with same amplitude) of the synthetic wave propagation example. (top left) Varimax criterion as a function of the rotation parameters θ and φ, the white star indicates the maximum, i.e., the optimal rotation. (right) Optimally rotated vectors $\mathbf{u}_k^{\text{opt1}}$ and $\mathbf{u}_k^{\text{opt2}}$. The unit-norm vectors are shown with a scaling factor $\lambda_k^{\text{opt}n}/\Lambda$, with $\Lambda = \sqrt{\sum_{n=1}^{N}\lambda_{n,k}^2/N}$. (bottom left) Reconstructed signal from the combination of the first (s^{-1}) optimally rotated wavelet PCs at all frequencies between 0.025 and 0.065.

For comparison the result of the rotation with the varimax criterion for case 2 is shown in Fig. 8. The varimax criterion ignores the phase information and the spatial structure of the eigenvectors. As it maximizes the variance of the square of the loadings, it promotes sparse solutions (i.e., rotated eigenvectors with a few high-magnitude loadings and many loadings of magnitude close to zero). However, in our synthetic case, as both waves propagate across the whole domain, their support is not sparse. The varimax criterion therefore selects a solution highlighting constructive and destructive interferences between the two waves rather than unmixing them. In fact, the varimax criterion can effectively help separate two waves only in the case when these have distinct spatial supports.

4. Analysis of global geopotential height and sea surface temperature

Identifying and extracting dominant climate modes across a range of space–time scales, such as interannual to decadal SST modes (Mantua et al. 1997; Trenberth 1997; Newman et al. 2016; Wang et al. 2017) and seasonal to subseasonal modes in the atmosphere (Madden and Julian 1971; Trenberth and Paolino 1981; Leathers et al. 1991; Trenberth and Hurrell 1994; Thompson and Wallace 1998; Feldstein 2000; Hurrell et al. 2013), has received a lot of attention. PCA methodologies have been a primary tool in such studies, with the many variants and extensions of the classical PCA method as discussed in the introduction. However, unmixing modes and accurately extracting propagating anomalies still present unique challenges. In this section, we apply the proposed rsPCA methodology to global daily SST and 500-mb (1 mb = 1 hPa) GPH time series, with the aim of highlighting its advantages in identifying and unmixing modes of climate variability at any desired temporal scale, as well as depicting the propagation of dynamical modes. We focus on subseasonal and interannual scales.

a. *Data used for analysis*

The analyzed data consist of a record of 71 years (1948–2019) of global daily 500-mb GPH from the NCEP–NCAR reanalysis project (Kalnay et al. 1996) and daily global SST observations from September 1981 to December 2014 from the NOAA Optimum Interpolation Sea Surface Temperature (OISST) analysis (Reynolds et al. 2007). The original GPH data are provided on 2.5° × 2.5° latitude/longitude grid while the original SST data are given on a 0.25° × 0.25° grid. For the purpose of this study, both datasets are reprojected on 220 km × 220 km equal-area pixels through a Mollweide projection (Snyder 1977). The reprojection on equal-area pixels is recommended before applying PCA or sPCA, to avoid giving excessive weight to high-latitude areas in terms of contribution to the system's variance.

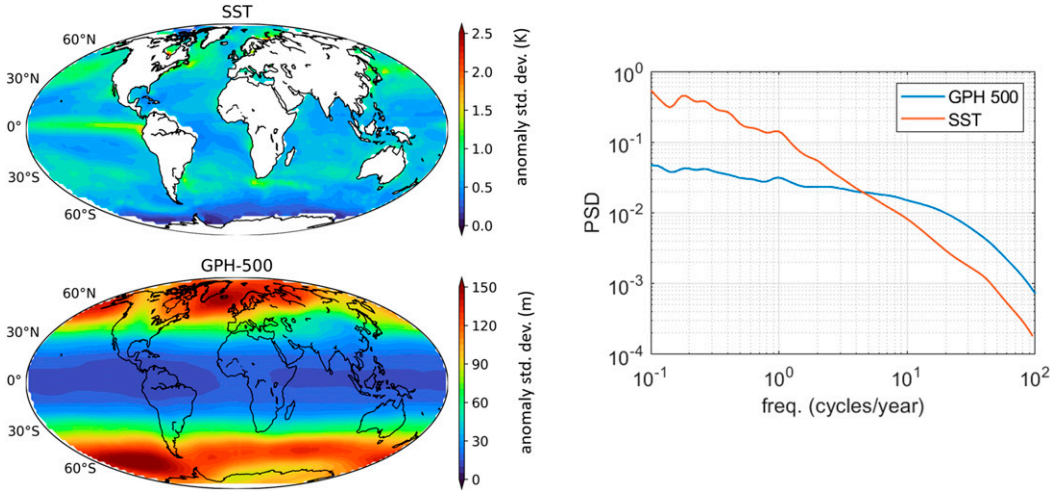

FIG. 9. Distribution of energy (variance) of SST and 500-mb GPH daily anomalies across space and across temporal frequencies. (left) Standard deviation of the anomaly time series in each pixel. (right) Temporal power spectral densities (PSD) of the GPH and SST time series. PSDs are estimated using the Morlet wavelet for each pixel and then averaged over all pixels. PSDs are normalized by the total variance of the system to be comparable.

Because the temporal variations of GPH and SST are dominated by the seasonal cycle, which is not what we are trying to characterize here, we removed the seasonal variations of the time series by subtracting the climatic mean of each calendar day in every pixel to obtain series of climatic anomalies. The climatic mean for each calendar day corresponds to the average over the whole time series of the 15-day period centered on this day. The studied systems are of dimension $N = 10\,284$ pixels and $L = 26\,298$ time steps for the GPH data, and $N = 7588$ pixels (land pixel being excluded) and $L = 12\,167$ time steps for the SST data.

Figure 9 shows how the energy (variance) of these variables is distributed across space and across temporal frequencies. The variability of the GPH is much lower in the tropics than at middle and high latitudes, whereas for the SST the variability is found more evenly distributed across the globe. Comparing the power spectral density of the two variables, we can see that SST shows relatively more variability than GPH at temporal scales between 3 months and 10 years while GPH shows more variability for periods shorter than 3 months (i.e., S2S time scales).

b. *Results*

Similarly to previous published studies (Weare et al. 1976; Wallace et al. 1993; Messié and Chavez 2011), the classical PCA is first applied to the global GPH and SST anomalies. Figure 10 shows how much of the variance is captured by the first 400 PCs for both variables. One can see that the first two PCs explain a higher fraction of the variance for the SST than for 500-mb GPH; however, the 50 first PCs explain 80% of the variability for GPH while they explain only 65% for SST. For

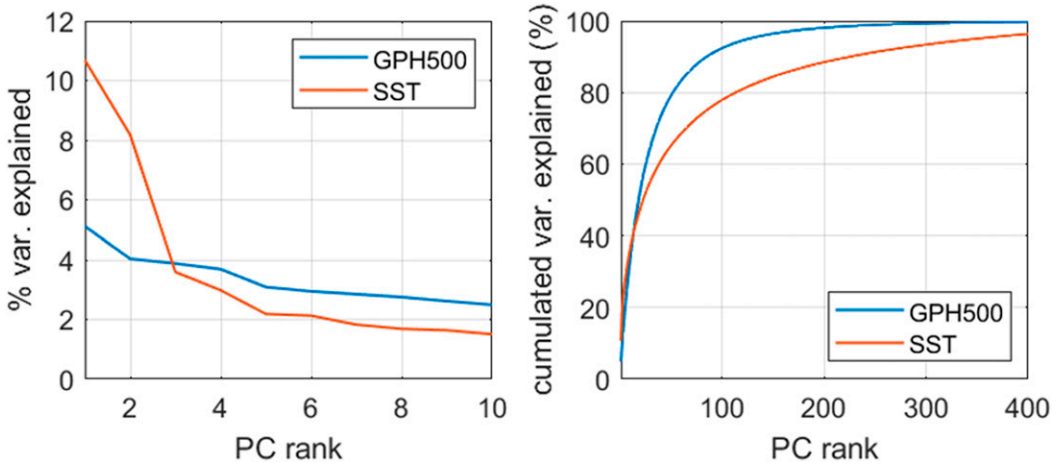

FIG. 10. (left) Percentage of the variance explained by the first 10 PCs for the classical PCA applied to daily SST and 500-mb GPH anomalies. (right) Cumulated variance explained by the first 400 PCs.

