

---

## Eigenvalues of autocovariance matrix: A practical method to identify the Koopman eigenfrequencies

Zhen Yicun <sup>1,\*</sup>, Chapron Bertrand <sup>1</sup>, Mémin Etienne <sup>2</sup>, Peng Lin <sup>3</sup>

<sup>1</sup> Institut Français de Recherche pour l'Exploitation de la Mer, 29280 Plouzané, France

<sup>2</sup> INRIA/IRMAR, Campus universitaire de Beaulieu, Rennes, 35042 Cedex, France

<sup>3</sup> Ocean University of China, 266100 Qingdao, China

\* Corresponding author : Yicun Zhen email address : [zhenyicun@protonmail.com](mailto:zhenyicun@protonmail.com)

---

### Abstract :

To infer eigenvalues of the infinite-dimensional Koopman operator, we study the leading eigenvalues of the autocovariance matrix associated with a given observable of a dynamical system. For any observable  $f$  for which all the time-delayed autocovariance exist, we construct a Hilbert space  $H_f$  and a Koopman-like operator  $K$  that acts on  $H_f$ . We prove that the leading eigenvalues of the autocovariance matrix has one-to-one correspondence with the energy of  $f$  that is represented by the eigenvectors of  $K$ . The proof is associated to several representation theorems of isometric operators on a Hilbert space, and the weak-mixing property of the observables represented by the continuous spectrum. We also provide an alternative proof of the weakly mixing property. When  $f$  is an observable of an ergodic dynamical system which has a finite invariant measure  $\mu$ ,  $H_f$  coincides with closure in  $L^2(X, d\mu)$  of Krylov subspace generated by  $f$ , and  $K$  coincides with the classical Koopman operator. The main theorem sheds light to the theoretical foundation of several semi-empirical methods, including singular spectrum analysis (SSA), data-adaptive harmonic analysis (DAHD), Hankel DMD, and Hankel alternative view of Koopman analysis (HAVOK). It shows that, when the system is ergodic and has finite invariant measure, the leading temporal empirical orthogonal functions indeed correspond to the Koopman eigenfrequencies. A theorem-based practical methodology is then proposed to identify the eigenfrequencies of  $K$  from a given time series. It builds on the fact that the convergence of the renormalized eigenvalues of the Gram matrix is a necessary and sufficient condition for the existence of  $K$ -eigenfrequencies. Numerical illustrating results on simple low dimensional systems and real interpolated ocean sea-surface height data are presented and discussed.

## I. INTRODUCTION

The dynamic mode decomposition (DMD) algorithm, is a powerful and versatile data-driven approach proposed by [1], ideally suited to analyze complex high-dimensional geophysical flows in terms of recurrent or quasi-periodic modes. The DMD is indeed related to the Koopman theory [2], stating that observables of an Hamiltonian system can always be described via a linear transformation. The original DMD algorithm has a lot of common points with the algorithm presented in [3]. For practical applications and real data analysis, several follow-up algorithms have been proposed. To name a few, it can be listed the optimized DMD [4], the optimal mode decomposition [5], the exact DMD [6], the Hankel DMD [7], the sparsity promoting DMD [8], the multi-resolution DMD [9], the extended DMD [10], DMD with control [11], total least squares DMD [12], dynamic distribution decomposition [13], etc. These DMD algorithms are generally mo-

---

tivated by different reasons, but a key overall objective is to help provide the most precise numerical approximation of the Koopman operator. When the system is ergodic and measure-preserving, it would indeed be equivalent to have a precise description of the spectrum  $\nu$  of Koopman operator restricted on  $\mathcal{H}_f$  and a precise mapping between  $\mathcal{H}_f$  and  $L^2(S^1, d\nu)$  (where  $\mathcal{H}_f$  is the linear subspace generated by a single observable  $f$  and  $S^1$  the unit complex circle, see section 2 for detailed definition of these spaces). The authors of [14] proved the convergence in the strong operator topology of extended DMD algorithm, provided a complete orthogonal basis of the space of square-integrable observables. In [7] the convergence of Hankel DMD algorithm is proved for the finite dimensional case, which corresponds to the finite truncation of the discrete part of the spectrum. Christoffel-Darboux kernel is exploited in [15] to directly identify the discrete component and the absolutely continuous component of the spectrum. Note, DMD algorithms are not the only way to approximate Koopman operator. In a series of papers ([16],[17], [18],[19] and [20]), the approximation of Koopman operator is performed by kernel methods. Recently, [20] showed the convergence of kernel methods for any measure preserving ergodic dynamical systems, the measure of which support lies on a compact manifold. A related issue for stochastic dynamics can be found in [21]

In this manuscript, we define a Hilbert space  $\mathcal{H}_f$  and the associated Koopman-like time-shift operator  $\mathcal{K}$  for each given time series the autocovariances of which exist. When the time series is given by an observable  $f$  on a ergodic dynamical system with finite invariant measure,  $\mathcal{H}_f$  coincides with the closure of the infinite dimensional Krylov subspace generated by the classical Koopman operator and  $f$ , and  $\mathcal{K}$  coincides with the classical Koopman operator on  $\mathcal{H}_f$ . We argue and prove that for any observable  $f$  such that all the time-delayed autocovariance exist, when the parameters of the autocovariance matrix  $C_{NM}(f)$  goes to infinity in the right order, the leading eigenvalues renormalized by the dimension of  $C_{NM}(f)$  converge to the energy of  $f$  that is represented by the eigenvector of  $\mathcal{K}$ . All other renormalized eigenvalues shall further converge to 0 uniformly. Despite its theoretical interests, the main theorem directly suggests a practical algorithm to explicitly identify the Koopman eigenfrequencies together with the associated energy from given time series. As a by-product, it also shows that the leading temporal empirical orthogonal functions calculated by singular spectrum analysis (SSA, [22]) method are indeed represented by the eigenfrequencies of  $\mathcal{K}$ . Similarly, this theorem also sheds light on the theoretical foundation of data-adaptive harmonic decomposition (DAHD,[23], and Hankel alternative view of Koopman analysis (HAVOK, [24]). Because all these methods are based either on trajectory matrix or on  $C_{NM}(f)$ .

The paper is organized as follows. In section 2, we present our main result and the necessary mathematical background knowledge. We also discuss about how the main theorem provides theoretical support to SSA, DAHD, and HAVOK. In section 3, we state the details of the proof of the main result. In section 4, we present the detailed algorithm and compare it with another numerical method based on Yosida's mean ergodic theorem ([25]). In section 5 we present numerical results on two simple low dimensional measure preserving ergodic dynamical systems and interpolated ocean sea-surface height data. Section 6 concludes this study and gives some perspectives. The necessary code and data that reproduces all the numerical results can be accessed at <https://doi.org/10.5281/zenodo.5585970>.

