Data-adaptive harmonic analysis of oceanic waves and turbulent flows

D. Kondrashov,[1, a] E. A. Ryzhov,[2, 3, b] and P. Berloff[2, c]

[1]*Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, 90095, USA*

[2]*Department of Mathematics, Imperial College London, London, SW7 2AZ, UK*

[3]*Pacific Oceanological Institute, Vladivostok, 690041, Russia*

(Dated: 20 May 2020)

We introduce new features of data-adaptive harmonic decomposition (DAHD) that are showcased to characterize spatiotemporal variability in high-dimensional datasets of complex and mutsicale oceanic flows, offering new perspectives and novel insights. First, we present didactic example with synthetic data for identification of coherent oceanic waves embedded in high amplitude noise. Then DAHD is applied to analyze turbulent oceanic flows simulated by Regional Oceanic Modeling System (ROMS) and an eddy-resolving three-layer quasigeostrophic (QG) ocean model, where resulting spectra exhibit a thin line capturing nearly all the energy at a given temporal frequency, and showing well defined scaling behavior across frequencies. DAHD thus permits sparse representation of complex, multiscale and chaotic dynamics by a relatively few data-inferred spatial patterns evolving with simple temporal dynamics, namely oscillating harmonically in time at a given single frequency. The detection of this low-rank behavior is facilitated by an eigendecomposition of Hermitian cross-spectral matrix and resulting eigenvectors that represent an orthonormal set of global spatiotemporal modes associated with specific temporal frequency, which in turn allows to rank these modes by their captured energy and across frequencies, and allow accurate space-time reconstruction. Furthermore, by using a correlogram estimator of the Hermitian cross-spectral density matrix, DAHD is both closely related and distinctly different from the spectral proper orthogonal decomposition (SPOD) that relies on Welch's periodogram as its estimator method.

---

[a]Electronic mail: dkondras@atmos.ucla.edu

[b]Electronic mail: ryzhovea@gmail.com

[c]Electronic mail: p.berloff@imperial.ac.uk

The turbulent oceanic flows consist of ubiquitous complex motions – jets, vortices and waves that co-exist on very different spatiotemporal scales but also without clear scale separation, and it brings natural challenge to characterize the whole complexity across the scales. In particular, the study of temporal scales has got less attention than of spatial scales. To that effect we offer fresh perspectives and novel insights by introducing new features of data-adaptive harmonic decomposition (DAHD) that are applied to analyze complex high-dimensional spatiotemporal datasets of oceanic flows, including synthetic example of identifying coherent oceanic waves embedded in high-amplitude noise, and turbulent flows simulated by a hierarchy of oceanic models. DAHD results reveal striking low-rank behavior and sparse representation of complex, multiscale and chaotic flows by a relatively few data-inferred spatial patterns evolving with simple temporal dynamics, as well as well-defined scaling behavior across temporal frequencies, such as exponential-like shape and power law.

I. INTRODUCTION

Over past decade, approaches based on Dynamical Mode Decomposition (DMD) and Koopman analysis, have quickly gained popularity to analyze datasets in the engineering fluids community (Schmid, 2010; Budišić, Mohr, and Mezić, 2012; Williams, Kevrekidis, and Rowley, 2015). DMD computes from the data the eigenvalues and eigenvectors of a linear model that approximates the underlying nonlinear dynamics. Tu *et al.* (2014) has established close connection of DMD with independently developed in climate science Principal Orthogonal Patterns (Penland, 1989) and Linear Inverse Modeling (Penland, 1996), that have been since generalized to include memory and nonlinear effects (Kondrashov, Chekroun, and Ghil, 2015; Mukhin *et al.*, 2015; Gavrilov *et al.*, 2016), as well as state-dependent noise (Kravtsov *et al.*, 2016; Martinez-Villalobos *et al.*, 2018).

Spectral proper orthogonal decomposition (SPOD) (Towne, Schmidt, and Colonius, 2018) and spectral empirical orthogonal function analysis (SEOF) (Schmidt *et al.*, 2019)), are based on an eigenvalue decomposition of the estimated cross-spectral density matrix, yielding at each frequency a set of time-harmonic and orthogonal modes; these SPOD/SEOF modes can be interpreted, in turn, as ensemble DMD modes. Data-adaptive harmonic decomposition (DAHD) (Chekroun and Kondrashov, 2017; Kondrashov, Chekroun, and Berloff, 2018) is based on spectral analysis of integral shift operator with two-point statistics kernel built from time-lagged cross-correlations, and it is realized numerically by an eigendecomposition that is briefly reviewed in Sec II A. Similar to SPOD/SEOF methods, DAHD is yielding orthogonal set of eigenvectors – data-adaptive harmonic modes (DAHMs) oscillating harmonically in time.

Furthermore, DAHD enables effective inverse modeling of the original dataset by a system of frequency-ranked nonlinear stochastic oscillators with memory effects, which has been successfully applied to challenging datasets across sciences, including oceanic turbulence (Kondrashov, Chekroun, and Berloff, 2018), Arctic sea ice (Kondrashov *et al.*, 2018; Kondrashov, Chekroun, and Ghil, 2018) and space physics (Kondrashov and Chekroun, 2018).

This study introduces important DAHD modifications and new features, namely an eigendecomposition of Hermitian cross-spectral density matrix in frequency-domain (Sec. II B) and auxiliary energy spectrum (Sec. II C) obtained by using judicious projection of the data onto DAHMs. DAHD connection and difference with SPOD/SEOF is established (Sec. II B).

Hermitian form of DAHD is applied first for identification of coherent oceanic waves in noisy synthetic data (Sec. III), and then to analyze several high-dimensional datasets of complex turbulent geophysical flows, namely Regional Oceanic Modeling System (ROMS) simulation of equatorial region (Sec. IV), and wind-driven gyres circulation by eddy-resolving three-layer quasigeostrophic ocean model (Sec. V). Discussion and conclusions follow in Sec VI.

II. DATA-ADAPTIVE HARMONIC DECOMPOSITION

A. Time-domain formulation

We consider a multivariate time series $\boldsymbol{X}(n) = (X_1(n), \ldots, X_d(n))$ formed with d spatial channels and $n = 1, \ldots, N$ time points (sampled evenly). Double-sided (unbiased) cross-correlation coefficients $\rho^{(p,q)}(m)$ are estimated for all pairs of channels p and q and time lag m up to a maximum $M-1$:

$$\rho^{(p,q)}(m) = \begin{cases} \frac{1}{N-m} \sum_{n=1}^{N-m} X_p(n+m) X_q(n), 0 \le m \le M-1, \\ \rho^{(q,p)}(-m), m < 0. \end{cases} \tag{1}$$

Here M is the embedding window and its size should be larger than typical decorrelation times in the data; it is the slowest temporal scale captured by DAHD. To reduce the biases of cross-correlations estimates, common lag windowing can be applied.

Combining all the lagged cross-correlation coefficients between a given pair of channels p and q

leads to a cross-correlation Hankel matrix $\boldsymbol{H}^{(p,q)}$, which is symmetric and obtained by a left shift of the row vector $(\rho^{(p,q)}_{-M+1}, \ldots, \rho^{(p,q)}_{0}, \ldots, \rho^{(p,q)}_{M-1})$, so every anti-diagonal consists of the same elements. Block-matrix $\boldsymbol{\mathcal{C}}$ of size $d(2M-1) \cdot d(2M-1)$ is made *symmetric by construction* and obtained by arranging all the Hankel matrices $\boldsymbol{H}^{(p,q)}$ for each pair of channels (p, q):

$$\begin{aligned} \boldsymbol{\mathcal{C}}^{(p,q)} &= \boldsymbol{H}^{(p,q)}, p \leq q, \\ \boldsymbol{\mathcal{C}}^{(p,q)} &= \boldsymbol{H}^{(q,p)}, \text{elsewhere} \end{aligned} \tag{2}$$

An important property of $\boldsymbol{\mathcal{C}}$ is that its eigenvalues λ come in pairs of opposite values, while the eigenvectors $\mathbf{W}_j = (\mathbf{E}^j_1, \ldots, \mathbf{E}^j_d)$ represent collection of global space-time patterns, namely data-adaptive harmonic modes (DAHMs) oscillating at a single temporal frequency. In particular, $\mathbf{E}^j_k$ is M'-long time series that corresponds to harmonic oscillation with a frequency f:

$$\mathbf{E}^j_k(s) = B^j_k \cos(2\pi f s + \theta^j_k), \quad 1 \leq s \leq M'; 1 \leq k \leq d, \tag{3}$$

where the amplitudes B^j_k and phases θ^j_k are data-adaptive, while frequency f is equally spaced in Nyquist interval [0 0.5] with M values:

$$f = \frac{(\ell - 1)}{M' - 1}, \quad \ell = 1, \ldots, \frac{M' + 1}{2} \tag{4}$$

In total, $j = 1, ..., d(2M-1)$ spectral DAHD eigenelements $(|\lambda_j|, \mathbf{W}_j)$ are computed, and DAHD spectrum is obtained by plotting eigenvalues $|\lambda|$ according to their frequency f.

Several DAHD studies (Chekroun and Kondrashov, 2017; Kondrashov, Chekroun, and Berloff, 2018) have relied on eigendecomposition of $\boldsymbol{\mathcal{C}}$ (Eq. 2) to perform the analysis on relatively low-dimensional datasets, i.e. d not exceeding 40. High-dimensional datasets can be compressed first by PCA, aiming to retain nearly all of the variance (98–99%) in leading PC's, which are then analyzed. In practice, DAHD results are invariant to the orthogonal rotation of the data and to PCA in particular. Still, even after PCA compression, matrix $\boldsymbol{\mathcal{C}}$ can be very large and its direct eigendecomposition is computationally prohibitive. For such cases, Ryzhov *et al.* (2019, 2020) have utilized frequency-domain DAHD formulation, described next.