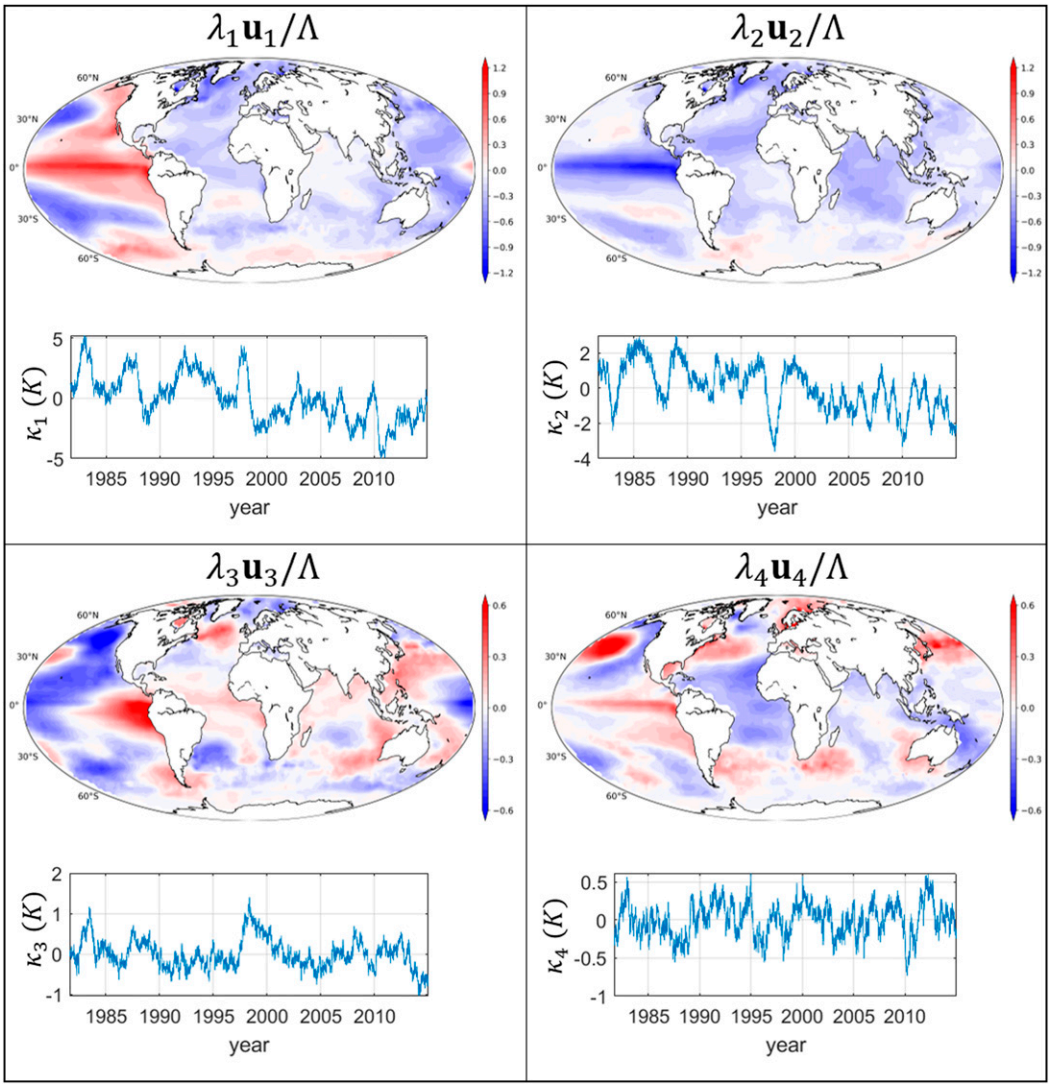

FIG. 11. Result of the classical PCA applied to daily SST anomalies, first four eigenvectors and associated PC time series. The unit-norm eigenvectors are shown with a scaling factor λ_n/Λ, with $\Lambda = \sqrt{\sum_{n=1}^{N}\lambda_n^2/N}$.

the SST, the first two modes, accounting respectively for 10.8% and 8.3% of the variance, appear to be related to the El Niño–Southern Oscillation (ENSO) signal (Trenberth 1997; Wang et al. 2017) (3–7-yr cycle) and to anthropogenic climate change (trend) with a strong signal over the tropical Pacific Ocean (Fig. 11). The third and fourth modes seem to resemble the patterns of the North Pacific gyre oscillation (Di Lorenzo et al. 2008, 2009) and the Pacific decadal oscillation (PDO) (Mantua et al. 1997; Newman et al. 2016). However, all these physical modes of variability and change are not perfectly separated by the PCA, as the first two PCs both reflect part of the ENSO and climate change signals, and the third and fourth PCs both reflect signals of the PDO and the North Pacific gyre oscillation. This is therefore a case for which a rotation of the PCs with an adapted criterion may allow a better physical attribution and interpretation (e.g., Chen and Wallace 2016; Chen et al. 2017; Wills et al. 2018). We note that all first four modes identified by the classical PCA are low-frequency modes corresponding to interannual to decadal time scales.

For the 500-mb GPH all first six PCs affect the medium and high latitudes (Fig. 12). The first one is located over Antarctica and the Southern Ocean. It forms a dipole as the polar region and the region between 50° and 70°S show an opposite response, which seems to resemble the pattern of the Antarctic Oscillation (AAO) (Thompson and Wallace 2000). The first PC shows a long-term trend (which is consistent with studies reporting current and projected trends in the AAO; Cai et al. 2003; Shindell and Schmidt 2004), together with high-frequency variations (typically 10–50-day oscillations). The other five first PCs only show high-frequency variations with no noticeable trend. The patterns shown by the eigenvectors with a succession of negative and positive values along

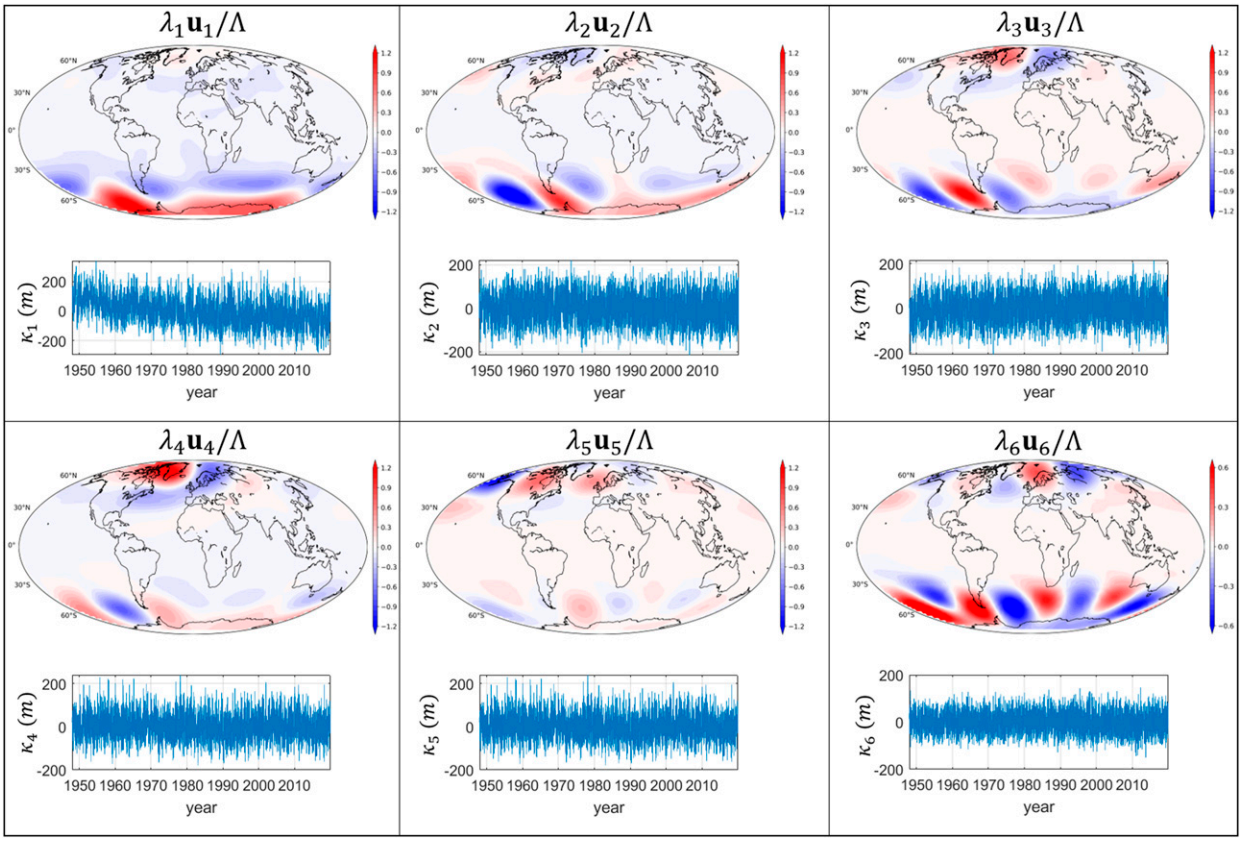

FIG. 12. Result of the classical PCA applied to daily 500-mb GPH anomalies, first six eigenvectors, and associated PC time series. The unit-norm eigenvectors are shown with a scaling factor λ_n/Λ, with $\Lambda = \sqrt{\sum_{n=1}^{N}\lambda_n^2/N}$.

latitudinal bands are typical wave propagation patterns (Kidson 1999; Roundy and Schreck 2009; Roundy 2015). However, nothing can be said about the direction and speed of propagation of the waves from the eigenvectors.