## II. PRELIMINARIES AND THE MAIN RESULT

Given a continuous-time dynamical system

$$X_t = \Phi_t(X_0),$$

and an observable  $f(X_t)$ , we have a time series  $\{f(X_t) : t = 0, \Delta t, 2\Delta t, \dots\}$ . We assume that the time-delayed autocovariance of  $f$  exists for all  $l \in \mathbb{N}$ :

$$\rho_{l\Delta t} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n f(X_{k\Delta t}) \bar{f}(X_{(k+l)\Delta t}). \quad (1)$$

To avoid misinterpretation, we use  $F_0 = \{f(X_0), f(X_{\Delta t}), \dots\}$  to denote time series associated to  $f$  and a given (fixed) orbit of the dynamical system. For  $l \in \mathbb{N}$ , we define  $F_l$  to be the time shifted time series  $\{f(X_{l\Delta t}), f(X_{(l+1)\Delta t}), \dots\}$ . For any  $a, b \in \mathbb{C}$  and any two time series, associated to the same dynamics

$$\begin{aligned} G &= \{g_0, g_1, \dots\} \\ H &= \{h_0, h_1, \dots\}, \end{aligned}$$

we define

$$aG + bH = \{ag_0 + bh_0, ag_1 + bh_1, \dots\}.$$

Let

$$\tilde{\mathcal{H}}_f = \left\{ \sum_{i=1}^n c_i F_{l_i} : n \geq 1, c_i \in \mathbb{C}, l_i \in \mathbb{N} \right\}. \quad (2)$$

Then  $\tilde{\mathcal{H}}_f$  is a linear space. Now we define the Koopman-like (or time shift) operator on  $\tilde{\mathcal{H}}_f$ . We start with for any  $l \in \mathbb{N}$

$$\mathcal{K}^{\Delta t} F_l = F_{l+1}, \quad (3)$$

and then generalize the action of  $\mathcal{K}^{\Delta t}$  to the whole  $\tilde{\mathcal{H}}_f$ . It is not hard to show that  $\mathcal{K}^{\Delta t} : \tilde{\mathcal{H}}_f \rightarrow \tilde{\mathcal{H}}_f$  is well-defined.

The existence of  $\rho_s$  allows us to define an inner product on  $\tilde{\mathcal{H}}_f$  by

$$\langle h, g \rangle = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} h_k \bar{g}_k, \quad (4)$$

where  $h, g \in \tilde{\mathcal{H}}_f$ . For any  $l_1, l_2 \in \mathbb{N}$ , it is obvious that

$$\rho_{l_1\Delta t} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n f(X_{(n+l_1)\Delta t}) \bar{f}(X_{(n+l_1+l_2)\Delta t}). \quad (5)$$

Hence  $\mathcal{K}^{\Delta t}$  preserves the inner product in  $\tilde{\mathcal{H}}_f$  and is hence continuous.

Let  $\mathcal{H}_f$  be the completion of  $\tilde{\mathcal{H}}_f$ .  $\mathcal{H}_f$  is a Hilbert space.  $\mathcal{K}^{\Delta t}$  preserves the inner product of  $\tilde{\mathcal{H}}_f$ . Therefore  $\mathcal{K}^{\Delta t}$  can be extended, by continuity, to an isometric operator that acts on  $\mathcal{H}_f$ . For sake of simplicity, keeping the same notation for the extended operator,  $\mathcal{K}^{\Delta t}$ , whose domain is  $\mathcal{H}_f$ , is the Koopman-like (or time shift) operator we study in this paper.

**Remark 1.** *The classical Koopman operator is defined to act on some function space on the whole phase space of some dynamical system, i.e.*

$$\mathcal{K}^{cl} g(x) = g(\Phi(x)), \quad (6)$$

where  $x \in X$  is a state and  $\Phi$  the discrete-time flow of the dynamical system.

In our setting, the definition of  $\mathcal{K}^{\Delta t}$  and  $\mathcal{H}_f$  purely relies on the time series. The dynamical system is hidden

behind. It is possible that the time series only reveals partial properties of the dynamical system. When the discrete-time dynamical system is ergodic and has finite invariant measure  $\mu$ , Birkhoff ergodic theorem guarantees that  $\mathcal{H}_f$  is isomorphic to the closure of Krylov subspace  $\overline{\text{Span}\{f, \mathcal{K}^{\text{cl}}f, \dots\}} \subset L^2(X, d\mu)$  as Hilbert spaces. The isomorphism is given by

$$\phi : \sum_{i=1}^n c_i F_{l_i} \rightarrow \sum_{i=1}^n c_i (\mathcal{K}^{\text{cl}})^{l_i} f \quad (7)$$

And it is not hard to see that  $\mathcal{K}^{\Delta t}$  acts on  $\mathcal{H}_f$  in a similar way as  $\mathcal{K}^{\text{cl}}$  on the space of observables  $L^2(X, d\mu)$ , i.e. for any  $h \in \mathcal{H}_f$ ,

$$\phi(\mathcal{K}h) = \mathcal{K}^{\text{cl}}(\phi(h)). \quad (8)$$

The elements in  $\tilde{\mathcal{H}}_f$  can always be represented as some time series. But in general we can not assert that any element in  $\mathcal{H}_f$  can be represented as some time series. In particular, we do not know if the eigenvectors  $v_i \in \mathcal{H}_f$  of  $\mathcal{K}$  can be represented as a time series in the form  $\{1, \xi_i, \xi_i^2, \dots\}$ . Nor in general can we identify  $\mathcal{H}_f$  with some space of functions on the phase space. Hence, without these assumptions, the time-shift operator and the Koopman operator cannot be strictly related. Ergodicity+finite invariant measure is a stronger assumption than the existence of autocovariance. As long as the autocovariances exist, the quantities mentioned in the main result are mathematically well-defined. Therefore we do not restrict ourselves to the case where the system is ergodic and has finite invariant measure. But we keep the generality in the definition of  $\mathcal{H}_f$  and  $\mathcal{K}$ . The physical meaning of  $\mathcal{H}_f$  and  $\mathcal{K}$  in the general case needs to be studied further. For readers who are interested in the classical Koopman operator  $\mathcal{K}^{\text{cl}}$ , we point out that  $\mathcal{K}$  indeed coincides with  $\mathcal{K}^{\text{cl}}$  when the system is ergodic and has finite invariant measure.

If  $\{f(X_t) : t \geq 0\}$  is a continuous-time time series, we assume that

$$\rho_s = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(X_t) \bar{f}(X_{t+s}) dt \quad (9)$$

exists for every  $s \geq 0$ . Similar to the discrete-time case, we define

$$F_s = \{f(X_t) : t \geq s\} \quad (10)$$

to be the continuous time series that starts from  $f(X_s)$ . We can define the linear space

$$\tilde{\mathcal{H}}_f^{\text{cont}} = \left\{ \sum_{i=1}^n c_i F_{s_i} : c_i \in \mathbb{C}, s_i \geq 0, n \geq 1 \right\}, \quad (11)$$

with inner product:

$$\langle h, g \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T h(t) \bar{g}(t) dt. \quad (12)$$

Let  $\mathcal{H}_f^{\text{cont}}$  be the completion of  $\tilde{\mathcal{H}}_f^{\text{cont}}$ .  $\mathcal{K}^s$  is an isometry on  $\mathcal{H}_f^{\text{cont}}$  for all  $s \geq 0$ .

We are interested in the eigenvalues and eigenvectors of  $\mathcal{K}^{\Delta t}$  (for discrete-time case) and  $\mathcal{K}^s$  (for continuous-time case). A natural question to ask is that whether the eigenfrequencies of  $\mathcal{K}^{\Delta t}$  and  $\mathcal{K}^s$  are the same. We have the following result.