B. Frequency-domain formulation

Theorem V.1 of (Chekroun and Kondrashov, 2017) has established that DAHD eigenvalues are related to singular values of *symmeterized complex* $d \times d$ cross-spectral matrix $\mathfrak{S}(f)$ whose elements are given by:

$$\mathfrak{S}_{p,q}(f) = \begin{cases} \widehat{\rho^{p,q}}(f) \text{ if } q \geq p, \\ \widehat{\rho^{q,p}}(f) \text{ if } q < p \end{cases} \tag{5}$$

where $\widehat{\rho^{p,q}}(f)$ is the Fourier transform at the frequency f of the cross-correlation sequence $\rho^{p,q}(m)$:

$$\widehat{\rho^{p,q}}(f) = \sum_{m=-M+1}^{M-1} \rho^{p,q}(m) e^{-2\pi i f} \tag{6}$$

In particular, for each singular value $\sigma_k(f)$ of $\mathfrak{S}(f)$ there exists, when $f \neq 0$, a pair of negative-positive DAHD eigenvalues $(\lambda_k^+(f), \lambda_k^-(f))$ of $\boldsymbol{\mathcal{C}}$ such that:

$$\lambda_k^+(f) = -\lambda_k^-(f) = \sigma_k(f), \quad 1 \leq k \leq d, \tag{7}$$

i.e., $2d$ eigenvalues are associated with each Fourier frequency $f \neq 0$, while respective DAHMs are shifted by the quarter of the period, i.e. $\theta_k^+ = \theta_k^- + \pi/2$. There are only d (not paired) eigenvalues for frequency $f = 0$.

As a point of departure from DAHD formulation presented above, it is important to note that matrix $\mathfrak{S}(f)$ in Eq. 5 is symmetrized by intent for practical numerical reasons, because $\widehat{\rho^{p,q}}(f)$ and $\widehat{\rho^{q,p}}(f)$ are generally different when $p \neq q$. However, $\widehat{\rho^{p,q}}(f)$ and $\widehat{\rho^{q,p}}(f)$ are necessarily related to each other due to the reversal property of estimated time-lagged cross-correlation sequences, namely:

$$\rho^{p,q}(m) = \rho^{q,p}(-m), \tag{8}$$

where $-M + 1 \leq m \leq M - 1$. This is due to the fact that for $\rho^{p,q}(m)$, channel p is leading channel q by m lags, and for $\rho^{q,p}(m)$ it is the opposite, see Eq. 1. Equivalently, Eq. 8 can be written also as

$$\rho^{p,q}(2M - m) = \rho^{q,p}(m), \tag{9}$$

where $1 \leq m \leq 2M - 1$. It is easy to show that Fourier transforms of $\rho^{q,p}(m)$ and $\rho^{p,q}(m)$ sequences are then related by the frequency-dependent phase shift:

$$\widehat{\rho^{p,q}}(f) = e^{i\phi} \overline{\widehat{\rho^{q,p}}}(f) \tag{10}$$

where $\overline{\widehat{\rho^{q,p}}}$ is complex conjugate, $f = 0.5\frac{(l-1)}{M-1}$, $\phi = \frac{2\pi(l-1)}{2M-1}$, $1 \leq l \leq M$. Furthermore, since Eq's (8–10) holds when $p = q$, we obtain for the diagonal elements of $\mathfrak{S}(f)$:

$$\mathfrak{S}_{p,p}(f) = e^{i\phi/2} |\widehat{\rho^{p,p}}(f)| \tag{11}$$

We can thus define cross-spectral density matrix $\boldsymbol{S}(f)$:

$$\boldsymbol{S}(f) = e^{-i\phi/2}\mathfrak{S}(f), \tag{12}$$

where $\mathfrak{S}(f)$ is *non-symmetrized*, i.e. $\mathfrak{S}_{p,q}(f) = \widehat{\rho^{p,q}}(f)$ for $1 \leq p, q \leq d$. Matrix $\boldsymbol{S}$ is Hermitian, i.e. and $\boldsymbol{S}$=$\boldsymbol{S}^*$=$\overline{\boldsymbol{S}^T}$ by the way of equations (10–11), and it is interpreted as a correlogram estimate of multidimensional Hermitian cross-spectral density $E\{\widehat{\boldsymbol{X}(f)}\widehat{\boldsymbol{X}(f)}^*\}$, where an expectation operator E is an ensemble average over different realizations of multidimensional random process $\boldsymbol{X}$. The phase-shift factor $e^{-i\phi/2}$ in Eq. 11 accounts for discrete case estimation. For univariate case in particular (i.e. $p = q = 1$), $S(f) = |\widehat{\rho(f)}|$ represents a well established correlogram method based on Wiener-Khinchin theorem, which states that the power spectral density is equal to the Fourier transform of its autocorrelation function (Percival and Walden, 1993). Furthermore, since $\boldsymbol{S}(f)$ is a Hermitian matrix, its eigendecomposition:

$$\boldsymbol{S}(f) = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^*, \tag{13}$$

yields set of d real eigenvalues $\Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_\mathrm{d})$ (their absolute values are used for the analysis and results presented), while associated d eigenvectors $\boldsymbol{U}$ form an orthonormal set at each frequency f in the Nyquist interval (see Eq. 10).

The Hermitian DAHD formulation (Eqs.11–13) thus closely resembles SPOD/SEOF approach (Towne, Schmidt, and Colonius, 2018; Schmidt *et al.*, 2019), with the key difference is that the latter considers Welch's overlapped averaging periodogram estimate of cross-spectral density matrix, while the former is based on correlogram estimate (Eq. 11). However, same as in SPOD, Hermitian DAHD eigenvectors $\boldsymbol{U}_k$ are associated with the Fourier transform of spatial wave which is harmonic in time at single frequency f, i.e. the time-domain DAHMs $\boldsymbol{W}_k$ (see Eq. 3) have the following representation:

$$\widehat{\boldsymbol{W}_{\boldsymbol{k}}}(f) = \boldsymbol{U}_{\boldsymbol{k}}(f), \tag{14}$$

For the purpose of reconstruction in time-domain and computing energy spectrum (see Sec. II C), for each λ_k, pairs $(\widehat{\boldsymbol{W}_{\boldsymbol{k}}^{+}}(f) = \boldsymbol{U}_{\boldsymbol{k}}(f), \widehat{\boldsymbol{W}_{\boldsymbol{k}}^{-}}(f) = e^{i\pi/2}\boldsymbol{U}_{\boldsymbol{k}}(f))$ are formed. In turn, $\boldsymbol{W}_{\boldsymbol{k}}^{+}(f)$ and $\boldsymbol{W}_{\boldsymbol{k}}^{-}(f)$ are obtained by inverse Fourier transform to yield space-time patterns at a single temporal frequency f and shifted by a quarter of a period (aka *sin* and *cos*), and thus provide an orthonormal basis set of modes $\mathbf{W}$ in time-domain across all frequencies, i.e. $\mathbf{W}^T\mathbf{W} = \mathbf{I}$. By analogy with SPOD/SEOF modes, DAHD modes can be interpreted then as optimal response modes of the forced linear system if the forcing is white in space and time (Towne, Schmidt, and Colonius, 2018; Schmidt *et al.*, 2019).

The frequency-domain approach is fully parallelizable and thus is very computationally efficient, since the eigendecomposition of each frequency can be performed in parallel.

C. Energy spectrum and space-time reconstruction

The maximum resolution in frequency (see Eq. 4) is achieved when $M = (N+1)/2$, and in this case DAHMs attain size Nxd, i.e. they have exactly same dimensions as the original data $\boldsymbol{X}$ in time and space. Since set of DAHM's is an orthonormal basis, i.e. $\mathbf{W}^T\mathbf{W} = \mathbf{I}$, it can be used to perform the following expansion of $\boldsymbol{X}$:

$$\mathbf{X} = \sum_{j=1}^{dN} \mu_j \mathbf{W_j} \tag{15}$$

Here, μ_j's are scalar weights – Data-Adaptive Harmonic coefficients (DAHCs) that are uniquely defined by projecting $\boldsymbol{X}$ onto DAHMs, namely:

$$\mu_j = \boldsymbol{W_j^T X} = \sum_{n=1}^{N}\sum_{k=1}^{d} E_k^j(n) X_k(n), \tag{16}$$

The original data can be then partially (or fully) reconstructed by selecting subset (whole set) of μ's and associated DAHMs in Eq (15). The physical meaning of DAHCs readily follows, namely when squared, they establish relative contributions of respective DAHMs to L^2 energy of $\mathbf{X}$, i.e. with above definitions:

$$\sum_{n=1}^{N}\sum_{k=1}^{d} X_k^2(n) = \sum_{j=1}^{Nd} \mu_j^2 \tag{17}$$

Because of pairing in DAHD eigenvectors and their association with a single temporal frequency f(see Sec. II B and Eq. 14), it is useful to isolate in the sum on right hand side of Eq. (17) energy contribution for any given pair:

$$\alpha_k^2(f) = (\mu_k^+(f))^2 + (\mu_k^-(f))^2;\ 1 \le k \le d, \tag{18}$$

The DAHD energy spectrum is then obtained by plotting collection of $\alpha^2(f)$ as a function of their frequency, similar to plot of $\lambda(f)$'s.