The wavelet-based spectral PCA is then applied to the GPH and SST datasets. Figure 13 shows how much of the spectral power of the system is explained by the four first spectral PCs as a function of the frequency. For the SST at periods between 2 and 10 years, the first spectral PC explains 50%–70% of the spectral power. For subannual frequencies, the fraction of spectral power explained by the first four PCs decreases at higher frequencies. For periods shorter than two months, each

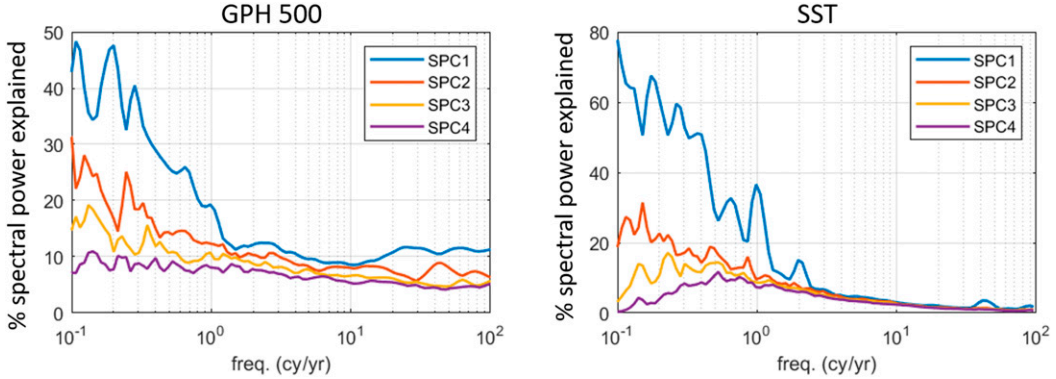

FIG. 13. Fraction of the spectral power explained by the first four spectral PCs of the sPCA as a function of the frequency for daily 500-mb GPH and SST anomalies.

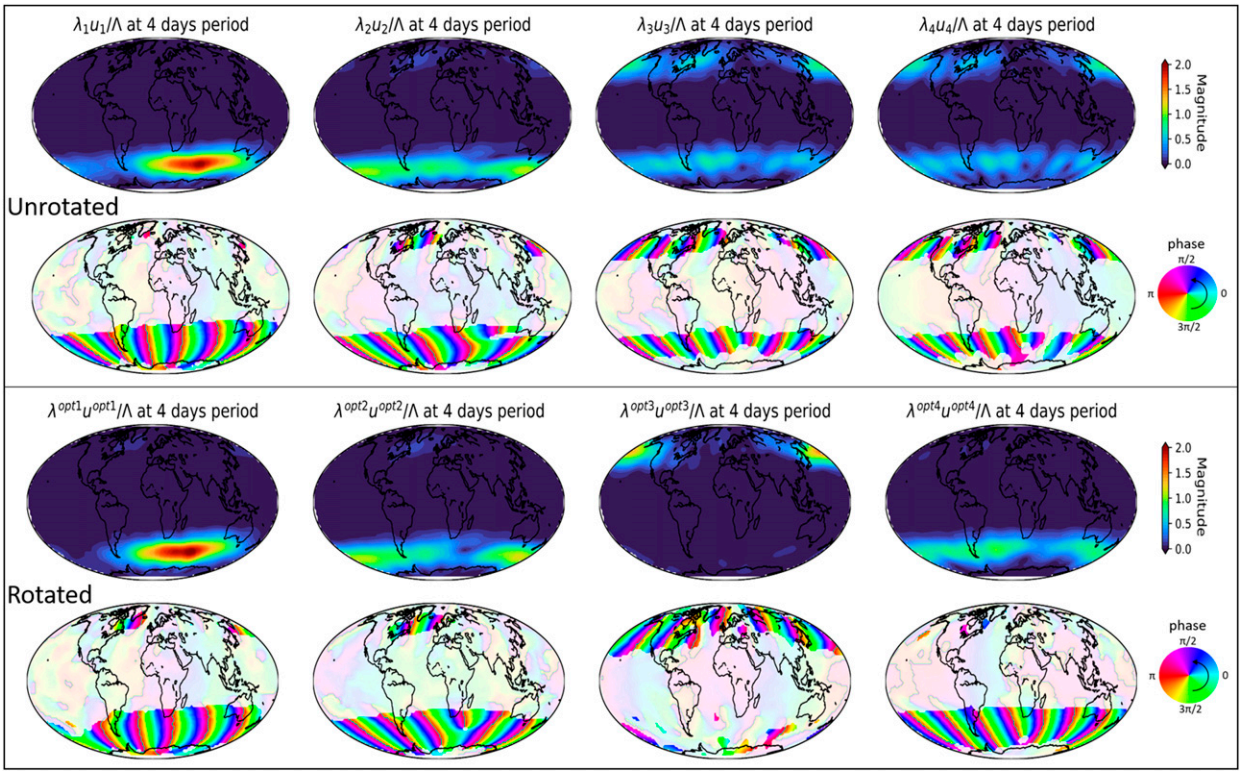

FIG. 14. Results of the sPCA and rsPCA applied to daily 500-mb GPH anomalies at the 4-day period. (top) First four unrotated eigenvectors (magnitude and phase); the unit-norm eigenvectors are shown with a scaling factor $\lambda_{n,k}/\Lambda$, with $\Lambda = \sqrt{\sum_{n=1}^{N} \lambda_{n,k}^2/N}$. The circular arrow in the color scale of the phase indicates the direction of propagation of the extracted waves. (bottom) First four optimally rotated vectors, using the proposed spatial Laplacian regularization (SLR) criterion. It is noted how the rotation "unmixes" the dynamic modes (i.e., separates the waves in the Northern and Southern Hemispheres).

one of the first four PCs explains less than 4% of the spectral power. For the 500-mb GPH, the first four PCs each explain between 5% and 12% of the spectral power for periods shorter than 6 months.

Figure 14 (top panel) shows the first four eigenvectors identified by the sPCA for the 500-mb GPH at the 4-day period, which is the highest frequency that can reasonably be analyzed with daily data. All the first four spectral PCs affect the latitudes poleward from 45°, specifically the first two PCs affect mostly the Southern Hemisphere, while the third and fourth ones affect both hemispheres. The SLR rotation procedure is performed among the first five eigenvectors to unmix their patterns (Fig. 14, bottom panel; the fifth vector is not shown). The first two eigenvectors are only slightly affected by the rotation, but the third and fourth are better differentiated after rotation: the third rotated vector affects the Northern Hemisphere and the fourth one affects the Southern Hemisphere. All four vectors depict waves of GPH anomalies with an eastward propagation, which are most likely the signature of Rossby wave propagation on top of the mean circulation in the atmosphere (i.e., the jet stream). At the 15-day period (Fig. 15), the first eigenvector also shows a wave with eastward propagation southward of 50°S, consistently corresponding to Rossby waves. This pattern in the Southern Hemisphere is consistently found among the first two eigenvectors of the sPCA for all periods between 4 and 60 days (not shown) with a spatial wavelength increasing with decreasing frequency, as expected from Rossby wave theory (Rossby 1945; Platzman 1968; Gill 1982). In the Northern Hemisphere, at the 15 days period, the second and fourth eigenvectors after rotation show a meandering pattern northward of 50°N with a westward propagation.