**Proposition 1.** *Let  $\{f(X_t) : t \geq 0\}$  be a continuous time process for which  $\rho_s$  exists for all  $s \geq 0$ . Assume that the curve  $\mathcal{K}^s : [0, \infty) \rightarrow \mathcal{H}_f^{\text{cont}}$  is continuous in  $s$ . Let  $\Delta t > 0$  be a time step. Assume that*

$$\begin{aligned} & \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(X_t) \bar{f}(X_{t+s}) dt \\ &= \lim_{T \rightarrow \infty} \frac{\Delta t}{T} \sum_{\mathbb{N} \ni n=0}^{T/\Delta t} f(X_{n\Delta t}) \bar{f}(X_{n\Delta t+s}), \end{aligned} \quad (13)$$

then  $\mathcal{H}_f \hookrightarrow \mathcal{H}_f^{\text{cont}}$ . Let  $q$  be an eigenfrequency of the discrete-time operator  $\mathcal{K}^{\Delta t}$ , i.e. there exists  $h \in \mathcal{H}_f \hookrightarrow \mathcal{H}_f^{\text{cont}}$  so that  $\mathcal{K}^{\Delta t}h = e^{iq}h$ . Then there exists an integer  $k$ , and  $h_k \in \mathcal{H}_f^{\text{cont}}$ , so that

$$\mathcal{K}^s h_k = e^{i \frac{q+2k\pi}{\Delta t} s} h_k \quad (14)$$

for all  $s \geq 0$ .

The proof of this proposition is given in the appendix. Proposition 1 guarantees that for every eigenfrequency  $q$  of  $\mathcal{K}^{\Delta t}$ , there always exists an eigenfrequency of  $\mathcal{K}^s$  for all  $s \geq 0$  which reduces to  $q$  at discrete time step.

Our discussion about continuous-time time series stops here. Hereinafter we always assume that the time series  $\{f(X_t)\}$  is discrete in time. For simplicity the time series is denoted by  $\{f(0), f(1), \dots\}$  and we use  $f$  to denote the whole time series generated by  $f$  as an element in  $\mathcal{H}_f$ . We use the notation  $\mathcal{K}$  to replace  $\mathcal{K}^{\Delta t}$ . Recall that the time-shift operator  $\mathcal{K}$  is an isometry on the Hilbert space  $\mathcal{H}_f$  of time-series associated to observable  $f$ .

**Remark 2.** *Due to Birkhoff ergodic theorem, autocovariance exists when the dynamical system is ergodic and has a finite invariant measure.*

The following theorem provides a very useful result to decompose any isometric operator and the Hilbert space on which it acts into unitary and non unitary part.

**Theorem 1** (Wold decomposition). *Let  $\mathcal{H}$  be a Hilbert space and  $\mathcal{V}$  any isometry of  $\mathcal{H}$ . Then we have an orthogonal decomposition  $\mathcal{H} = \mathcal{H}_{\text{NU}} \oplus \mathcal{H}_{\text{U}}$ , and  $\mathcal{V} = \mathcal{V}_{\text{NU}} \oplus \mathcal{V}_{\text{U}}$ , such that  $\mathcal{H}_{\text{NU}} = \bigoplus_{s \in I} \mathcal{H}_s$ ,  $\mathcal{V}_{\text{NU}}$  acts on*

$\mathcal{H}_s$  for some index set  $I$  as a unilateral shift, i.e.  $\mathcal{V}_{\text{NU}}(v_0, v_1, \dots) = (0, v_0, v_1, \dots)$ . And  $\mathcal{V}_{\text{U}}$  acts on  $\mathcal{H}_{\text{U}}$  and is unitary.  $\mathcal{H}_{\text{NU}}$  is called the completely non unitary part of  $\mathcal{H}$  as it does not contain closed subspaces of  $\mathcal{H}$  on which  $\mathcal{V}$  acts as a unitary operator.

Wold theorem is a particular case of (Szökefalvi-Nagy–Foia s) theorem for contraction operator.

**Theorem 2** (Szökefalvi-Nagy–Foias). *Let  $\mathcal{T}$  be a contraction operator (i.e.  $\|\mathcal{T}\| \leq 1$ ) on a Hilbert space  $\mathcal{H}$  then*

$$\mathcal{H}_U := \bigcap_{k \geq 0} (\text{fix}(\mathcal{T}^k \mathcal{T}^{*k}) \cap \text{fix}(\mathcal{T}^{*k} \mathcal{T}^k))$$

*is the largest space among all closed  $\mathcal{T}$ -invariant and  $\mathcal{T}^*$ -invariant subspaces of  $\mathcal{H}$  on which  $\mathcal{T}$  restricts to a unitary operator. The orthogonal complement  $\mathcal{H}_U^\perp = \mathcal{H}_{NU}$  is the completely non unitary part of  $\mathcal{H}$ . Here  $\text{fix}(A)$  refers to the subspace spanned by all the invariant vectors of operator  $A$ .*

Theorem 1 implies an orthogonal decomposition  $f = f_{NU} + f_U$ . Note that  $\mathcal{H}_{f,NU}$  and  $\mathcal{H}_{f,U}$  are invariant under  $\mathcal{K}$  and  $\mathcal{H}_{f,NU} \perp \mathcal{H}_{f,U}$ . Then the fact that  $\mathcal{H}_f$  is generated by  $f$  implies that  $\mathcal{H}_{f,U} = \overline{\text{Span}_{\mathbb{C}}\{f_U, \mathcal{K}f_U, \dots\}}$ . Note that for any eigenvector  $h \in \mathcal{H}_f$  of  $\mathcal{K}$  associated to eigenvalue  $\lambda$ , we have an orthogonal decomposition  $h = h_{f,U} + h_{f,NU}$ , where  $h_{f,U} \in \mathcal{H}_{f,U}$ ,  $h_{f,NU} \in \mathcal{H}_{f,NU}$ . Then  $\mathcal{K}h = \lambda h = \mathcal{K}h_{f,U} + \mathcal{K}h_{f,NU}$ . Because  $\mathcal{H}_{f,U}$  and  $\mathcal{H}_{f,NU}$  are invariant subspaces,  $h = \mathcal{K}h_{f,U}/\lambda + \mathcal{K}h_{f,NU}/\lambda$  is an orthogonal decomposition for which  $\mathcal{K}h_{f,U}/\lambda \in \mathcal{H}_{f,U}$ ,  $\mathcal{K}h_{f,NU}/\lambda \in \mathcal{H}_{f,NU}$ , implying that  $h_{f,U}$  and  $h_{f,NU}$  are both  $\mathcal{K}$ -eigenvectors of the same eigenvalue.  $h_{f,NU}$  must be zero because  $\mathcal{K}$  acts on  $\mathcal{H}_{f,NU}$  as the unilateral shift operator. Hence a  $\mathcal{K}$ -eigenvector must be inside  $\mathcal{H}_{f,U}$ .