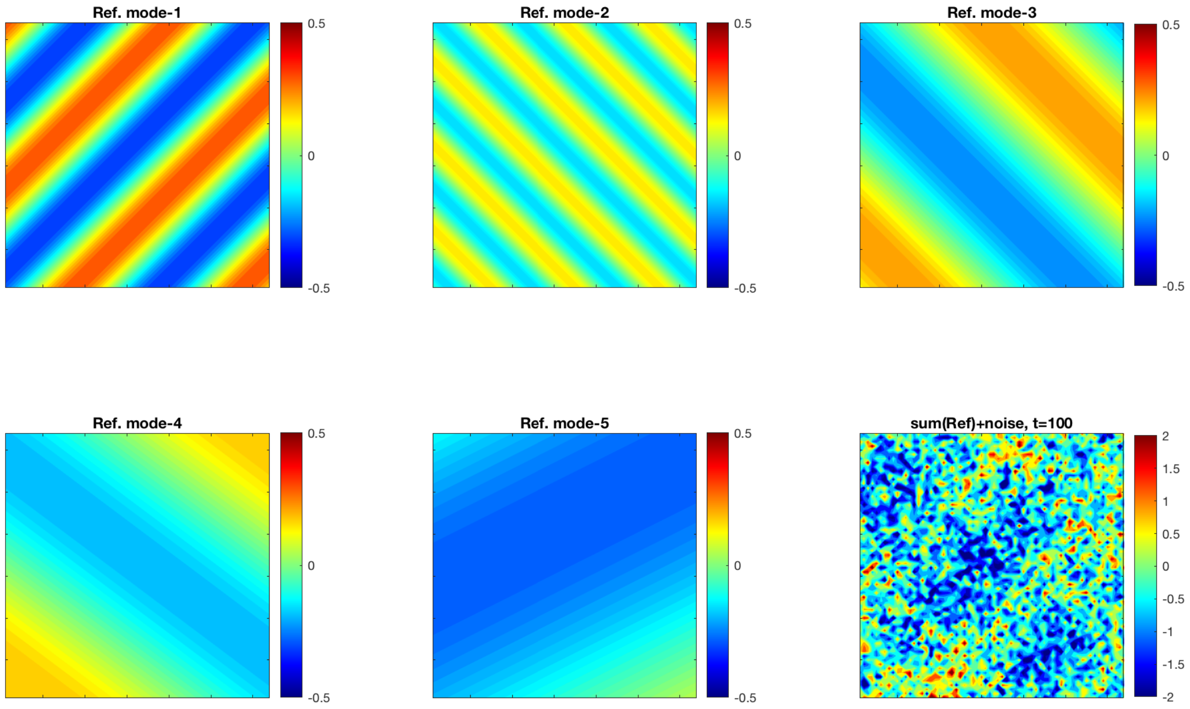


FIG. 1: Temporal snapshots of 2-D reference waves and their sum contaminated by noise. (Multimedia view).

III. IDENTIFICATION OF ROSSBY WAVES

We consider here a synthetic example of several propagating 2-D waves:

$$u_n(x,y,t) = A_n cos(k_n x/L_x + l_n y/L_y + \omega_n t) \tag{19}$$

where A_n are the weights, while frequency ω_n and wave numbers (k_n, l_n) obey Rossby dispersion relation:

$$\omega_n = -\frac{\beta k_n}{k_n^2 + l_n^2 + R^{-2}} \tag{20}$$

The full dataset $\mathbf{v}(x,y,t)$ consists of a coherent component $\mathbf{s}(x,y,t) = \sum_{n=1}^{K} u_n(x,y,t)$ embedded in a *temporally correlated, and spatially uncorrelated* noise $\mathbf{r}(x,y,t)$ represented locally by independent AR(1) processes and randomly chosen AR(1) cofficient:

$$\mathbf{v}(x,y,t) = (1-\nu)^{1/2}\,\mathbf{s}(x,y,t) + \nu^{1/2}\,\mathbf{r}(x,y,t)\,. \tag{21}$$

For this example we choose $L_x = L_y = 64$ and uniform spacing in x and y with $N_x = N_y = 64$ of grid points; $R = 5$, $\beta = -2 \cdot 10^{-7}$, and $K = 5$ waves with the spatial wave numbers $k_1 = -10$;

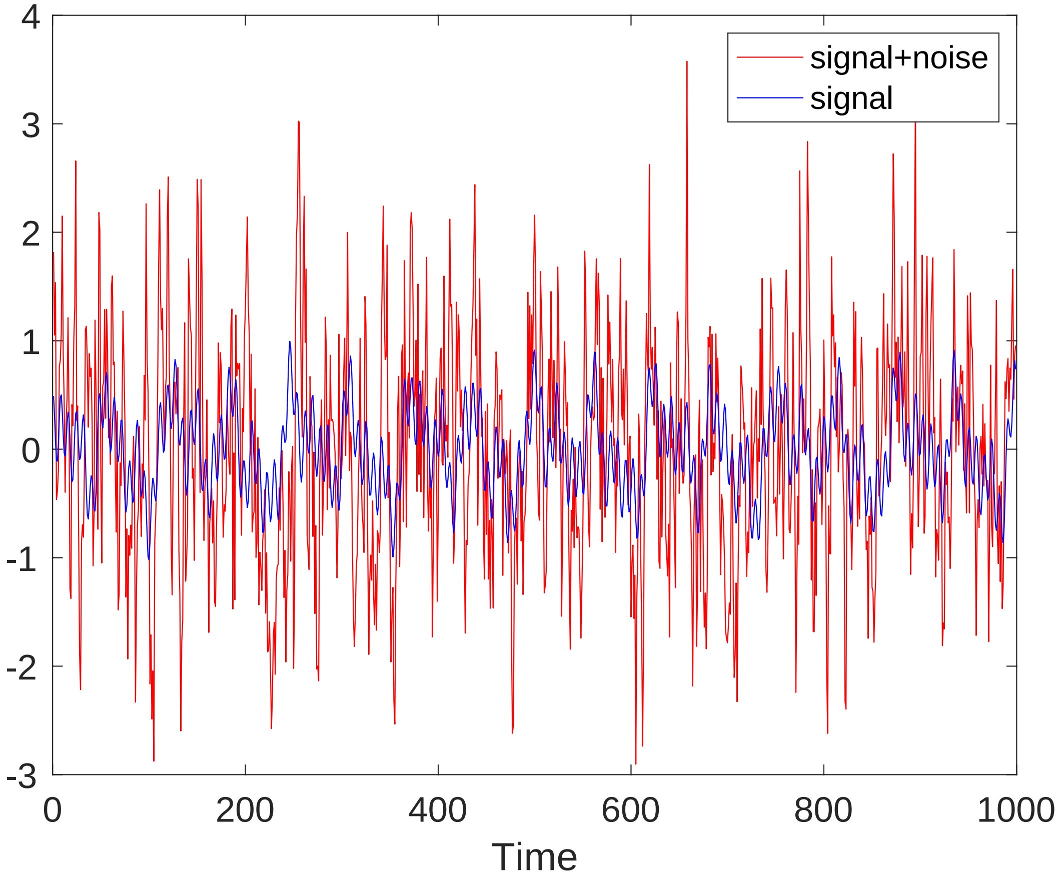


FIG. 2: Time series of the wavy signal and full data with imposed red noise at selected (x,y) point.

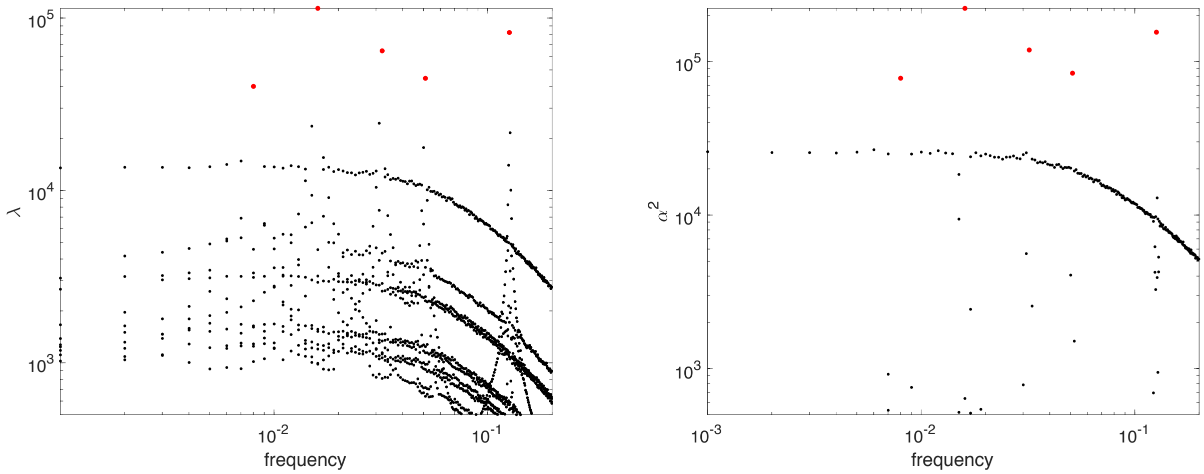


FIG. 3: DAHD spectra of eigenvalues λ (left panel) and energy α^2 (right panel) for the 2-D synthetic dataset: red circles – pairs of modes associated with largest values at reference wave frequencies.