For the SST, the first eigenvector at frequencies 1/4 and 1/6 yr^{-1} is shown in Fig. 16. The spectral PCs associated with the first eigenvector represent respectively 59% and 68% of the spectral power at these two frequencies. The patterns at these two frequencies are consistent and correspond to the ENSO patterns that also appeared in the first two eigenvectors of the classical PCA. The tropical Pacific Ocean is the region of the globe showing the strongest response to this mode. Regions in the South and North Pacific also respond to this mode with an opposite phase (π shift) to what is found in the tropical region, meaning anomalies of opposite sign. Outside of the Pacific, some regions of the Atlantic and Indian Oceans seem to also have a weaker response to the ENSO signal; however, for these regions the patterns are not consistent between the 4- and 6-yr period,

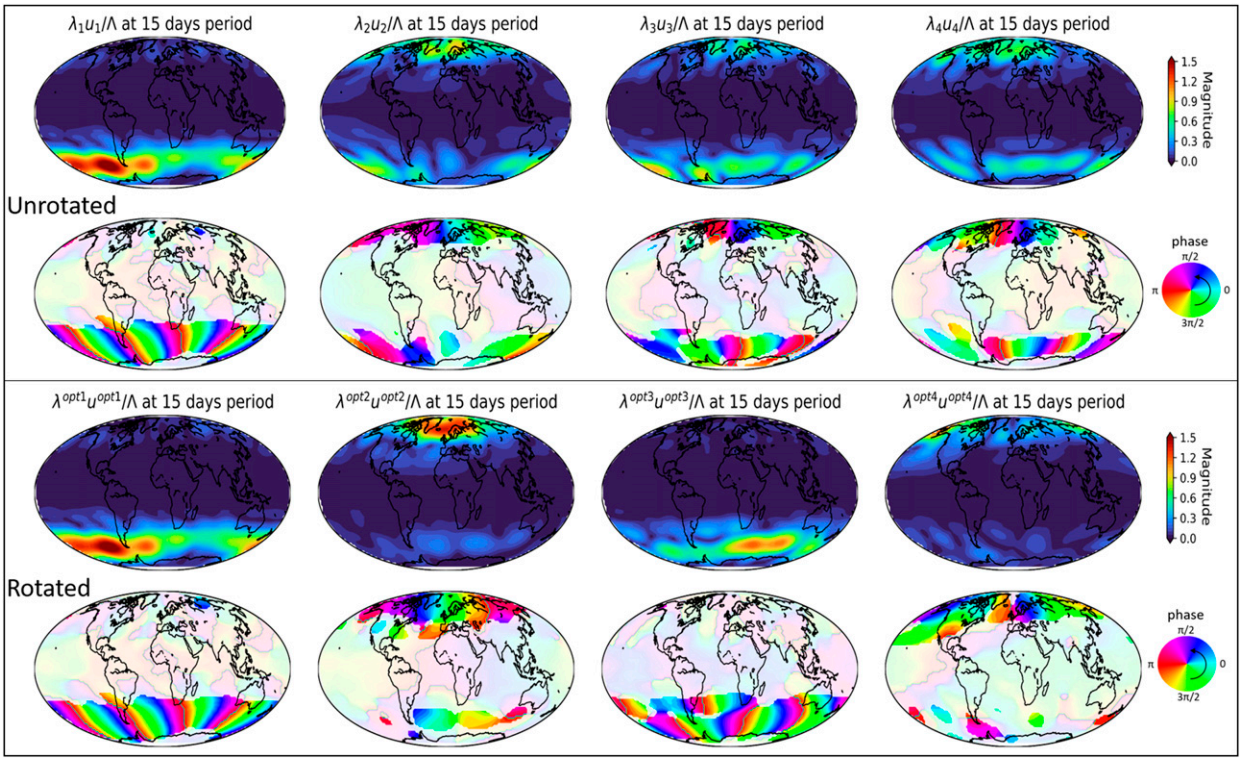

FIG. 15. As in Fig. 14, but at the 15-day period.

particularly in terms of phase. No neat propagation patterns can be found in the phase of the first eigenvector at the two considered frequencies. ENSO is therefore characterized mostly as a stationary wave (a dipole). The bottom panel of Fig. 16 shows the reconstructed signal through the inverse wavelet transform of the first wavelet PC at all frequencies between 1/2 and 1/7 yr^{-1}. As explained in section 2b, the reconstructed signal is a linear combination of the bandpass filtered and phase-synchronized SST anomaly time series. The bandpass filtering through the (inverse) wavelet transform allows a smooth signal and well-defined peaks. The phase synchronization also fosters well-defined peaks as the combined phase-synchronized time series will generally all peak at the same time. The reconstructed signal is consistent with the well-known historical variations of ENSO.

At the 15-day period, the first spectral PC of the SST only accounts for 2% of the spectral power (Fig. 13). However, even if noisy, the spatial pattern of the associated eigenvector is coherent. It appears as a wave around 50°S with an eastward propagation (Fig. 17). It is consistent with the Rossby wave pattern identified by the rsPCA in the 500-mb GPH at that same frequency of 15 days. Therefore, this mode is interpreted as the signature of an atmospheric Rossby wave on the SST and reveals ocean–atmosphere coupling. This shows the ability of the spectral PCA to extract coherent modes of low amplitude from random variability and noise.

The results shown in this section clearly illustrate the advantage of the wavelet-based sPCA against the classical PCA, in separating modes that correspond to different frequencies, and in extracting propagation information. Moreover, it is demonstrated that, in the presence of several competing modes at subannual scales, rotation of the eigenvectors and optimal unmixing of the underlying climate modes via the proposed rsPCA methodology is essential for physical interpretability.

5. Conclusions

The need for understanding patterns of variability and change in climate signals for the purpose of predictive and diagnostic analysis (e.g., for regional prediction, untangling the forced signal from internal variability, and diagnosing the performance of climate models) has never been more imperative. Classical PCA is a well-developed mathematical analysis tool that has been used extensively in climate studies. Its extension in the Fourier frequency domain, the spectral PCA (sPCA), has seen more limited application even though it can potentially better handle dynamical modes of variability thanks to the phase information. We show that the implementation of spectral PCA through the continuous Morlet analytic wavelet transform offers several advantages in terms of simplicity and robustness. In the present work, particular interest is given to the phase of the eigenvectors, which contains the information for making the patterns physically interpretable. Moreover, when several modes of similar amplitude exist within the same frequency band, the rotation of the eigenvectors procedure can help interpret the patterns of the emerging modes. Our proposed criterion for optimal rotation is to look for the rotated eigenvectors having the simplest spatial structure, which is

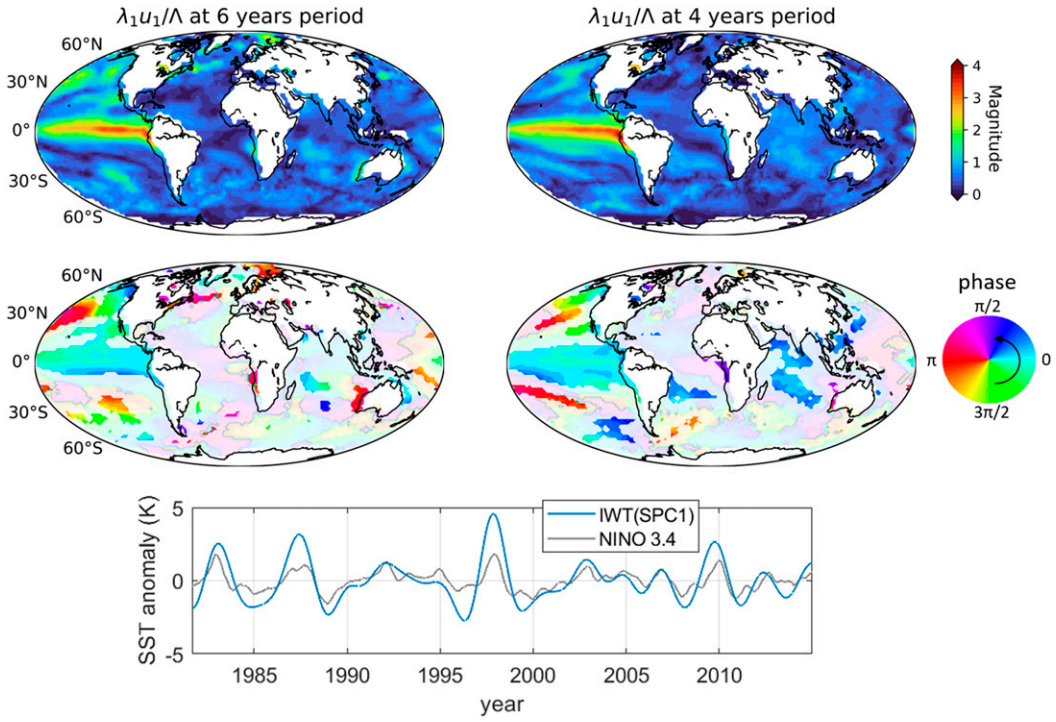

FIG. 16. Results of the sPCA applied to SST daily anomalies at the periods between 2 and 7 years. (top) First unrotated eigenvector (magnitude and phase) at 6- and 4-yr periods. (bottom) Reconstructed signal through the inverse wavelet transform of the first wavelet PC at every frequency corresponding to periods between 2 and 7 years along with the Niño-3.4 index (Deser et al. 2010). The linear correlation between the reconstructed signal and the Niño-3.4 index is 0.83.

achieved by a regularized estimation that minimizes the integral over the studied spatial area of the absolute value (i.e., the L1 norm) of the spatial Laplacian of the rotated vectors. This criterion, termed spatial Laplacian regularization (SLR), is found to be very efficient when applied to a synthetic example with two waves propagating in opposite directions, even in the presence of high variance spatiotemporally correlated noise.