**Definition 1.** *A Hilbert space  $\mathcal{H}$  with an unitary operator  $U$  is called  $f$ -cyclic if  $\mathcal{H} = \overline{\text{Span}_{\mathbb{C}}\{f, Uf, \dots\}}$  for some  $f \in \mathcal{H}$ .*

**Theorem 3** (Spectral theorem for unitary operator). *Let  $\mathcal{H}$  be a Hilbert space and  $U$  an unitary operator on  $\mathcal{H}$ . Assume that  $\mathcal{H}$  is  $f$ -cyclic for some  $f \in \mathcal{H}$ . Then there exists a finite measure  $\nu_f$  on the unit circle  $S^1 \subset \mathbb{C}$ , and an isomorphism  $\phi$*

$$\phi : \mathcal{H} \rightarrow L^2(S^1, d\nu_f) \quad (15)$$

*such that  $\phi \circ U \circ \phi^{-1}(g)(z) = zg(z)$ , for any  $g \in L^2(S^1, d\nu_f)$  and any  $z \in S^1$ . In particular,  $\phi(f) = 1$ .*

See lemma 5.4 in [26] for a mathematical proof. Note that lemma 5.4 in [26] assumes that  $\mathcal{H} = \overline{\text{Span}_{\mathbb{C}}\{\dots, U^{-1}f, f, Uf, \dots\}}$ , which is a weaker assumption than  $\mathcal{H}$  being  $f$ -cyclic in the sense of Definition 1. Therefore Theorem 3 applies to  $\mathcal{H}_{f,U}$  and  $\mathcal{K}$ . In general, the spectrum measure  $\nu$  consists of the discrete component, the singular-continuous component, and the absolutely continuous component (with respect to Lebesgue measure):  $\nu = \nu_d + \nu_{sc} + \nu_{ac}$ . The three components are pairwise-orthogonal, in the sense that for any  $g \in L^2(S^1, d\nu)$ , we can write  $g = g_d + g_{sc} + g_{ac}$  such that  $\nu_d(g_d) = \nu(g_d)$ ,  $\nu_d(g_{sc}) = \nu_d(g_{ac}) = 0$ , similarly for  $\nu_{sc}$  and  $\nu_{ac}$ . Together with theorem 1, this suggests the

orthogonal decomposition  $\mathcal{H}_f = \mathcal{H}_{f,d} + \mathcal{H}_{f,NU} + \mathcal{H}_{f,sc} + \mathcal{H}_{f,ac}$ , and

$$f = f_{NU} + f_d + f_{sc} + f_{ac}. \quad (16)$$

It is easy to see that  $\mathcal{H}_{f,d}$ ,  $\mathcal{H}_{f,NU}$ ,  $\mathcal{H}_{f,sc}$  and  $\mathcal{H}_{f,ac}$  are invariant subspaces of  $\mathcal{K}$ . In particular, the discrete part  $\nu_d$  is a finite or countable sum of Dirac measures  $\nu_d = \sum_i |a_i|^2 \delta_{\xi_i}$ , where  $\xi_i \in S^1$  is the support of  $\delta_{\xi_i}$ . Hence we can write

$$\phi(f_d) = \sum_i \mathbf{1}_{\xi_i}, \quad (17)$$

where  $\mathbf{1}_{\xi_i}(z) = 1$  if  $z = \xi_i$  and 0 otherwise.

As such, for every  $\xi_i \in \text{Supp}(\nu_d)$ ,  $\phi^{-1}(\mathbf{1}_{\{\xi_i\}})$  is an eigenvector of  $\mathcal{K}$ . On the other hand, let  $h \in \mathcal{H}_{f,U}$  be an eigenvector of  $\mathcal{K}$ , i.e.  $\mathcal{K}h = \xi h$  for some  $\xi \in S^1$ . Then  $\|\xi \phi(h) - z \phi(h)\|^2 = 0$ . Let  $A = \{z : z \phi(h)(z) \neq \xi \phi(h)(z)\}$ , then  $\nu(|\mathbf{1}_A \phi(h)|^2) = 0$ , meaning that  $\mathbf{1}_A \phi(h) = 0$  in  $L^2(S^1, d\nu)$ . Hence  $\phi(h) = \mathbf{1}_{\{\xi\}} \phi(h)$ , and  $\nu(\{\xi\}) > 0$ . This shows that there is a one-to-one correspondence between the support of the discrete measure  $\nu_d$  and the  $\mathcal{K}$ -eigenvectors inside  $\mathcal{H}_{f,U}$ . In particular, all the eigenvectors of  $\mathcal{K}$  are simple.

Let  $v_i$  be the corresponding normalized  $\mathcal{K}$ -eigenvectors, then we have

$$f = \sum_i a_i v_i + f_{sc} + f_{ac} + f_{NU}. \quad (18)$$

Our goal is to evaluate  $|a_i|^2$ , the numerical tool is the Gram matrix  $G_{NM}(f)$ , where the  $(i, j)$ -entry is

$$G_{NM,ij}(f) = \frac{1}{M} \sum_{t=0}^M f(i+t) \bar{f}(j+t). \quad (19)$$

We also define the autocovariance matrix

$$C_N(f) := \begin{pmatrix} \rho_0 & \rho_1 & \cdots & \rho_N \\ \bar{\rho}_1 & \rho_0 & \cdots & \rho_{N-1} \\ \vdots & & & \\ \bar{\rho}_N & \bar{\rho}_{N-1} & \cdots & \rho_0 \end{pmatrix} = \begin{pmatrix} \langle f, f \rangle & \langle f, \mathcal{K}f \rangle & \cdots & \langle f, \mathcal{K}^N f \rangle \\ \langle \mathcal{K}f, f \rangle & \langle \mathcal{K}f, \mathcal{K}f \rangle & \cdots & \langle \mathcal{K}f, \mathcal{K}^N f \rangle \\ \vdots & & & \\ \langle \mathcal{K}^N f, f \rangle & \langle \mathcal{K}^N f, \mathcal{K}f \rangle & \cdots & \langle \mathcal{K}^N f, \mathcal{K}^N f \rangle \end{pmatrix}$$

It is obvious that for any  $N \in \mathbb{N}$

$$\lim_{M \rightarrow \infty} G_{NM}(f) = C_N(f). \quad (20)$$

Recall that  $f = f_d + f_{NU} + f_{sc} + f_{ac}$  and that each of these four components is orthogonal to all other components. Hence

$$C_N(f) = C_N(f_d) + C_N(f_{NU}) + C_N(f_{sc}) + C_N(f_{ac}). \quad (21)$$

Let  $d_{NM,1} \geq d_{NM,2} \geq \dots \geq 0$  be the eigenvalues of  $G_{NM}(f)$ . Let  $d_{N,1} \geq d_{N,2} \geq \dots \geq 0$  be the eigenvalues of  $C_N(f)$ . It is clear that

$$\lim_{M \rightarrow \infty} d_{NM,i} = d_{N,i} \quad (22)$$

Our main result states that:

**Theorem 4** (Main result). *Assume that the autocovariance  $\rho_s$  exists for all  $s \in \mathbb{N}$ . Let  $\{v_i\}$  be the time shift operator eigenvectors of unit length. Let  $f \in \mathcal{H}_f$  the Hilbert space of observable time-series with  $f = \sum_{i=1}^{\infty} a_i v_i + f_{sc} + f_{ac} + f_{NU}$ , where  $f_{sc}$  and  $f_{ac}$  are the components of  $f$  in the space spanned by the singular-continuous spectrum and absolute-continuous spectrum, and  $f_{NU}$  the component of  $f$  in the completely non unitary subspace (i.e. direct sum of unilateral shift spaces). Assume that  $|a_1| \geq |a_2| \geq \dots \geq 0$ . Then for any  $i$ :*

$$\lim_{N \rightarrow \infty} \lim_{M \rightarrow \infty} \frac{d_{NM,i}}{N} = |a_i|^2. \quad (23)$$