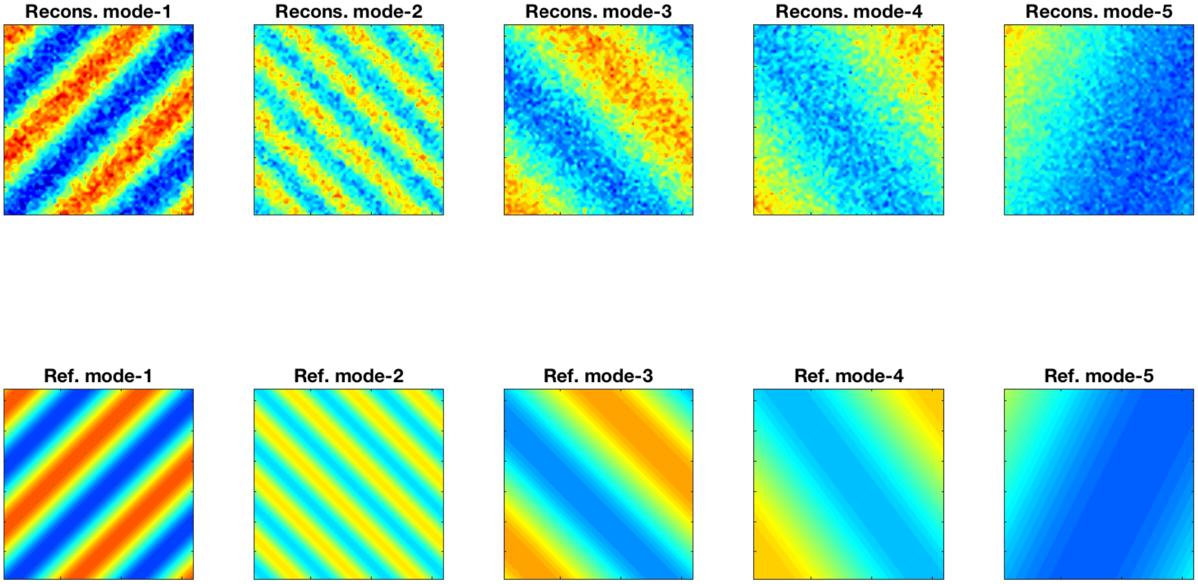


FIG. 4: Temporal snapshot for DAHD reconstruction of waves by using DAHMs associated with spectral peaks in Fig. 3, and comparison with respective reference patterns (see Fig. 1). (Multimedia view).

$k_2 = 20$; $k_3 = -5$; $k_4 = 4$; $k_5 = -2$; $l_1 = 10$; $l_2 = 20$; $l_3 = -5$; $l_4 = 3$; $l_5 = 1$, and the wave weights $A_1 = 1$; $A_2 = 0.5$, $A_3 = 0.75$; $A_4 = 0.6$ and $A_5 = 0.9$. With an appropriately chosen time step, the wave periodicities $|2\pi/\omega_n|$ in sampling units are $T_1 \approx 63$, $T_2 \approx 126$, $T_3 \approx 32$, $T_4 \approx 20$ and $T_5 \approx 8$, and we generate dataset with $N = 999$ points in time. By using coefficient $\nu = 0.9$ in Eq. 21, we consider a case of low signal-to-noise ratio, i.e. the wavy signal accounts only for 14% of the variance in the full data $\mathbf{v}(x, y, t)$, see comparison of "clean" 2-D wave snapshots with their noisy sum (Fig. 1 (Multimedia view), and time series at (x,y) point (Fig. 2).

Then we apply frequency-domain DAHD Hermitian algorithm (Sec. II B) to the noisy data $\mathbf{v}(x, y, t)$ with the goal to diagnose and reconstruct reference waves $u_n(x, y, t)$ in space and time. We use embedding window $M = (N + 1)/2 = 500$ in sampling units to obtain maximum spectral resolution.

Figure 3 shows DAHD eigenvalues ($\lambda(f)$, see Eq. 13) and energy spectrum $\alpha^2(f)$ (see Eq. 18)), that is evenly spaced with $M = 500$ bins in the Nyquist range [0 0.5]. Both $\lambda(f)$ and $\alpha^2(f)$ spectra are characterized by sharp peaks at the frequencies of reference waves, located high above noise background and associated with a pair of DAHMs ($\mathbf{W}$'s, Eq. 3). Furthermore, the energy spectra $\alpha^2(f)$ has an additional advantage by providing cleaner identification of coherent waves from noise,

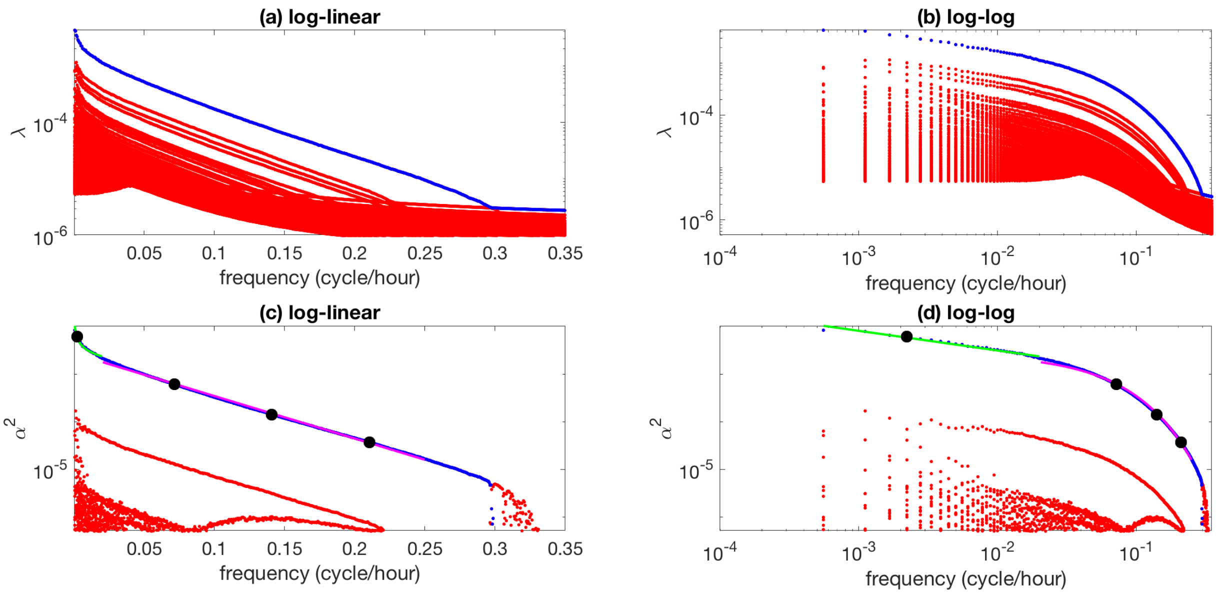


FIG. 5: DAHD spectra of eigenvalues $\lambda(f)$ (panels a,b) and energy $\alpha^2(f)$ (panels c,d) of the surface vorticity field in ROMS simulation: blue – modes with the largest λ at each frequency; green - power law fit $\approx f^{-0.4}$, magenta - exponential law fit $e^{-2\tau f}$ with characteristic time scale $\tau \approx$ 10 hours; black dots – modes at selected frequencies shown in Fig. 6.

i.e. without multiple background lines present in $\lambda(f)$ spectra.

Figure 4 (Multimedia view) shows fairly accurate reconstruction (see Eq. 15) of reference wave patterns, by using pair of DAHM's ($\mathbf{W}$'s, Eq. 3) and DAHCs (μ's, Eq. 16) associated with the spectral peaks in Fig. 3.

IV. REGIONAL OCEANIC MODELING SYSTEM

The dataset consists of $N = 1803$ hourly snapshots of surface vorticity field in the square domain (with 1002x1002 spatial resolution) of equatorial region simulated by the Regional Oceanic Modeling System (Srinivasan *et al.*, 2017) – a primitive equation ocean model. To enable efficient DAHD analysis for such high-dimensional dataset, it was first compressed by PCA. The leading $d = 1000$ empirical orthogonal function (EOF) modes and principal components (PCs) (from a total of 1803 modes) that capture $\approx$99% of variance. Next, we apply Hermitian DAHD algorithm (Sec II B) to the dataset composed by the time series of $d = 1000$ PCs and with a window $M = (N + 1)/2 = 902$ (in sampling units) to obtain maximum spectral resolution.