When applied to global 500-mb GPH reanalysis data, the Morlet wavelet–based spectral PCA is able to identify signatures of Rossby waves at periods between 4 and 60 days. The SLR rotation allows to better separate the several components of the Rossby waves and leads to more easily interpreted eigenvectors. When applied to global SST observations, the sPCA can extract the ENSO signal for period between 2 and 7 years. This signal accounts for about 60% of the spectral power in the corresponding frequency band. Even without the rotation procedure, it does well in isolating the ENSO signal from other physical modes (unlike the classical PCA, which mixes the ENSO and climate change signals). This can be attributed both to the frequency localization (two modes of variability operating at different frequencies are naturally separated by the sPCA) and to the fact that the Morlet wavelet is not sensitive to linear trends. At the 15-day period, the wavelet-based sPCA is able to identify the signature of an atmospheric Rossby wave on the SST anomaly time series even if the corresponding spectral PC accounts for only 2% of the spectral power showing the ability of the method to extract low-amplitude propagating waves from a highly variable random signal.

When analyzing the spatial structure of the mapped complex eigenvectors, the argument (phase) is particularly informative. It informs us about the propagation of the wave-type modes at the corresponding frequency, and the phase velocity in

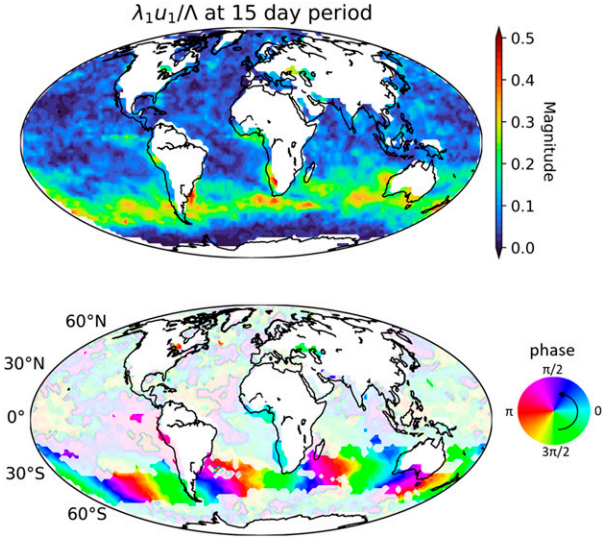

FIG. 17. First unrotated eigenvector (magnitude and phase) at the 15-day period of the sPCA applied to daily SST anomalies. The eastward propagating signal is interpreted as the signature of an atmospheric Rossby wave.

particular can be estimated directly from the mapped eigenvector. The phase information of the sPCA can help understand dynamics and causal relations when classical PCA can only determine correlations. One must note however that the propagation of the phase information does not necessarily correspond to energy or mass flow, particularly when it comes to non-monochromatic dispersive waves, for which phase velocity and group velocity are different (Gill 1982). That is why it is important to analyze the full range of frequencies when non-monochromatic signals are involved.

As the SLR regularization has been shown particularly efficient in improving the interpretability of the patterns of the complex mapped eigenvectors in the present study, it could also be used on real-valued eigenvectors such as those resulting from the classical PCA (e.g., Figs. 11 and 12). Even in the absence of phase information it would impose spatially smooth eigenvectors. For complex eigenvectors, a possible alternative to minimizing the L1 norm of the Laplacian is to minimize only the norm of the Laplacian of the argument of the mapped eigenvector. This would impose smooth and spatially linear phase, while ignoring the local magnitude of the waves. This alternative criterion has been found efficient when the sought waves affect the whole analyzed domain (not shown); however, when some regions of the spatial domain do not show any response to the wave signal, the criterion is negatively affected by the fact that the phase is undefined (random in practice) in those unaffected areas.

The rsPCA method holds a great potential for evaluating and comparing climate model simulations and separate climatic signal from noise by applying the method to ensembles of realizations. We note that the method may be employed using an analytic wavelet different from the Morlet wavelet. While the Morlet wavelet transform has the advantage of being relatively simply related to the Fourier transform, making the Morlet wavelet and Fourier cross-spectral matrices interpretable in a similar way, other wavelets may be more adapted for identifying modes that are not necessarily oscillatory and show irregular periodicity. The proposed rsPCA methodology is expected to benefit from further experimentation and evaluation, which may lead to potential improvements in implementation. For example, an open question concerning this particular rsPCA and rotated PCAs in general, is how to choose the number of PCs to retain when performing rotations (Horel 1981, 1984; Richman 1981, 1986; White et al. 1991).

Acknowledgments. The authors acknowledge support provided by the National Science Foundation (NSF) under the EAGER program (Grant ECCS-1839441) and the TRIPODS+X program (Grant DMS-1839336), as well as by NASA's Global Precipitation Measurement program (Grant 80NSSC19K0684). L.V. was supported by a NASA Earth and Space Science Fellowship (Grant 80NSSC18K1409). Upon request, the data and code that support the findings of this study can be provided by the corresponding author. The support by NSF (Grant EAR-1928724) and NASA (Grant 80NSSC19K0726) for the organization of the 12th International Precipitation Conference (IPC12) and production of the IPC12 special collection of papers is gratefully acknowledged. We also thank the Editor Dr. DelSole and three anonymous reviewers for their constructive comments which improved the presentation of our work.

Data availability statement. The NOAA Optimum Interpolation 1/4 Degree Daily Sea Surface Temperature (OISST) Analysis, version 2 is freely available at https://doi.org/10.7289/V5SQ8XB5. The NCEP–NCAR reanalysis data are freely available at https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html.

APPENDIX A

Notations

Operators are given as follows:

y' denotes the transpose conjugate of y; if y is a complex scalar, y' simply denotes the complex conjugate;
$\tilde{y}$ denotes the empirical estimate of the quantity y;
$\hat{y}$ denotes the Fourier transform of y;
Δ is the Laplacian operator.

Indices are denoted as follows:

n in $(1, \ldots, N)$ is the variable/location index. It also denotes the rank of an eigenvalue and associated eigenvector;
l in $(1, \ldots, L)$ is the observation/time index;
k in $(1, \ldots, K)$ is the scale/frequency index.

Matrix and vector notation is given as follows:

$\mathbf{X}$ is the $N \times L$ data matrix;
$\mathbf{x}_n$ is the nth row of the matrix $\mathbf{X}$;
$\mathbf{C}$ is the $N \times N$ covariance matrix;
$\mathbf{W}_k$ is the $N \times L$ matrix of wavelet coefficients at scale ν_k;
$\mathbf{S}_k$ is the $N \times N$ cross-spectral matrix at frequency f_k;
$\mathbf{u}_n$ is the nth eigenvector of $\tilde{\mathbf{C}}$;
$\mathbf{u}_{n,k}$ is the nth eigenvector of $\widetilde{\mathbf{S}_k}$;
$\mathbf{u}_k^{\text{opt}n}$ is the nth optimally rotated vector;
$\boldsymbol{\kappa}_n$ is the nth principal component time series for the classical PCA;
$\boldsymbol{\kappa}_{n,k}$ is the nth principal component series of wavelet coefficients at scale ν_k of the sPCA.