For a given observable  $f$ , the trajectory matrix is defined as:

$$A_{NM}(f) = \begin{pmatrix} f(0), & f(1), & \dots & f(M) \\ f(1), & f(2), & \dots & f(M+1) \\ \vdots & & & \\ f(N), & f(N+1), & \dots & f(N+M) \end{pmatrix}. \quad (24)$$

Then it is clear that

$$G_{NM}(f) = \frac{1}{M} A_{NM}(f) A_{NM}(f)^* \quad (25)$$

where  $A^*$  refers to the conjugate transpose of  $A$ . Let  $\delta_{NM,1} \geq \delta_{NM,2} \geq \dots \geq 0$  be the singular values of  $A_{NM}(f)$ . Then directly we have that  $\delta_{NM,i} = \sqrt{M d_{NM,i}}$ .

**Corollary 1** (Trajectory matrix version). *Assume  $\rho_s$  exists for all  $s \in \mathbb{N}$ . Then*

$$\lim_{N \rightarrow \infty} \lim_{M \rightarrow \infty} \frac{\delta_{NM,i}^2}{NM} = |a_i|^2. \quad (26)$$

**Remark 3.** *The Gramian matrix is used by singular spectrum analysis methods ([22]) to construct temporal modes of the given time series. The eigenfunctions of  $G_{NM}$  are called temporal empirical orthogonal functions (EoFs). Theorem 4 implies that the leading temporal EoFs are due to theoretical eigenfrequencies of  $\mathcal{K}$ . Similarly, the data-adaptive harmonic decomposition (DAHD, [23]) and Hankel alternative view of Koopman analysis (HAVOK, [24]) are based on Gramian matrix and trajectory matrix, respectively. The main theorem and the corollary directly provides a way to identify which features extracted by SSA, DAHD, or HAVOK are related to eigenfrequencies of  $\mathcal{K}$ , and which features are not.*

The main result can be summarized as the following abstract mathematical theorem with respect to isometric operators Hilbert space. According to our knowledge, this mathematical result is not shown in any previous literature.

**Theorem 5** (Main result in pure mathematical form). *Let  $\mathcal{H}$  be a Hilbert space and  $\mathcal{K}$  an isometry of  $\mathcal{H}$ . The inner product in  $\mathcal{H}$  is denoted by  $\langle \cdot, \cdot \rangle$ . Let  $f \in \mathcal{H}$  be any vector. Let*

$$C_N(f) = \begin{pmatrix} \langle f, f \rangle & \langle f, \mathcal{K}f \rangle & \dots & \langle f, \mathcal{K}^N f \rangle \\ \langle \mathcal{K}f, f \rangle & \langle \mathcal{K}f, \mathcal{K}f \rangle & \dots & \langle \mathcal{K}f, \mathcal{K}^N f \rangle \\ \vdots & & & \\ \langle \mathcal{K}^N f, f \rangle & \langle \mathcal{K}^N f, \mathcal{K}f \rangle & \dots & \langle \mathcal{K}^N f, \mathcal{K}^N f \rangle \end{pmatrix}. \quad (27)$$

Let  $v_i$  be the eigenvectors of  $\mathcal{K}$  in  $\mathcal{H}$ . And let

$$f = \sum_i a_i v_i + f^\perp \quad (28)$$

be the decomposition of  $f$ , where  $f^\perp$  is perpendicular to the subspace spanned by all the eigenvectors of  $\mathcal{K}$ . There maybe uncountably many  $v_i$  in  $\mathcal{H}$  but only countably many are included in the summation. Let  $d_{N,1} \geq d_{N,2} \geq \dots \geq 0$  be the eigenvalues of  $C_N$ . And assume that  $|a_1| \geq |a_2| \geq \dots \geq 0$ . Then

$$\lim_{N \rightarrow \infty} \frac{d_{N,i}}{N} = |a_i|^2 \quad (29)$$

### III. PROOF OF THE MAIN THEOREM

We first present several lemmas which are independent of the language of Koopman theory.

**Lemma 1.** *Let  $\mathcal{T}$  be a contraction on a Hilbert space  $\mathcal{H}$ . Then for every  $f, g \in \mathcal{H}_{NU}$*

$$\lim_{n \rightarrow \infty} \langle \mathcal{T}^n f, g \rangle_{\mathcal{H}} = 0$$

*Proof of lemma 1.* For every  $f \in \mathcal{H}$  the sequence  $(\|\mathcal{T}^n f\|)_{n \in \mathbb{N}}$  is decreasing thus convergent. For any  $k \in \mathbb{N}$ , we have

$$\begin{aligned} \|\mathcal{T}^{*k} \mathcal{T}^k \mathcal{T}^n f - \mathcal{T}^n f\|^2 &= \|\mathcal{T}^{*k} \mathcal{T}^k \mathcal{T}^n f\|^2 - \\ &2\text{Re}(\langle \mathcal{T}^{*k} \mathcal{T}^k \mathcal{T}^n f, \mathcal{T}^n f \rangle) + \|\mathcal{T}^n f\|^2 \\ &= \|\mathcal{T}^{*k} \mathcal{T}^{n+k} f\|^2 - 2\|\mathcal{T}^{n+k} f\|^2 + \|\mathcal{T}^n f\|^2 \\ &\leq \|\mathcal{T}^{n+k} f\|^2 - 2\|\mathcal{T}^{n+k} f\|^2 + \|\mathcal{T}^n f\|^2 \\ &= \|\mathcal{T}^n f\|^2 - \|\mathcal{T}^{n+k} f\|^2 \rightarrow 0 \text{ as } n \rightarrow \infty \end{aligned}$$

Hence  $\langle (I - \mathcal{T}^{*k} \mathcal{T}^k) \mathcal{T}^n f, g \rangle_{\mathcal{H}} \rightarrow 0$  for every  $f, g \in \mathcal{H}$  as  $n \rightarrow \infty$ , therefore,

$\langle \mathcal{T}^n f, g \rangle_{\mathcal{H}} \rightarrow 0$  for every  $g \in \text{ran}(I - \mathcal{T}_c^{*k} \mathcal{T}_c^k)$  as  $n \rightarrow \infty$

The same argument for  $\mathcal{T}^*$  yields

$$\langle \mathcal{T}^n f, g \rangle_{\mathcal{H}} = \langle f, \mathcal{T}^{*n} g \rangle_{\mathcal{H}} \rightarrow 0 \text{ for every } g \in \text{ran}(\mathbb{I} - \mathcal{T}^k \mathcal{T}^{*k})$$

as  $n \rightarrow \infty$ .