Figure 5 shows DAHD eigenvalues ($\lambda(f)$, see Eq. 13) and energy spectrum $\alpha^2(f)$ (see Eq. 18))

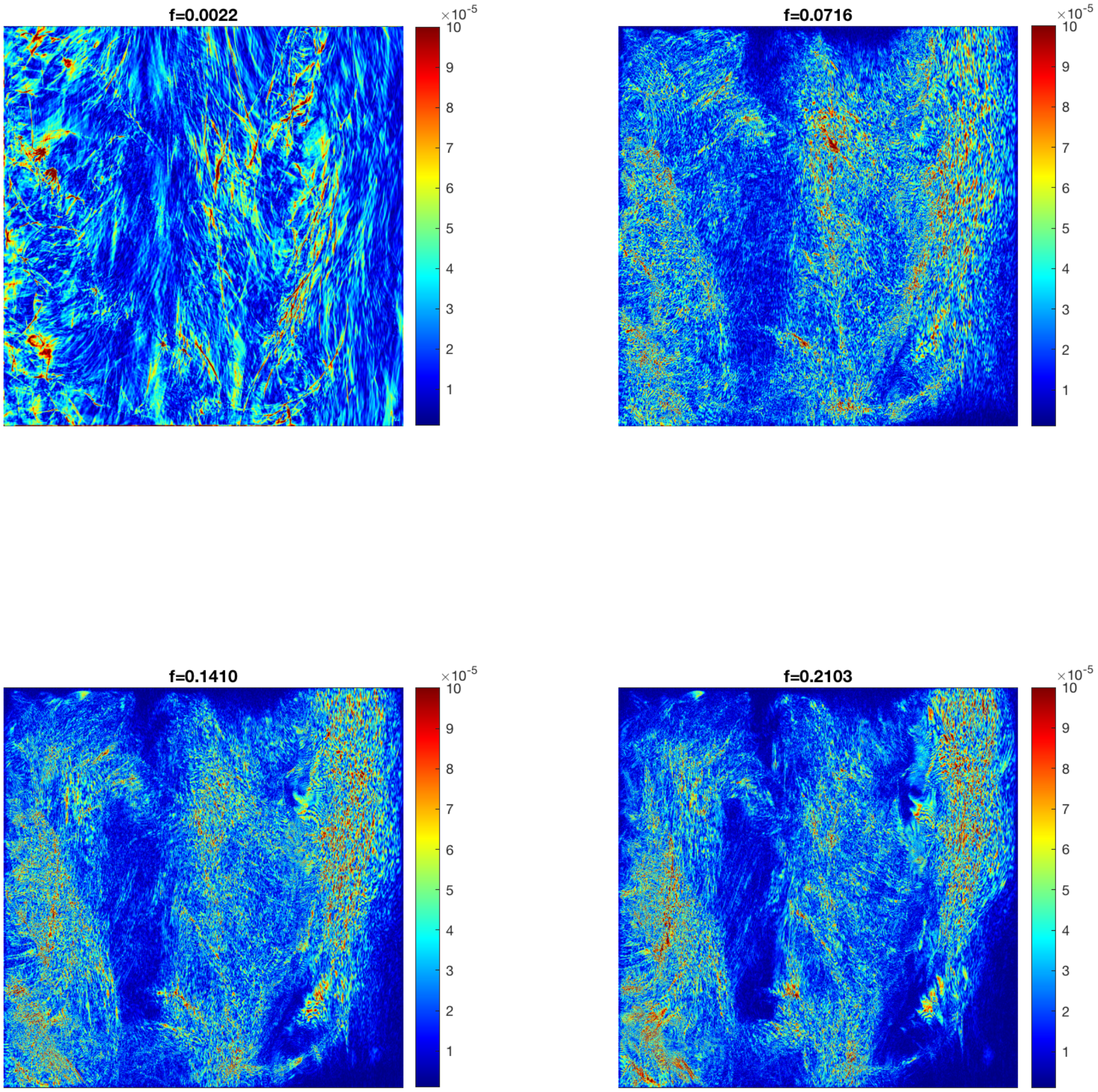


FIG. 6: Magnitude $|B|$ of DAHM (see Eq. 3) associated with largest $\alpha^2(f)$ and $\lambda(f)$ at selected frequencies (black dots in Figure 5).

that is evenly spaced with $M = 902$ bins in the Nyquist range [0 0.5]hr^{-1}. At each frequency, 100 largest values of α^2 is shown from the maximum possible $d = 1000$. The $\alpha^2(f)$ spectrum is characterized by a thin spectral line formed by a pair of DAHMs capturing largest energy per frequency, and located high above diffuse background. In comparison, the eigenspectrum $\lambda(f)$ is more compact and with smaller gap from the background, similar to results of Sec. III.

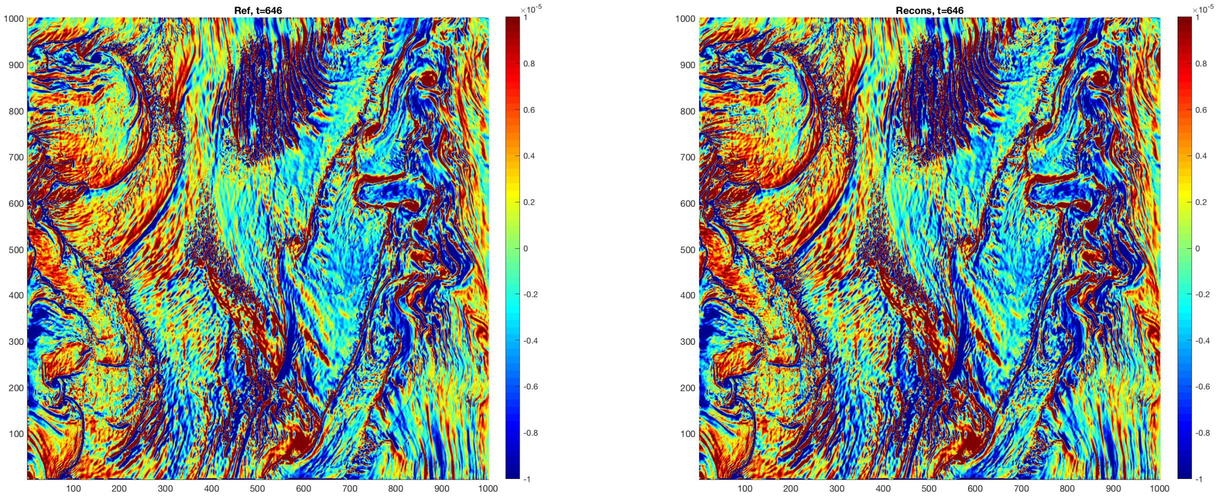


FIG. 7: An instantaneous snapshot of surface vorticity in the original data ROMS (left) vs. reconstruction (right) obtained by using DAHCs and DAHMs associated with the top spectral energy line, i.e. largest $\alpha^2(f)$ at given frequency, see text for details, units (s^{-1}). (Multimedia view).

Comparison of top spectral α^2 line in log-log and log-linear plots in Fig. 5, reveals power-like dependence at very low frequencies $f < 0.02\mathrm{hr}^{-1}$ (green line) and exponential decay for $f \in [0.02\ 0.3]\mathrm{hr}^{-1}$ (magenta line). Several spectral structures with much smaller magnitude and varying scaling laws are also revealed in the diffuse background for $f \in [0.15\ 0.35]\mathrm{hr}^{-1}$.

It should be noted that power spectra with exponential law have not been widely reported or discussed in the context of turbulent or chaotic data. Maggs and Morales (2012) have observed exponential spectra in UCLA plasma physics experiments, and have interpreted it as nonlinear signatures of chaotic dynamics in *Lorentzian pulses*, given by $g(t) = \frac{\tau^2}{(t-t_0)^2+\tau^2}$ with power spectra $\approx e^{\frac{-2f}{f_s}}$, where $f_s = \frac{1}{\tau}$ is the scaling frequency. By fitting values in the top spectral line we obtain $\tau \approx 10$ hours for exponential spectra, and $\approx f^{-0.4}$ for the power-law spectra at very low frequencies.

Figure 6 shows magnitude $|B|$ of the DAHM after transformation into physical space from EOF space and associated with the largest value of $\alpha^2(f)$ at selected frequencies – black circles in Fig. 5. While each of the DAHMs oscillates harmonically in time (see Eq. 3), it is inherently *multiscale* with large and small spatial features. Their dynamical interpretation as optimal response modes of linear system (Sec. II B) hinges on the validity of approximating effective forcing of the vorticity field, by the white noise in space and time, and it is clearly beyond the scope of this study.

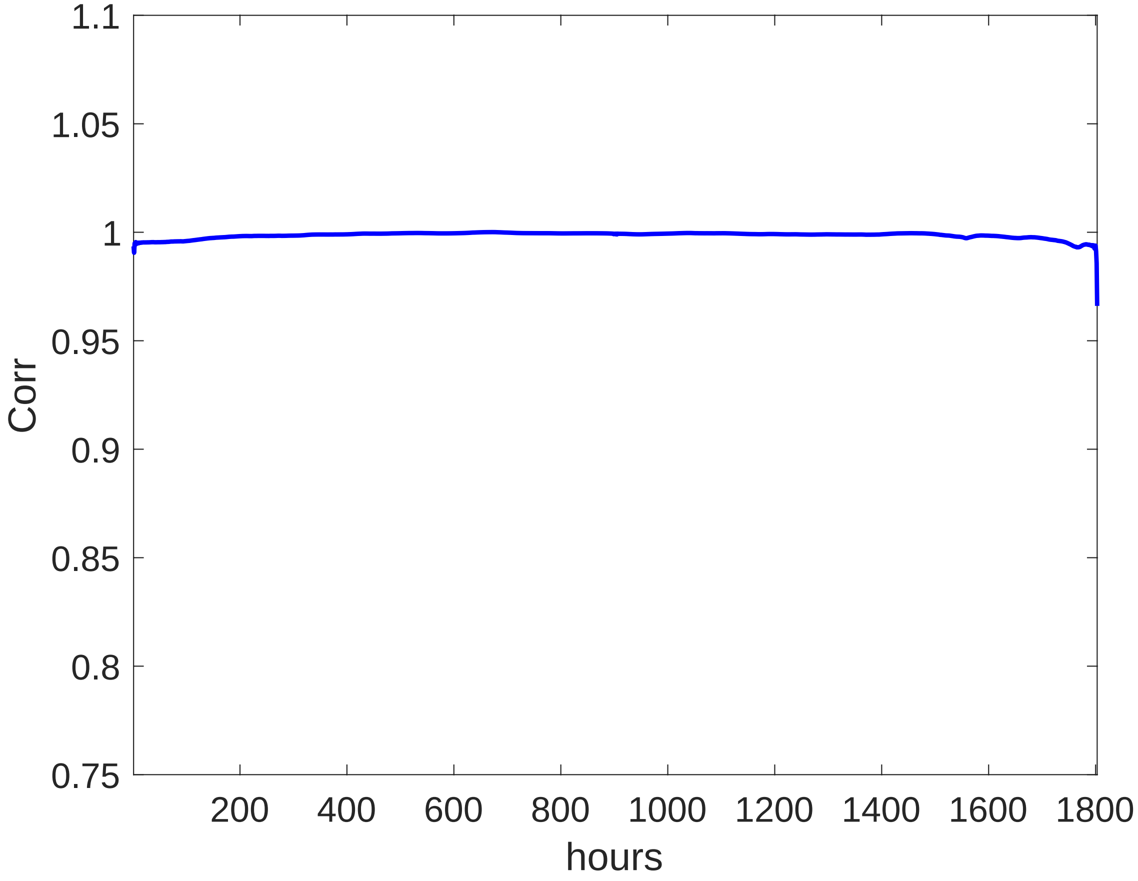


FIG. 8: Correlation of instantaneous spatial patterns of reconstruction and reference ROMS vorticity fields (see Fig. 7).