Variable and parameters are denoted as follows:

λ_n^2 is the nth eigenvalue of $\tilde{\mathbf{C}}$;
$\lambda_{n,k}^2$ is the nth eigenvalue of $\widetilde{\mathbf{S}_k}$;
ν_k is the scale parameter;
f_k is the frequency parameter;
$b_k(f)$ is the frequency band associated to the Morlet wavelet at scale ν_k;
(θ, φ) are the rotation parameters in the space generated by two complex vectors;
$J(\theta, \varphi)$ is the objective function of the rotation;
$(\theta_{\text{opt}}, \varphi_{\text{opt}})$ are the rotation parameters minimizing the objective function;
$a(t)$ and $b(t)$ are the wave signals of the synthetic example;
$\mu_{n_1,n_2}(t)$ is the noise in the synthetic example;

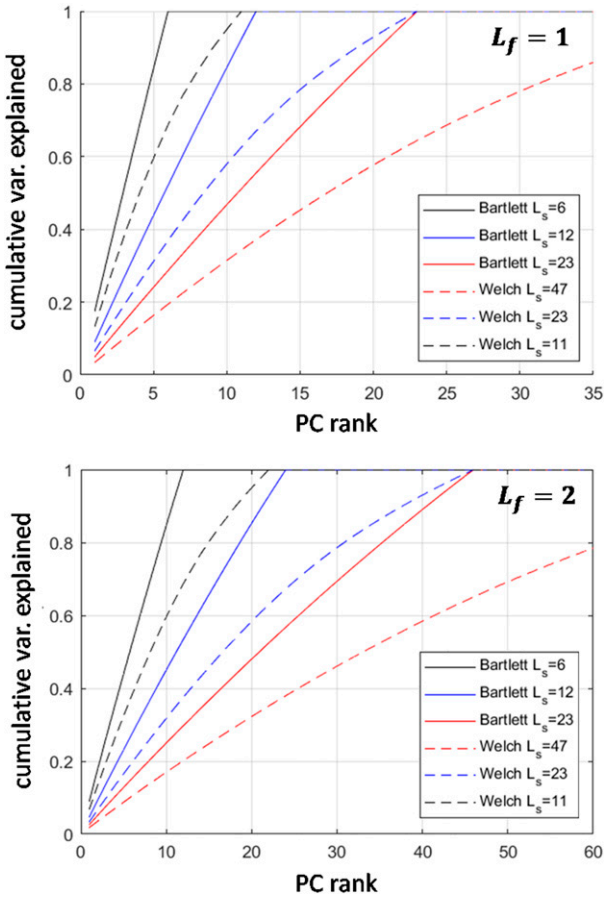

FIG. B1. Cumulative variance explained as a function of the PC rank for periodogram-based spectral PCAs with different windowing setups applied to a system composed of 3960 independent time series (white noise); L_f is the number of discrete frequencies in each frequency band b_k, and L_s is the number of windows of the periodograms. For the Welch periodograms, the windowing function is a Daniell window and the overlapping factor is 50%. Note that for the periodogram-based PCA with independent white noise time series, these results are independent of the frequency.

d_a is the distance to the origin of the wave a in the synthetic example;
d_b is the distance to the origin of the wave b in the synthetic example;
γ_1, γ_2, d_0, and c are the parameters controlling the waves and the noise amplitude and the waves propagation in the synthetic example [Eq. (20)].

APPENDIX B

Rank of Empirical Cross-Spectral Matrix

The empirical Fourier cross-spectral matrix computed directly from the Fourier coefficients over the b_k frequency band is

$$\widetilde{\mathbf{S}_k} = \frac{1}{L_s L_f}\begin{bmatrix} \sum_{f_l\in b_k}\sum_{j=1}^{L_s}\widehat{x_{1,j}}(f_l)\widehat{x_{1,j}}'(f_l) & \cdots & \sum_{f_l\in b_k}\sum_{j=1}^{L_s}\widehat{x_{1,j}}(f_l)\widehat{x_{N,j}}'(f_l) \\ \vdots & \ddots & \vdots \\ \sum_{f_l\in b_k}\sum_{j=1}^{L_s}\widehat{x_{N,j}}(f_l)\widehat{x_{1,j}}'(f_l) & \cdots & \sum_{f_l\in b_k}\sum_{j=1}^{L_s}\widehat{x_{N,j}}(f_l)\widehat{x_{N,j}}'(f_l) \end{bmatrix}, \tag{B1}$$

where $\widehat{x_{1,j}}(f_l)$ is the empirical Fourier coefficient (from a discrete Fourier transform) derived from the jth subsample of the time series $\mathbf{x}_1$ at frequency f_l; L_f is the number of number of discrete frequencies within the b_k frequency band and L_s is the number of subsamples. Each subsample is obtained by windowing or tapering the time series. The prime ($'$) denotes the complex conjugate operator.

In the limit case where $L_f = 1$ and $L_s = 1$ (e.g., when using the discrete Fourier transform without averaging, windowing or tapering in the time domain or in the frequency domain), with $\widehat{x_{1,1}}(f_l) = \alpha_1 e^{i\theta_1}$ we obtain

$$\widetilde{\mathbf{S}_k} = \begin{bmatrix} \alpha_1^2 & \alpha_1\alpha_2 e^{i(\theta_1-\theta_2)} & \cdots & \alpha_1\alpha_N e^{i(\theta_1-\theta_N)} \\ \alpha_2\alpha_1 e^{i(\theta_2-\theta_1)} & \alpha_2^2 & & \\ \vdots & & \ddots & \vdots \\ \alpha_N\alpha_1 e^{i(\theta_N-\theta_1)} & & \cdots & \alpha_N^2 \end{bmatrix}, \tag{B2}$$

which is by construction a rank-1 matrix. More generally, the rank of the empirical cross-spectral matrix is at most of rank $L_s \times L_f$ since from (B1) it can be decomposed as a sum of $L_s \times L_f$ rank-1 matrices. The degree of independence between the subsamples and the adjacent frequency bands also affects the decrease rate of the eigenvalues (see Fig. B1).

APPENDIX C

Fourier Spectrum of the Morlet Wavelet

The Fourier transform of the Morlet mother wavelet Ψ is (Addison 2017)

$$\hat{\Psi}(f) = \pi^{1/4}\sqrt{2}e^{-(2\pi f - 2\pi f_o)^2/2}. \tag{C1}$$

The frequency band b_0 associated with the Morlet mother wavelet Ψ is therefore a Gaussian function centered at the frequency f_o. The Fourier transform of the Morlet daughter wavelet $\Psi_{t,\nu}(u) = 1/\sqrt{\nu}\Psi(u - t/\nu)$ is

$$\widehat{\Psi_{t,\nu}}(f) = \sqrt{\nu}\hat{\Psi}(\nu f)e^{-2i\pi ft}. \tag{C2}$$

The frequency band b_k associated with the Morlet daughter wavelet at scale ν_k is therefore a Gaussian function centered at the frequency $f_k = f_0/\nu_k$.

APPENDIX D

Cone of Influence and Edge Effects

When computing wavelet coefficients from finite length time series, a padding operation is needed to compute the coefficients at the beginning and at the end of the series (Torrence and Compo 1998). Possible solutions are zero-padding, repeating, or mirroring the time series. Repeating and zero-padding are not recommended since they are likely to create a sharp discontinuity (particularly when the time series show a trend). An efficient solution, which is used in the present article, is to pad the series with the values corresponding to its first and last time steps, thus avoiding creating discontinuities. The most conservative option would be to not consider all the coefficients inside the ''cone of influence'', that is, all the coefficients potentially affected by edge effects (e.g., by setting them to zero). However, this would reduce the length of the series of wavelet coefficients available for computing the cross-spectral matrix (particularly at coarse scales/low frequencies).

APPENDIX E

Laplacian Operator and Phase Unwrapping

In image processing the discrete Laplacian operator $\hat{\Delta}$ is defined as a convolution kernel of the following form:

$$\begin{matrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{matrix}. \tag{E1}$$

REFERENCES

Addison, P. S., 2017: *Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*. CRC Press, 446 pp.

Banskota, A., M. J. Falkowski, A. M. Smith, E. S. Kane, K. M. Meingast, L. L. Bourgeau-Chavez, M. E. Miller, and N. H. French, 2017: Continuous wavelet analysis for spectroscopic determination of subsurface moisture and water-table height in Northern peatland ecosystems. *IEEE Trans. Geosci. Remote Sens.*, **55**, 1526–1536, https://doi.org/10.1109/TGRS.2016.2626460.

Bartlett, M. S., 1950: Periodogram analysis and continuous spectra. *Biometrika*, **37**, 1–16, https://doi.org/10.2307/2332141.

Bloomfield, P., and J. M. Davis, 1994: Orthogonal rotation of complex principal components. *Int. J. Climatol.*, **14**, 759–775, https://doi.org/10.1002/joc.3370140706.

Bueso, D., M. Piles, and G. Camps-Valls, 2020: Nonlinear PCA for spatio-temporal analysis of Earth observation data. *IEEE Trans. Geosci. Remote Sens.*, **58**, 5752–5763, https://doi.org/10.1109/TGRS.2020.2969813.