We obtain that  $\langle \mathcal{T}^n f, g \rangle_{\mathcal{H}} \rightarrow 0$  as  $n \rightarrow \infty$  for every

$$\begin{aligned} f, g &\in \sum_{k \geq 1} (\text{ran}(\mathbb{I} - \mathcal{T}^{*k} \mathcal{T}^k) + \text{ran}(\mathbb{I} - \mathcal{T}^k \mathcal{T}^{*k})) \\ &= (\cap_{k \geq 1} (\text{ran}(\mathbb{I} - \mathcal{T}^{*k} \mathcal{T}^k)^\perp \cap \text{ran}(\mathbb{I} - \mathcal{T}^k \mathcal{T}^{*k})^\perp))^\perp \\ &= (\cap_{k \geq 1} (\text{fix}(\mathcal{T}^{*k} \mathcal{T}^k) \cap \text{fix}(\mathcal{T}^k \mathcal{T}^{*k})))^\perp = \mathcal{H}_U^\perp = \mathcal{H}_{NU} \end{aligned} \quad \square$$

**Lemma 2.** *Let  $\mathcal{H}$  be a Hilbert space over  $\mathbb{C}$  of finite dimension or infinite dimension.  $a_0, a_1, \dots \in \mathcal{H}$  such that  $\sum \|a_i\|^2 < \infty$ . Then*

$$\lim_{k \rightarrow \infty} \sum_i |\langle a_i, a_{i+k} \rangle| = 0 \quad (30)$$

*Proof of Lemma 2.* Without loss of generality, we may assume that  $\sum_i \|a_i\|^2 = 1$ . For any  $\epsilon > 0$ , there exists  $N$  such that  $\sum_{i \geq N} \|a_i\|^2 \leq \epsilon/2$ . Further there exists  $M > N$ , such that for any  $i < N$  and  $j > M$ ,  $\|a_j\| < \frac{\epsilon}{2} \|a_i\|$ . Now for any  $k > M$ ,

$$\begin{aligned} \sum_i |\langle a_i, a_{i+k} \rangle| &= \sum_{i=1}^N |\langle a_i, a_{i+k} \rangle| + \sum_{i > N} |\langle a_i, a_{i+k} \rangle| \\ &\leq \sum_{i \leq N} \|a_i\| \|a_{i+k}\| + \frac{1}{2} \sum_{i > N} (\|a_i\|^2 + \|a_{i+k}\|^2) \\ &\leq \frac{\epsilon}{2} \sum_{i \leq N} \|a_i\|^2 + \sum_{i > N} \|a_i\|^2 \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned} \quad (31) \quad \square$$

**Lemma 3** (The weakly mixing property). *Let  $\nu$  be a continuous measure on  $S^1$ . Then for any  $f, g \in L^2(S^1, d\nu)$  and  $z \in S^1$*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^N |\nu(z^i f \bar{g})|^2 = 0 \quad (32)$$

*Proof.* The proof of this mixing theorem can be found page 39 of [27]. An alternative proof is given in the appendix.  $\square$

**Lemma 4.** *Let  $L > 0$  and  $\xi_1, \dots, \xi_L \in S^1$  such that  $\xi_i \neq \xi_j$ , and  $c_1 \geq \dots \geq c_L > 0$ . Then*

$$\lim_{N \rightarrow \infty} \max_{\sum_{i=1}^N |\alpha_i|^2 = 1} \frac{1}{N} \sum_{k=1}^L c_k^2 \left| \sum_{i=0}^N \alpha_i \xi_k^i \right|^2 = c_1^2 \quad (33)$$

*Proof of lemma 4.* Let

$$E_N = \frac{1}{\sqrt{N+1}} \begin{pmatrix} 1 & 1 & \dots & 1 \\ \xi_1 & \xi_2 & \dots & \xi_L \\ \vdots & \vdots & \ddots & \vdots \\ \xi_1^N & \xi_2^N & \dots & \xi_L^N \end{pmatrix} \quad (34)$$

$$= (\Phi_{N,1}, \dots, \Phi_{N,L}), \quad (34)$$

$$\text{and } C = \text{diag}(c_1, \dots, c_L) \quad (35)$$

Then  $\|\Phi_{N,k}\| = 1$  and  $\langle \Phi_{N,k}, \Phi_{N,l} \rangle = \frac{1 - (\xi_k \xi_l^{-1})^{N+1}}{(N+1)(1 - \xi_k \xi_l^{-1})} \rightarrow 0$  as  $N \rightarrow \infty$ . Applying the Gram-Schmidt process to  $E_N$ , we get a matrix  $\tilde{E}_N = E_N S_N$ , so that the columns of  $\tilde{E}_N$  are orthogonal to each other and of unit length.  $S_N \rightarrow I_L$  as  $N \rightarrow \infty$ , where  $I_L$  refers to the identity matrix. Hence  $\lim_{N \rightarrow \infty} S_N C = C$ . As a consequence the leading singular value  $d_{N,1}$  of  $E_N C$  converges to  $c_1$ . On the other hand, for any  $\alpha = (\alpha_0, \dots, \alpha_N) \in \mathbb{C}^N$  with  $\|\alpha\| = 1$ , direct computation shows that

$$\|\alpha^\top E_N C\|^2 = \frac{1}{N+1} \sum_{k=1}^L c_k^2 \sum_{i=0}^N |\alpha_i \xi_k^i|^2. \quad (36)$$

Hence

$$\max_{\|\alpha\|=1} \frac{1}{N+1} \sum_{k=1}^L c_k^2 \sum_{i=0}^N |\alpha_i \xi_k^i|^2 = \|E_N C\|^2 = d_{N,1}^2 \rightarrow c_1^2 \quad (37)$$

$\square$

Now we can start to prove the main result. Recall that  $f = f_d + f_{sc} + f_{ac} + f_{NU}$ , and  $C_N(f) = C_N(f_{NU}) + C_N(f_d) + C_N(f_{ac}) + C_N(f_{sc})$ . For any semi positive-definite matrix  $A \in \mathbb{C}^{N \times N}$ , we define  $\|A\| = \max_{v \in \mathbb{C}^N} \frac{v^\top A v}{\|v\|^2}$ . The maximal eigenvalue  $d_1$  of  $A$  is equal to  $\|A\|$ . If Theorem 4 holds for  $i = 1$ , we can then recursively deduce Theorem 4 for all  $i$  by removing  $a_i v_i$  from  $f$  at each step. Since

$$\lim_{M \rightarrow \infty} G_{NM}(f) = C_N(f), \quad (38)$$

It is thus sufficient to prove that:

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{\|C_N(f_{NU})\|}{N} &= \lim_{N \rightarrow \infty} \frac{\|C_N(f_{ac})\|}{N} \\ &= \lim_{N \rightarrow \infty} \frac{\|C_N(f_{sc})\|}{N} = 0, \end{aligned} \quad (39)$$

and that

$$\lim_{N \rightarrow \infty} \frac{\|C_N(f_d)\|}{N} = |a_1|^2. \quad (40)$$

Now fix  $N$ , for any  $g \in \mathcal{H}_f$ ,

$$\|C_N(g)\| = \max_{\alpha} \frac{\|\sum_{i=0}^N \alpha_i \mathcal{K}^i g\|^2}{|\alpha_1|^2 + \dots + |\alpha_N|^2} \quad (41)$$