Figure 7 (Multimedia view) demonstrates that the instantaneous full vorticity flow field can be reconstructed with high accuracy (see Eq 15) by using a small subset of DAHD eigenelements in the top spectral energy line across frequencies, i.e. a pair of DAHM's ($\mathbf{W}$'s, Eq. 3) and associated DAHCs (μ's, Eq. 16) which yield the largest energy contribution ($\alpha^2(f)$) at each frequency $f \in$ [0.0 0.35]hr^{-1}, and in total accounting for $\approx 97\%$ of the variance in the vorticity field. The correlation between reconstruction and reference patterns is very high over whole time, see Fig. 8.

Such high-degree of information compression in the reconstruction is achieved by $\approx 2 \cdot 900 = 1800$ DAHD-inferred spatial patterns (i.e. DAHMs and taking into account phase-quadrature property of the latter) which evolve with trivial temporal dynamics, namely harmonic oscillation. In contrast, PCA compression yields 1000 spatial EOFs but with very complex temporal dynamics in associated PCs.

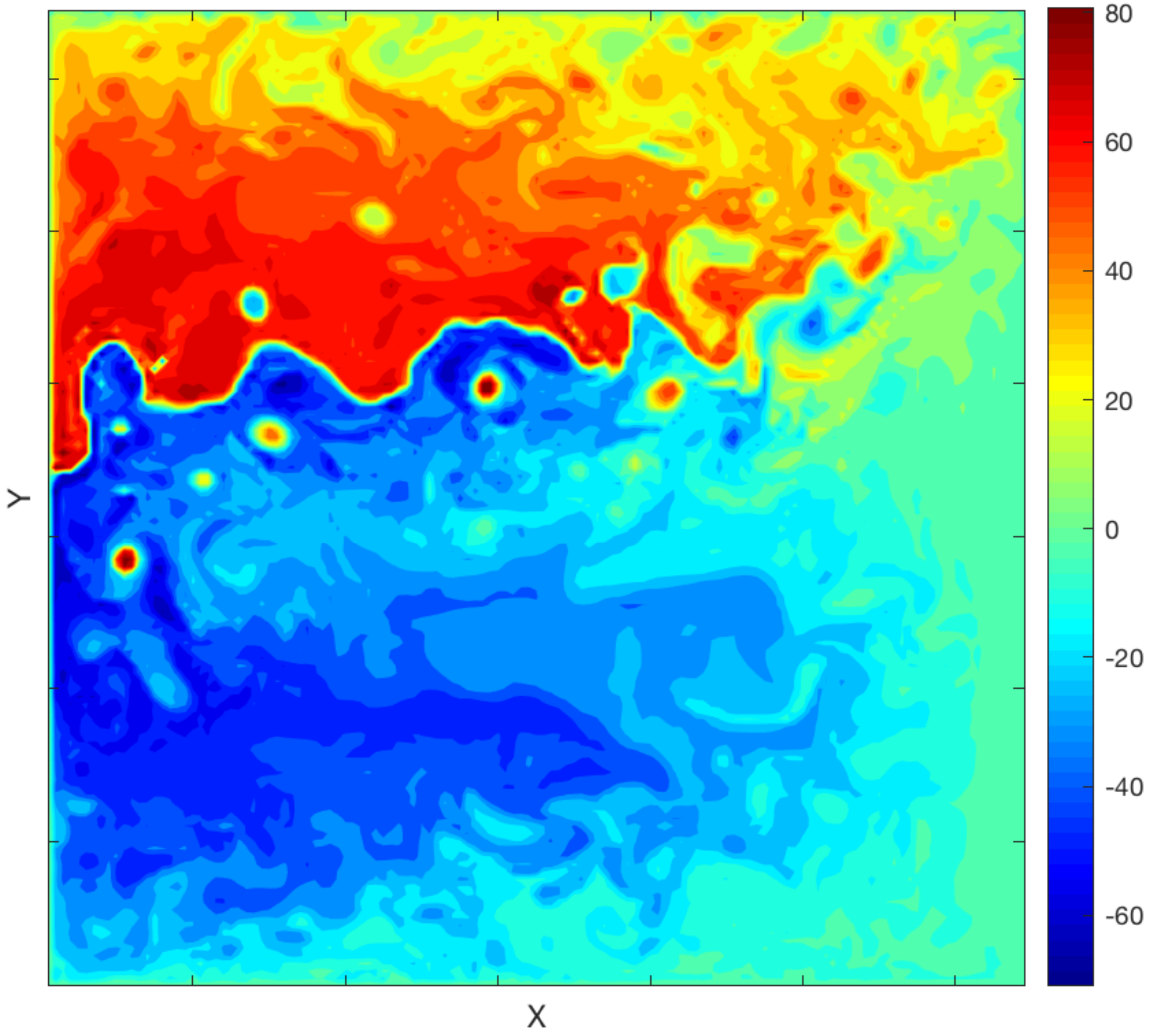
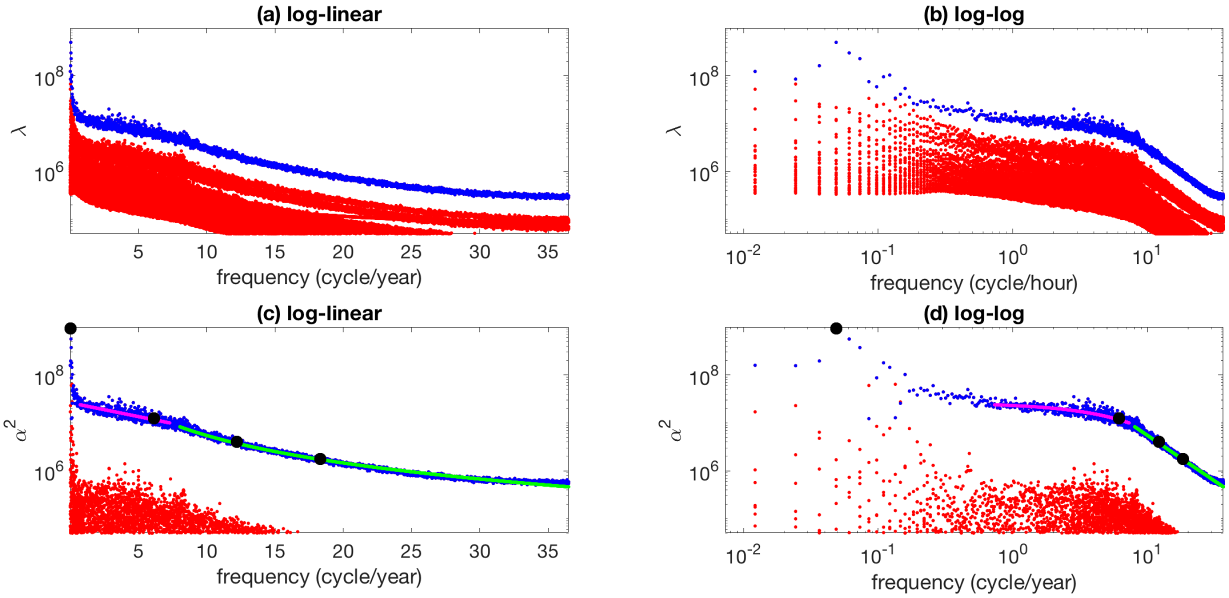


FIG. 9: Snapshot of upper-layer potential vorticity (PV) field in the three-layer quasigeostrocphic oceanic model simulation of wind-driven gyres (Ryzhov *et al.*, 2019). Nondimensional color scale units, the same for Fig. 12.

V. EDDY-RESOLVING OCEANIC MODEL OF WIND-DRIVEN GYRES

The dataset is obtained by the eddy-resolving oceanic quasigeostrophic (QG) model (Shevchenko and Berloff, 2015; Kondrashov and Berloff, 2015; Ryzhov *et al.*, 2019, 2020) that produces a double-gyre flow pattern, characterized by well-developed and turbulent eastward jet extension of the western boundary currents with its adjacent recirculation zones. Dataset consists of $N = 5999$ 5-day snapshots of potential vorticity in the upper layer (Fig. 9) solution with $513\text{x}513$ spatial resolution. The leading $d = 2000$ EOFs and PCs from PCA capture $\approx 99\%$ of the total variance.


(a) log-linear

(b) log-log

(c) log-linear

(d) log-log

FIG. 10: DAHD spectra of eigenvalues $\lambda(f)$ (panels a,b) and energy $\alpha^2(f)$ (panels c,d) of the PV field (Fig. 9); blue – modes with the largest λ at each frequency, green - power-law fit $\approx f^{-5/3}$; magenta - exponential law fit $e^{-2t_s f}$ with characteristic time scale $\tau_s \approx 25$ days, black dots – DAHD modes shown in Fig. 11.