Cai, W., P. H. Whetton, and D. J. Karoly, 2003: The response of the Antarctic Oscillation to increasing and stabilized atmospheric CO_2. *J. Climate*, **16**, 1525–1538, https://doi.org/10.1175/1520-0442-16.10.1525.

Chen, X., and J. M. Wallace, 2016: Orthogonal PDO and ENSO indices. *J. Climate*, **29**, 3883–3892, https://doi.org/10.1175/JCLI-D-15-0684.1.

——, ——, and K. Tung, 2017: Pairwise-rotated EOFs of global SST. *J. Climate*, **30**, 5473–5489, https://doi.org/10.1175/JCLI-D-16-0786.1.

Cottis, R. A., A. M. Homborg, and J. M. C. Mol, 2016: The relationship between spectral and wavelet techniques for noise analysis. *Electrochim. Acta*, **202**, 277–287, https://doi.org/10.1016/j.electacta.2015.11.148.

Deser, C., M. A. Alexander, S.-P. Xie, and A. S. Phillips, 2010: Sea surface temperature variability: Patterns and mechanisms. *Annu. Rev. Mar. Sci.*, **2**, 115–143, https://doi.org/10.1146/annurev-marine-120408-151453.

Di Lorenzo, E., and Coauthors, 2008: North Pacific Gyre Oscillation links ocean climate and ecosystem change. *Geophys. Res. Lett.*, **35**, L08607, https://doi.org/10.1029/2007GL032838.

——, and Coauthors, 2009: Nutrient and salinity decadal variations in the central and eastern North Pacific. *Geophys. Res. Lett.*, **36**, L14601, https://doi.org/10.1029/2009GL038261.

Feldstein, S. B., 2000: The timescale, power spectra, and climate noise properties of teleconnection patterns. *J. Climate*, **13**, 4430–4440, https://doi.org/10.1175/1520-0442(2000)013<4430:TTPSAC>2.0.CO;2.

Gabor, D., 1946: Theory of communication. Part I: The analysis of information. *J. Inst. Electr. Eng.*, **93**, 429–441, https://doi.org/10.1049/ji-3-2.1946.0074.

Georgiou, T. T., and A. Lindquist, 2019: Dynamic relations in sampled processes. *IEEE Control Syst. Lett.*, **3**, 144–149, https://doi.org/10.1109/LCSYS.2018.2859481.

Ghil, M., and Coauthors, 2002: Advanced spectral methods for climatic time series. *Rev. Geophys.*, **40**, 1003, https://doi.org/10.1029/2000RG000092.

Gill, A. E., 1982: Planetary waves. *Atmosphere–Ocean Dynamics*, Academic Press, 500–507.

Hannachi, A., I. T. Jolliffe, and D. B. Stephenson, 2007: Empirical orthogonal functions and related techniques in atmospheric science: A review. *Int. J. Climatol.*, **27**, 1119–1152, https://doi.org/10.1002/joc.1499.

Hasselmann, K., 1988: PIPs and POPs: The reduction of complex dynamical systems using principal interaction and oscillation patterns. *J. Geophys. Res.*, **93**, 11 015–11 021, https://doi.org/10.1029/JD093iD09p11015.

Hirahara, S., M. Ishii, and Y. Fukuda, 2014: Centennial-scale sea surface temperature analysis and its uncertainty. *J. Climate*, **27**, 57–75, https://doi.org/10.1175/JCLI-D-12-00837.1.

Horel, J. D., 1981: A rotated principal component analysis of the interannual variability of the Northern Hemisphere 500 mb height field. *Mon. Wea. Rev.*, **109**, 2080–2092, https://doi.org/10.1175/1520-0493(1981)109<2080:ARPCAO>2.0.CO;2.

——, 1984: Complex principal component analysis: Theory and examples. *J. Appl. Meteor. Climatol.*, **23**, 1660–1673, https://doi.org/10.1175/1520-0450(1984)023<1660:CPCATA>2.0.CO;2.

Hsieh, W. W., 2004: Nonlinear multivariate and time series analysis by neural network methods. *Rev. Geophys.*, **42**, RG1003, https://doi.org/10.1029/2002RG000112.

Hudgins, L., C. A. Friehe, and M. E. Mayer, 1993: Wavelet transforms and atmospheric turbulence. *Phys. Rev. Lett.*, **71**, 3279–3282, https://doi.org/10.1103/PhysRevLett.71.3279.

Hurrell, J. W., Y. Kushnir, G. Ottersen, and M. Visbeck, 2013: An overview of the North Atlantic Oscillation. *The North Atlantic Oscillation: Climatic Significance and Environmental Impact*, *Geophys. Monogr.*, Vol. 134, Amer. Geophys. Union, 1–35.

Jiang, X., and S. Mahadevan, 2011: Wavelet spectrum analysis approach to model validation of dynamic systems. *Mech. Syst. Signal Process.*, **25**, 575–590, https://doi.org/10.1016/j.ymssp.2010.05.012.

Johnson, E. S., and M. J. Mc Phaden, 1993: Structure of intraseasonal Kelvin waves in the equatorial Pacific Ocean. *J. Phys. Oceanogr.*, **23**, 608–625, https://doi.org/10.1175/1520-0485(1993)023<0608:SOIKWI>2.0.CO;2.

Jolliffe, I. T., 1986: Principal components in regression analysis. *Principal Component Analysis*, Springer-Verlag, 129–155.

Kaiser, H. F., 1958: The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187–200, https://doi.org/10.1007/BF02289233.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–472, https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Keiner, L. E., and X. H. Yan, 1997: Empirical orthogonal function analysis of sea surface temperature patterns in Delaware Bay. *IEEE Trans. Geosci. Remote Sens.*, **35**, 1299–1306, https://doi.org/10.1109/36.628796.

Kessler, W. S., 2001: EOF representations of the Madden–Julian oscillation and its connection with ENSO. *J. Climate*, **14**, 3055–3061, https://doi.org/10.1175/1520-0442(2001)014<3055:EROTMJ>2.0.CO;2.

Kidson, J. W., 1999: Principal modes of Southern Hemisphere low-frequency variability obtained from NCEP–NCAR reanalyses. *J. Climate*, **12**, 2808–2830, https://doi.org/10.1175/1520-0442(1999)012<2808:PMOSHL>2.0.CO;2.

Kirby, J. F., 2005: Which wavelet best reproduces the Fourier power spectrum? *Comput. Geosci.*, **31**, 846–864, https://doi.org/10.1016/j.cageo.2005.01.014.

Kumar, P., and E. Foufoula-Georgiou, 1997: Wavelet analysis for geophysical applications. *Rev. Geophys.*, **35**, 385–412, https://doi.org/10.1029/97RG00427.

Leathers, D. J., B. Yarnal, and M. A. Palecki, 1991: The Pacific/North American teleconnection pattern and United States climate. Part I: Regional temperature and precipitation associations. *J. Climate*, **4**, 517–528, https://doi.org/10.1175/1520-0442(1991)004<0517:TPATPA>2.0.CO;2.

Lian, T., and D. Chen, 2012: An evaluation of rotated EOF analysis and its application to tropical Pacific SST variability. *J. Climate*, **25**, 5361–5373, https://doi.org/10.1175/JCLI-D-11-00663.1.

Lilly, J. M., and J. Park, 1995: Multiwavelet spectral and polarization analyses of seismic records. *Geophys. J. Int.*, **122**, 1001–1021, https://doi.org/10.1111/j.1365-246X.1995.tb06852.x.

Lorenz, E. N., 1956: Empirical orthogonal functions and statistical weather prediction. Massachusetts Institute of Technology Science Rep. 1, 49 pp.

Madden, R. A., and P. R. Julian, 1971: Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28**, 702–708, https://doi.org/10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2.

Mann, M. E., and J. Park, 1994: Global-scale modes of surface temperature variability on interannual to century timescales. *J. Geophys. Res.*, **99**, 25 819–25 833, https://doi.org/10.1029/94JD02396.

——, and ——, 1999: Oscillatory spatiotemporal signal detection in climate studies: A multiple-taper spectral domain approach. *Advances in Geophysics*, Vol. 41, Academic Press, 1–131, https://doi.org/10.1016/S0065-2687(08)60026-6.

——, B. A. Steinman, and S. K. Miller, 2020: Absence of internal multidecadal and interdecadal oscillations in climate model simulations. *Nat. Commun.*, **11**, 49, https://doi.org/10.1038/s41467-019-13823-w.

Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and C. R. Francis, 1997: A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.*, **78**, 1069–1079, https://doi.org/10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2.

Messié, M., and F. Chavez, 2011: Global modes of sea surface temperature variability in relation to regional climate indices. *J. Climate*, **24**, 4314–4331, https://doi.org/10.1175/2011JCLI3941.1.