Hence Eq.(39) and (40) are equivalent to the following:

$$\lim_{N \rightarrow \infty} \max_{\alpha} \frac{\|\sum_{i=0}^N \alpha_i \mathcal{K}^i g\|^2}{N(|\alpha_1|^2 + \dots + |\alpha_N|^2)} = \begin{cases} |a_1|^2 & \text{if } g = f_d \\ 0 & \text{if } g = f_{sc}, f_{ac} \text{ or } f_{NU} \end{cases} \quad (42)$$

The case when  $g = f_{NU}$  can be quickly proved:

**Proposition 2** (The case for  $f_{NU}$ ).

$$\lim_{N \rightarrow \infty} \max_{\alpha} \frac{\|\sum_{i=0}^N \alpha_i \mathcal{K}^i f_{NU}\|^2}{N(|\alpha_1|^2 + \dots + |\alpha_N|^2)} = 0 \quad (43)$$

*Proof of proposition 2.* Without loss of generality, we may assume that  $\|\alpha\| = 1$ . Since  $f_{NU} \in \mathcal{H}_{NU} = \bigoplus_{s \geq 0} \mathcal{H}_s$ , we can write  $f_{NU} = (a_0, a_1, \dots)$ , where  $a_i \in \mathcal{H}_s$ . For  $k > 0$ , let  $c_k = |\langle \mathcal{K}^i f_{NU}, \mathcal{K}^{i+k} f_{NU} \rangle| = |\langle \mathcal{K}^{i+k} f_{NU}, \mathcal{K}^i f_{NU} \rangle|$ , which does not depend on  $i$ . Lemma 1 implies that  $\lim_{|i-j| \rightarrow \infty} c_{|i-j|} = 0$ . Therefore for any  $\epsilon > 0$ , there exists  $M_\epsilon$  such that  $c_{|i-j|} \leq \epsilon/4$  for any  $|i-j| > M_\epsilon$ . Now for any  $N > 2M_\epsilon \|f_{NU}\|^2/\epsilon$ , and any  $|\alpha_1|^2 + \dots + |\alpha_N|^2 = 1$ ,

$$\begin{aligned} \|\sum_{i=0}^N \alpha_i \mathcal{K}^i f_{NU}\|^2 &= \sum_{i,j} \alpha_i \bar{\alpha}_j \langle \mathcal{K}^i f_{NU}, \mathcal{K}^j f_{NU} \rangle \\ &\leq 2 \sum_{k=0}^N \sum_{i=0}^{N-k} \alpha_i \bar{\alpha}_{i+k} c_k \leq \sum_{k=0}^N \sum_{i=0}^{N-k} (|\alpha_i|^2 + |\alpha_{i+k}|^2) c_k \\ &\leq 2 \sum_{k=0}^N c_k \leq 2 \sum_{k=0}^{M_\epsilon} c_k + 2 \sum_{k > M_\epsilon} c_k \\ &\leq M_\epsilon \|f_{NU}\|^2 + (N - M_\epsilon) \epsilon / 2 \leq \frac{N\epsilon}{2} + \frac{N\epsilon}{2} = N\epsilon \quad (44) \end{aligned}$$

□

Recall the notations in Theorem 3, for any  $g \in \mathcal{H}_{f,U}$ ,

$$\begin{aligned} \|\sum_{i=0}^N \alpha_i \mathcal{K}^i g\|_{\mathcal{H}_f}^2 &= \|\sum_{i=0}^N \alpha_i z^i \phi(g)\|_{L^2(S^1, d\nu_f)}^2 \\ &= \int_{S^1} |\sum_{i=0}^N \alpha_i z^i|^2 |\phi(g)(z)|^2 d\nu_f(z) \quad (45) \end{aligned}$$

This proves that

**Proposition 3.** For any  $g \in \mathcal{H}_{f,U}$ ,

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{\|C_N(g)\|}{N} \\ = \lim_{N \rightarrow \infty} \max_{\|\alpha\|=1} \frac{1}{N} \int_{S^1} |\sum_{i=0}^N \alpha_i z^i|^2 |\phi(g)(z)|^2 d\nu_f(z) \quad (46) \end{aligned}$$

To prove Eq.(42) for  $g = f_d, f_{sc}$  and  $f_{ac}$ , we start with the following lemma.

**Lemma 5.** Let  $f, \mathcal{H}_{f,U}, \nu_f, \phi$  be the same as in Theorem 3. For simplicity, we denote  $\nu_f$  by  $\nu$ .  $\nu = \nu_d + \nu_{sc} + \nu_{ac}$ . Let  $\nu_{d,1}$  be a purely discrete finite measure on  $S^1$ , such that  $\{\xi_1, \dots, \xi_L\} = \text{Supp}(\nu_{d,1}) \subset \text{Supp}(\nu_d)$  and  $\nu_{d,1}(\{\xi_i\}) = \nu_d(\{\xi_i\})$  for any  $0 \leq i \leq L$ . Let  $c_k = \sqrt{\nu_d(\{\xi_k\})}$ . Let  $f_k = \phi^{-1}(\mathbb{1}_{\{\xi_k\}})$ , and set  $h = \sum_{k=1}^L f_k$ . Let  $d_{N,1}(h)$  be the leading eigenvalue of  $C_N(h)$ . Then

$$\lim_{N \rightarrow \infty} \frac{d_{N,1}(h)}{N} = \max_k c_k^2. \quad (47)$$

*Proof of proposition 5.* According to proposition 3,

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{d_{N,1}(h)}{N} &= \lim_{N \rightarrow \infty} \frac{\|C_N(h)\|}{N} \\ &= \lim_{N \rightarrow \infty} \max_{\|\alpha\|=1} \frac{1}{N} \int_{S^1} |\sum_{i=0}^N \alpha_i z^i|^2 |\phi(h)(z)|^2 d\nu_f(z) \\ &= \lim_{N \rightarrow \infty} \max_{\|\alpha\|=1} \frac{1}{N} \int_{S^1} |\sum_{i=0}^N \alpha_i z^i|^2 d\nu_{d,1}(z) \\ &= \lim_{N \rightarrow \infty} \max_{\|\alpha\|=1} \frac{1}{N} \sum_{k=1}^L c_k^2 |\sum_{i=0}^N \alpha_i \xi_k^i|^2 \quad (48) \end{aligned}$$

Then lemma 4 implies what we want to prove. □

**Proposition 4** (The case for  $f_d$ ). Eq.(40) holds.

*Proof of proposition 4.* For any  $\epsilon > 0$ , we choose a truncation  $\nu_d = \nu_{d,1} + \nu_{d,\epsilon}$ , so that  $\nu_{d,\epsilon}(S^1) < \epsilon$ ,  $|\text{Supp}(\nu_{d,1})| < \infty$ , and that  $\nu_{d,1}(\{\xi_k\}) = \nu_d(\{\xi_k\})$  whenever  $\xi_k \in \text{Supp}(\nu_{d,1})$ . This decomposition of measure also induces an orthogonal decomposition of  $L^2(S^1, d\nu_d) = L^2(S^1, d\nu_{d,1}) \oplus L^2(S^1, d\nu_{d,\epsilon})$ . Note that these two components are invariant under  $\mathcal{K}$ . When  $\epsilon$  is small enough,  $|a_1|^2 = \max_{\xi} \nu_{d,1}(\xi)$ .