Next, we apply Hermitian DAHD algorithm (Sec II B) to the dataset composed by the time series of $d = 2000$ PCs and with a window $M = (N+1)/2 = 3000$ (in sampling units).

Figure 10 shows DAHD eigenvalues ($\lambda(f)$, see Eq. 13) and energy spectrum $\alpha^2(f)$ (see Eq. 18)). For energy spectra, only largest 30 $\alpha^2(f)$ values are shown from the total number of $d = 2000$ at each frequency. Shape of the spectra is characterized by a narrow spectral line of one $\alpha^2(f)$ value per frequency with a wide gap above diffuse background of lower values. Furthermore, the top spectral line exhibits sharp spectral peak at ≈ 20 years, followed by exponential law $e^{-2\tau f}$ with characteristic time scale $\tau \approx 25$ days, while Kolmogorov's $(-\frac{5}{3})$ power law is diagnosed in the high-frequency range (green).

Figure 11 shows magnitude $|B|$ of the DAHM associated with the largest value of $\alpha^2(f)$ at selected frequencies (see black circles in Fig. 10). Notably, the DAHM associated with decadal spectral peak, represents a coherent spatial pattern along the jet, while the spatial patterns in exponential and Kolmogorov ranges are patchy and inherently multiscale.

Figure 12 (Multimedia view) shows that instantaneous PV anomaly field can be reconstructed very accurately (see Eq 15) by using DAHD eigenelements in the top spectral line (i.e. blue dots

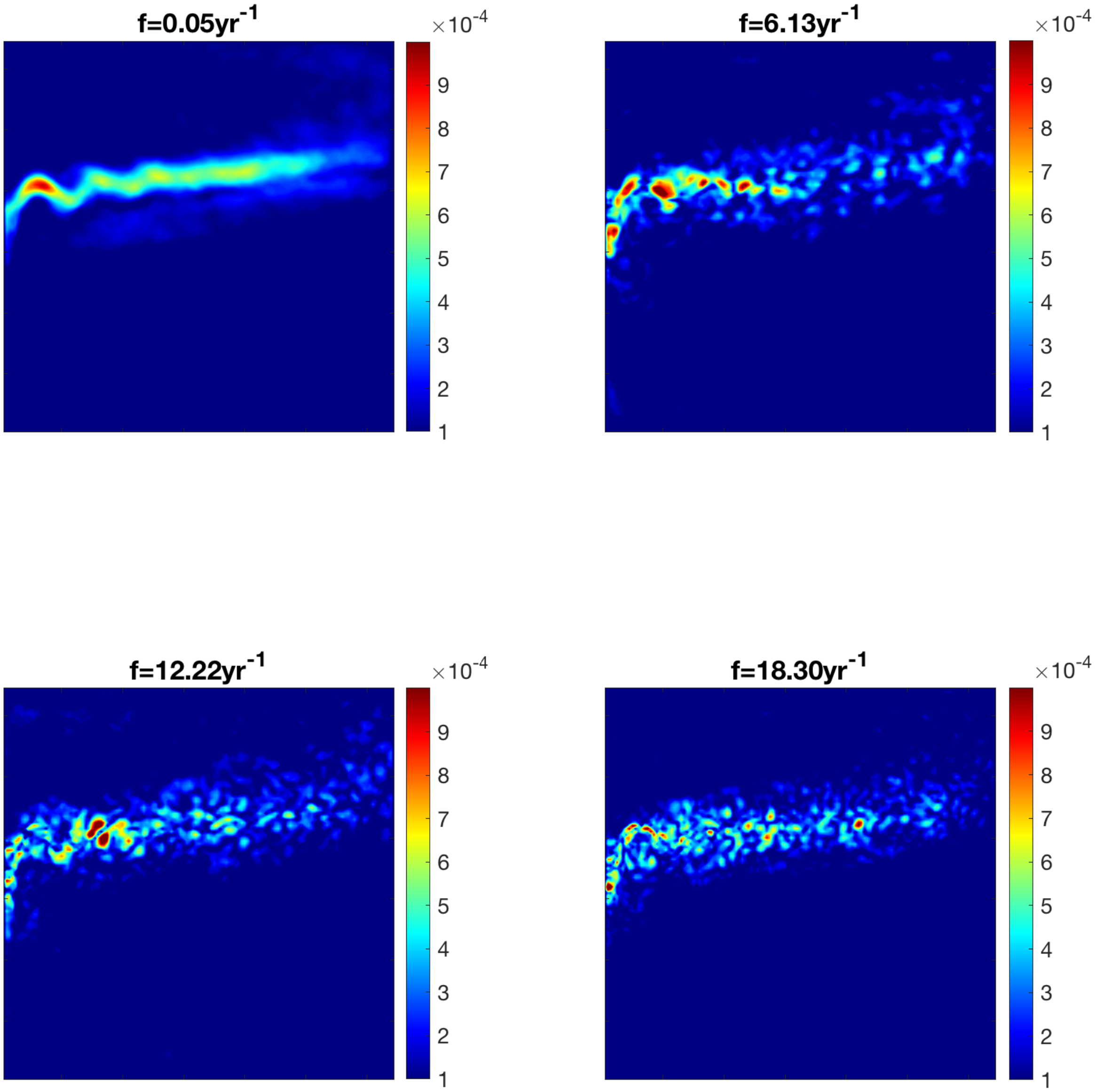

f=0.05yr^{-1}
f=6.13yr^{-1}
f=12.22yr^{-1}
f=18.30yr^{-1}

FIG. 11: Magnitude $|B|$ of DAHM (see Eq. 3) associated with largest $\alpha^2(f)$ at selected frequencies f (cycle/year), see black dots in top spectral line in Figure 10).

in Fig. 10), accounting for $\approx 96\%$ of variance in the reference field, while correlation between reconstruction and reference PV patterns is high over whole time interval, see Fig. 13.

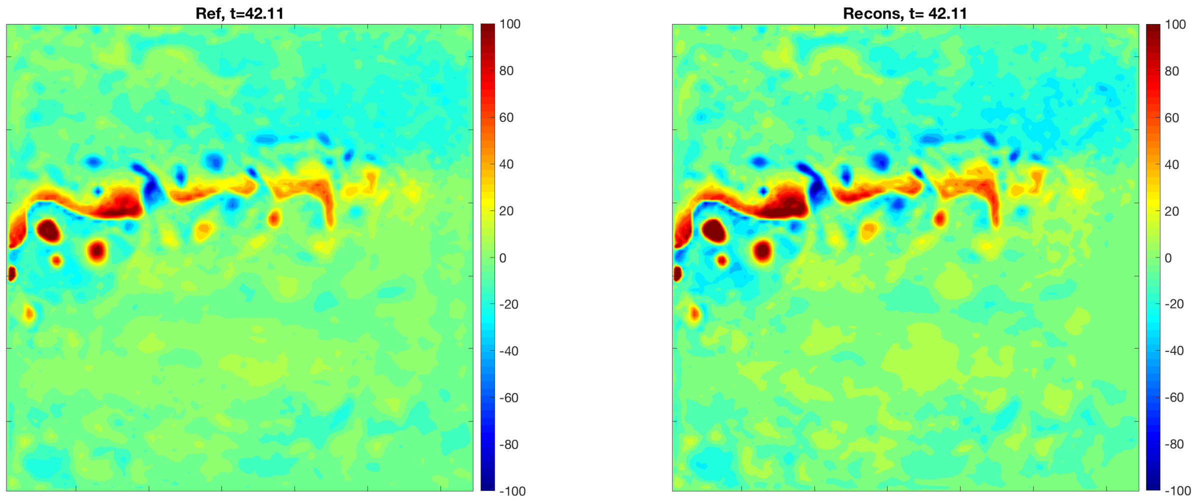


FIG. 12: An instantaneous snapshot of PV anomaly in the original QG data (left) vs. reconstruction (right) obtained by using DAHD elements in the top spectral line (blue dots in Fig. 10). (Multimedia view).

VI. DISCUSSION AND CONCLUSIONS

The energy spectra and scaling laws of turbulent atmospheric and oceanic flows (Delsole, 2004; Callies and Ferrari, 2013; McWilliams, 2016; Chapman, 2017) are commonly analyzed in terms of spatial scales (wavenumber), implicitly implying ergodic assumption, i.e. larger/smaller spatial scales corresponding to their slower/faster evolution. In comparison, DAHD explicitly focuses on the temporal scales but without any built-in assumptions on their spatial content, and thus can provide new insights, such as an exponential shape of energy spectra not widely reported before. Furthermore, DAHD yields sparse representation of complex, multiscale and chaotic dynamics by relatively few data-inferred spatial patterns oscillating harmonically in time. The observed low-rank behavior can be interpreted as dominance of given physical mechanism of energy distribution and transfer across temporal frequencies, and revealed by scaling laws of DAHD spectra. By utilizing correlogram estimator of Hermitian cross-spectral density matrix, DAHD is closely related but distinctly different from SPOD that relies on Welch's periodogram.