Mestas-Nuñez, A. M., and D. B. Enfield, 1999: Rotated global modes of non-ENSO sea surface temperature variability. *J. Climate*, **12**, 2734–2746, https://doi.org/10.1175/1520-0442(1999)012<2734:RGMONE>2.0.CO;2.

Morlet, J., G. Arens, E. Fourgeau, and D. Giard, 1982: Wave propagation and sampling theory—Part II: Sampling theory and complex waves. *Geophysics*, **47**, 222–236, https://doi.org/10.1190/1.1441329.

Navarra, A., and V. Simoncini, 2010: *A Guide to Empirical Orthogonal Functions for Climate Data Analysis*. Springer Science & Business Media, 151 pp.

Newman, M., and Coauthors, 2016: The Pacific decadal oscillation, revisited. *J. Climate*, **29**, 4399–4427, https://doi.org/10.1175/JCLI-D-15-0508.1.

North, G. R., T. L. Bell, and R. F. Cahalan, 1982: Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.*, **110**, 699–706, https://doi.org/10.1175/1520-0493(1982)110<0699:SEITEO>2.0.CO;2.

Park, J., and M. E. Mann, 2000: Interannual temperature events and shifts in global temperature: A ''multiwavelet'' correlation approach. *Earth Interact.*, **4**, https://doi.org/10.1175/1087-3562(2000)004<0001:ITEASI>2.3.CO;2.

Perrier, V., T. Philipovitch, and C. Basdevant, 1995: Wavelet spectra compared to Fourier spectra. *J. Math. Phys.*, **36**, 1506–1519, https://doi.org/10.1063/1.531340.

Picci, G., and S. Pinzoni, 1986: Dynamic factor-analysis models for stationary processes. *J. Math. Control Inf.*, **3**, 185–210, https://doi.org/10.1093/imamci/3.2-3.185.

Platzman, G. W., 1968: The Rossby wave. *Quart. J. Roy. Meteor. Soc.*, **94**, 225–248, https://doi.org/10.1002/qj.49709440102.

Power, S., F. Tseitkin, V. Mehta, B. Lavery, S. Torok, and N. Holbrook, 1999: Decadal climate variability in Australia during the twentieth century. *Int. J. Climatol.*, **19**, 169–184, https://doi.org/10.1002/(SICI)1097-0088(199902)19:2<169:AID-JOC356>3.0.CO;2-Y.

Proakis, J. G., 2001: Nonparametric methods for power spectrum estimation. *Digital Signal Processing: Principles, Algorithms and Applications*, Pearson Education India, 908–919.

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution-blended analyses for sea surface temperature. *J. Climate*, **20**, 5473–5496, https://doi.org/10.1175/2007JCLI1824.1.

Richman, M. B., 1981: Obliquely rotated principal components: An improved meteorological map typing technique? *J. Appl. Meteor.*, **20**, 1145–1159, https://doi.org/10.1175/1520-0450(1981)020<1145:ORPCAI>2.0.CO;2.

——, 1986: Rotation of principal components. *J. Climatol.*, **6**, 293–335, https://doi.org/10.1002/joc.3370060305.

Rossby, C. G., 1945: On the propagation of frequencies and energy in certain types of oceanic and atmospheric waves. *J. Meteor.*, **2**, 187–204, https://doi.org/10.1175/1520-0469(1945)002<0187:OTPOFA>2.0.CO;2.

Roundy, P. E., 2015: On the interpretation of EOF analysis of ENSO, atmospheric Kelvin waves, and the MJO. *J. Climate*, **28**, 1148–1165, https://doi.org/10.1175/JCLI-D-14-00398.1.

——, and C. J. Schreck, 2009: A combined wave-number–frequency and time-extended EOF approach for tracking

the progress of modes of large-scale organized tropical convection. *Quart. J. Roy. Meteor. Soc.*, **135**, 161–173, https://doi.org/10.1002/qj.356.

Shindell, D. T., and G. A. Schmidt, 2004: Southern Hemisphere climate response to ozone changes and greenhouse gas increases. *Geophys. Res. Lett.*, **31**, L18209, https://doi.org/10.1029/2004GL020724.

Snyder, J. P., 1977: A comparison of pseudocylindrical map projections. *Amer. Cartogr.*, **4**, 59–81, https://doi.org/10.1559/152304077784080031.

Thompson, D. W. J., and J. M. Wallace, 1998: The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.*, **25**, 1297–1300, https://doi.org/10.1029/98GL00950.

——, and ——, 2000: Annular modes in the extratropical circulation. Part I: Month-to-month variability. *J. Climate*, **13**, 1000–1016, https://doi.org/10.1175/1520-0442(2000)013<1000:AMITEC>2.0.CO;2.

Thornhill, N. F., S. L. Shah, B. Huang, and A. Vishnubhotla, 2002: Spectral principal component analysis of dynamic process data. *Control Eng. Pract.*, **10**, 833–846, https://doi.org/10.1016/S0967-0661(02)00035-7.

Torrence, C., and G. P. Compo, 1998: A practical guide to wavelet analysis. *Bull. Amer. Meteor. Soc.*, **79**, 61–78, https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2.

Trenberth, K. E., 1997: The definition of El Niño. *Bull. Amer. Meteor. Soc.*, **78**, 2771–2778, https://doi.org/10.1175/1520-0477(1997)078<2771:TDOENO>2.0.CO;2.

——, and D. A. Paolino, 1981: Characteristic patterns of variability of sea level pressure in the Northern Hemisphere. *Mon. Wea. Rev.*, **109**, 1169–1189, https://doi.org/10.1175/1520-0493(1981)109<1169:CPOVOS>2.0.CO;2.

——, and J. W. Hurrell, 1994: Decadal atmosphere–ocean variations in the Pacific. *Climate Dyn.*, **9**, 303–319, https://doi.org/10.1007/BF00204745.

Vetterli, M., and J. Kovacevic, 1995: *Wavelets and Subband Coding*. Prentice-Hall, 488 pp.

Wallace, J. M., 1971: Spectral studies of tropospheric wave disturbances in the tropical western Pacific. *Rev. Geophys.*, **9**, 557–612, https://doi.org/10.1029/RG009i003p00557.

——, 1972: Empirical orthogonal representation of time series in the frequency domain. Part II: Application to the study of tropical wave disturbances. *J. Appl. Meteor.*, **11**, 893–900, https://doi.org/10.1175/1520-0450(1972)011<0893:EOROTS>2.0.CO;2.

——, and R. E. Dickinson, 1972: Empirical orthogonal representation of time series in the frequency domain. Part I: Theoretical considerations. *J. Appl. Meteor.*, **11**, 887–892, https://doi.org/10.1175/1520-0450(1972)011<0887:EOROTS>2.0.CO;2.

——, and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.*, **109**, 784–812, https://doi.org/10.1175/1520-0493(1981)109<0784:TITGHF>2.0.CO;2.

——, Y. Zhang, and K. H. Lau, 1993: Structure and seasonality of interannual and interdecadal variability of the geopotential height and temperature fields in the Northern Hemisphere troposphere. *J. Climate*, **6**, 2063–2082, https://doi.org/10.1175/1520-0442(1993)006<2063:SASOIA>2.0.CO;2.

Wang, C., C. Deser, J. Y. Yu, P. DiNezio, and A. Clement, 2017: El Niño and Southern Oscillation (ENSO): A review. *Coral Reefs of the Eastern Tropical Pacific*, P. Glymn, D. Manzello, and I. Enochs, Eds., Springer, 85–106.

Weare, B. C., A. R. Navato, and R. E. Newell, 1976: Empirical orthogonal analysis of Pacific sea surface temperatures. *J. Phys. Oceanogr.*, **6**, 671–678, https://doi.org/10.1175/1520-0485(1976)006<0671:EOAOPS>2.0.CO;2.

Welch, P., 1967: The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.*, **15**, 70–73, https://doi.org/10.1109/TAU.1967.1161901.

Wheeler, M. C., and H. H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, **132**, 1917–1932, https://doi.org/10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2.

White, D., M. Richman, and B. Yarnal, 1991: Climate regionalization and rotation of principal components. *Int. J. Climatol.*, **11**, 1–25, https://doi.org/10.1002/joc.3370110102.

Wills, R. C., T. Schneider, J. M. Wallace, D. S. Battisti, and D. L. Hartmann, 2018: Disentangling global warming, multidecadal variability, and El Niño in Pacific temperatures. *Geophys. Res. Lett.*, **45**, 2487–2496, https://doi.org/10.1002/2017GL076327.