The results of this study pave way to several promising DAHD applications, such as (i) obtaining highly accurate spectral fingerprint of scaling laws in geophysical and astrophysical turbulent fluid flows, (ii) global spectral diagnostics for comparison and characterization of model simulations and observations, (iii) parsimonious extraction and reconstruction of mutiscale features of interest

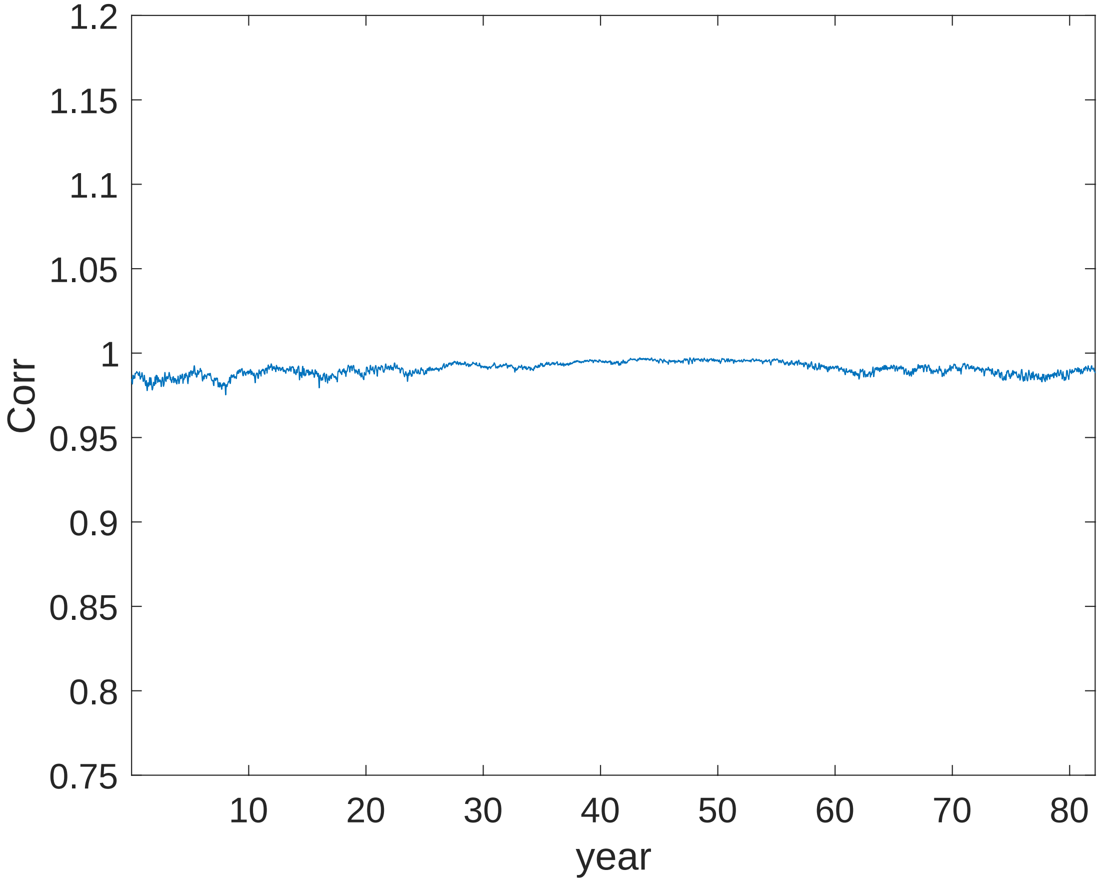


FIG. 13: Correlation of instantaneous spatial patterns of reconstruction and reference PV anomaly fields (see Fig. 12).

enabled by DAHD compression of energetic content, (iv) uncovering invariant laws across different spatial and temporal scales and phenomena, (v) predictive capability enabled by simple temporal dynamics contained in DAHMs, (vi) identification of coherent structures in noisy data.

ACKNOWLEDGMENTS

The authors would like to thank Prof. James McWilliams and Dr. Kaushik Srinivasan for many useful discussions and providing ROMS dataset. Two anonymous reviewers provided valuable input which helped to improve the manuscript.

This research was supported by the National Science Foundation (NSF) grants $OCE-1658357$ and the NERC grant $NE/R011567/1$. P.B. also gratefully acknowledges funding by NERC Grant No. NE/T002220/1 and Leverhulme Grant No. RPG-2019-024. We also would like to acknowledge high-performance computing support from Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by NSF.

DAHD Toolbox is available at: http://research.atmos.ucla.edu/tcd/dkondras/Software.html

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

Budišić, M., Mohr, R., and Mezić, I., "Applied Koopmanism," Chaos **22**, 047510 (2012).

Callies, J. and Ferrari, R., "Interpreting energy and tracer spectra of upper-ocean turbulence in the submesoscale range (1–200 km)," Journal of Physical Oceanography **43**, 2456–2474 (2013).

Chapman, C. C., "New perspectives on frontal variability in the Southern Ocean," Journal of Physical Oceanography **47**, 1151–1168 (2017).

Chekroun, M. D. and Kondrashov, D., "Data-adaptive harmonic spectra and multilayer Stuart-Landau models," Chaos **27**, 093110 (2017).

Delsole, T., "Stochastic models of quasigeostrophic turbulence," Surveys in Geophysics **25**, 107–149 (2004).

Gavrilov, A., Mukhin, D., Loskutov, E., Volodin, E., Feigin, A., and Kurths, J., "Method for reconstructing nonlinear modes with adaptive structure from multidimensional data," CHAOS **26**, 123101 (2016).

Kondrashov, D. and Berloff, P., "Stochastic modeling of decadal variability in ocean gyres," Geophysical Research Letters **42**, 1543–1553 (2015).

Kondrashov, D. and Chekroun, M. D., "Data-adaptive harmonic analysis and modeling of solar wind-magnetosphere coupling," Journal of Atmospheric and Solar-Terrestrial Physics (2018), 10.1016/j.jastp.2017.12.021.

Kondrashov, D., Chekroun, M. D., and Berloff, P., "Multiscale Stuart-Landau emulators: Application to wind-driven ocean gyres," Fluids **3**, 21 (2018).

Kondrashov, D., Chekroun, M. D., and Ghil, M., "Data-driven non-Markovian closure models," Physica D. **297**, 33–55 (2015).

Kondrashov, D., Chekroun, M. D., and Ghil, M., "Data-adaptive harmonic decomposition and prediction of Arctic sea ice extent," Dynamics and Statistics of the Climate System (2018), 10.1093/climsys/dzy001.

Kondrashov, D., Chekroun, M. D., Yuan, X., and Ghil, M., "Data-adaptive harmonic decomposition and stochastic modeling of Arctic sea ice," in *Advances in Nonlinear Geosciences*, edited by A. Tsonis (Springer, 2018).

Kravtsov, S., Tilinina, N., Zyulyaeva, Y., and Gulev, S. K., "Empirical modeling and stochastic simulation of sea level pressure variability," Journal of Applied Meteorology and Climatology **55**, 1197–1219 (2016).

Maggs, J. E. and Morales, G. J., "Exponential power spectra, deterministic chaos and lorentzian pulses in plasma edge dynamics," Plasma Physics and Controlled Fusion **54**, 124041 (2012).

Martinez-Villalobos, C., Vimont, D. J., Penland, C., Newman, M., and Neelin, J. D., "Calculating State-Dependent Noise in a Linear Inverse Model Framework," Journal of the Atmospheric Sciences **75**, 479–496 (2018).

McWilliams, J. C., "Submesoscale currents in the ocean," Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **472**, 20160117 (2016).

Mukhin, D., Gavrilov, A., Feigin, A., Loskutov, E., and Kurths, J., "Principal nonlinear dynamical modes of climate variability," Sci. Rep. **5**, 15510 (2015).

Penland, C., "Random forcing and forecasting using principal oscillation pattern analysis," Mon. Weather Rev. **117**, 2165–2185 (1989).

Penland, C., "A stochastic model of IndoPacific sea surface temperature anomalies," Physica D **98**, 534–558 (1996).

Percival, D. B. and Walden, A. T., *Spectral analysis for physical applications* (Cambridge university press, 1993).

Ryzhov, E., Kondrashov, D., Agarwal, N., and Berloff, P., "On data-driven augmentation of low-resolution ocean model dynamics," Ocean Modelling **142**, 101464 (2019).

Ryzhov, E., Kondrashov, D., Agarwal, N., and Berloff, P., "On data-driven induction of the low-frequency variability in a coarse-resolution ocean model," Ocean Modelling (2020), submitted.

Schmid, P. J., "Dynamic mode decomposition of numerical and experimental data," J. Fluid Mech. **656**, 5–28 (2010).

Schmidt, O. T., Mengaldo, G., Balsamo, G., and Wedi, N. P., "Spectral empirical orthogonal functio n analysis of weather and climate data," Monthly Weather Review **147**, 2979–2995 (2019).

Shevchenko, P. and Berloff, P., "Multi-layer quasi-geostrophic ocean dynamics in eddy-resolving regimes," Ocean Modelling **84**, 1–14 (2015).

Srinivasan, K., McWilliams, J. C., Renault, L., Hristova, H. G., Molemaker, J., and Kessler, W. S., “Topographic and mixed layer submesoscale currents in the near-surface southwestern tropical pacific,” Journal of Physical Oceanography **47**, 1221–1242 (2017).

Towne, A., Schmidt, O. T., and Colonius, T., “Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis,” Journal of Fluid Mechanics **847**, 821–867 (2018).

Tu, J. H., Rowley, C. W., Luchtenburg, D. M., Brunton, S. L., and Kutz, J. N., “On dynamic mode decomposition: Theory and applications,” J. Comput. Dyn. **1**, 391–421 (2014).

Williams, M. O., Kevrekidis, I. G., and Rowley, C. W., “A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition,” J. Nonlin. Sci. **25**, 1307–1346 (2015).