Apply lemma 5 to  $\nu_{d,1}$ , and let  $h$  be defined as in lemma 5. Then  $f = h + f_\epsilon$  and

$$C_N(f) = C_N(h) + C_N(f_\epsilon), \quad (49)$$

and

$$\lim_{N \rightarrow \infty} \frac{\|C_N(h)\|}{N} = c_1^2 = \max_{\xi} \nu_d(\{\xi\}) = |a_1|^2 \quad (50)$$

And note that, applying Cauchy-Schwartz inequality,

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{\|C_N(f_\epsilon)\|}{N} \\ = \lim_{N \rightarrow \infty} \max_{\|\alpha\|=1} \int_{S^1} \frac{|\sum_{i=0}^N \alpha_i z^i|^2}{N} d\nu_{d,\epsilon} \leq \nu_{d,\epsilon}(S^1) < \epsilon \quad (51) \end{aligned}$$

Eq.(50)(51) implies Eq.(40) by letting  $\epsilon \rightarrow 0$ . □

**Proposition 5.** Let  $\nu_c$  be a continuous finite measure on  $S^1$ . Then

$$\lim_{N \rightarrow \infty} \max_{\|\alpha\|^2=1} \int_{S^1} \frac{1}{N} \left| \sum_{i=0}^N \alpha_i z^i \right|^2 d\nu_c = 0. \quad (52)$$

*Proof of proposition 5.* Let  $c_k = |\nu_c(z^k)|$ . In lemma 3, let  $f = g = 1$ , it implies that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^N c_k^2 = 0. \quad (53)$$

Therefore for  $\|\alpha\| = 1$ ,

$$\begin{aligned} \int_{S^1} \frac{1}{N} \left| \sum_{i=0}^N \alpha_i z^i \right|^2 d\nu_c &\leq \frac{2}{N} \sum_{k=0}^N \sum_{i=0}^{N-k} |\alpha_{i+k} \bar{\alpha}_i| c_k \\ &\leq \frac{1}{N} \sum_{k=0}^N \sum_{i=0}^{N-k} (|\alpha_{i+k}|^2 + |\alpha_i|^2) c_k \leq \frac{1}{N} \sum_{k=0}^N 2c_k \\ &\leq \frac{2}{N} \sqrt{(N+1) \sum_{k=0}^N c_k^2} = 2 \sqrt{\frac{N+1}{N} \frac{1}{N} \sum_{k=0}^N c_k^2} \rightarrow 0, \end{aligned} \quad (54)$$

as  $N \rightarrow \infty$ .  $\square$

**Corollary 2** (The case for  $f_{sc}$  and  $f_{ac}$ ). Eq.(39) holds for  $f_{sc}$  and  $f_{ac}$ .

*Proof.* This is the direct consequence of proposition 3 and 5.  $\square$

#### IV. ALGORITHM AND DISCUSSION

A direct application of the main theorem is to determine whether or not the given finite data set is sufficient enough for determining the  $i$ -th  $\mathcal{K}$ -eigenfrequency using Gramian matrix. For this purpose, we provide the following algorithm.

- Given a time series of data  $\{f(t)\}_{0 \leq t \leq T}$  where  $t, T$  are non-negative integers that represent the iteration number, choose  $N_k, M_{k,j}$  where  $1 \leq j \leq l_k$ , such that  $N_k + M_{k,j} \leq T$ ,  $M_{k,1} < M_{k,2} < \dots < M_{k,l_k} \gg N_k$ .
- For each  $N_k, M_{k,j}$ , compute the renormalized eigenvalues of  $G_{N_k M_{k,j}}$ , denoted by  $\sigma_{k,j,i} = \frac{d_{N_k M_{k,j}, i}}{N_k}$ .
- Given  $i$ , for each  $N_k$  check if  $\sigma_{k,j,i}$  converges as  $j$  increases. If for some  $k$  it does not converge, it means that the  $i$ -th  $\mathcal{K}$ eigenfrequency is not well represented by this dataset.
- Given  $i$ , if for all  $k$ ,  $\sigma_{k,j,i}$  shows good convergence, then check if  $\sigma_{k,l_k,i}$  converges as  $k$  increases. If  $\sigma_{k,l_k,i}$  converges to some nonzero number, then the

energy of the  $i$ -th  $\mathcal{K}$ -eigenfrequency is well represented by this data set. Otherwise, the  $i$ -th  $\mathcal{K}$ -eigenfrequency is not well-represented by this data set.

**Remark 4** (Identification of the  $\mathcal{K}$ -eigenfrequencies). Assume convergence for sufficiently enough  $i$ , choose  $G_{NM}$  so that  $d_{NM,i}/N$  is close enough to  $|a_i|^2$  for sufficiently many  $i$ . Assume that  $|a_{k-1}| > |a_k| = |a_{k+1}| = \dots = |a_{k+L}| > |a_{k+L+1}|$ .  $L$  must be finite because  $\|f\|^2 \geq \sum_i |a_i|^2$ . Let  $\xi_k, \dots, \xi_{k+L}$  be the corresponding theoretical  $\mathcal{K}$ eigenfrequencies. Our goal is to identify  $\xi_k, \dots, \xi_{k+L}$ . Let  $\eta_{k+i} = (1, \xi_{k+i}, \xi_{k+i}^2, \dots, \xi_{k+i}^N)$ . Let  $\{v_{NM,k}, v_{NM,k+1}, \dots, v_{NM,k+L}\}$  be the corresponding eigenvectors of  $G_{NM}$ . Then each of  $v_{NM,k}, \dots, v_{NM,k+L}$  is approximately a linear combination of  $\eta_k, \dots, \eta_{k+L}$ . Then  $\xi_{k+i}$  can be identified by applying Fourier analysis. In following two cases, the eigenfrequency can be approximated by counting the local maximums of  $v_{NM,i}$ .

- Case 1:  $f$  is a real valued observable. And for each  $v_i$  there is no eigen-vector except the conjugate of  $v_i$  that has the same energy as  $v_i$ ;
- Case 2: for each  $v_i$  there does not exist other eigen-vectors that has the same energy as  $v_i$ .

#### A. Implication to Hankel DMD

In [7] a Hankel DMD algorithm has been proposed and the authors showed that  $d_{N,i}$  can be used to identify Koopman and non-Koopman eigenfunctions for fixed  $N$  under the conditions that 1), the Hilbert space  $\mathcal{H}_f$  is finite dimensional and 2),  $N$  is larger than the dimension of  $\mathcal{H}_f$ . More precisely, they showed that  $d_{N,i} > 0$  if and only if  $d_{N,i}$  corresponds to a Koopman eigenfunction. However, this assumption is already too strong even for the case where  $f(x)$  is the observation of the first component of the 3 dimensional Lorenz system. In the case for which the dimension of  $\mathcal{H}_f$  is infinite, their method unfortunately fails. Because  $d_{N,i}$  can be positive even if there is no Koopman eigenfunctions. Therefore Theorem 4 can be thought of as a completion of the method posed in [7] under a much weaker assumption, by letting  $N \rightarrow \infty$ .

#### B. Comparison with Yosida's formula

Yosida mean ergodic theorem [25] provides a formula to calculate  $a_\omega$ , the coefficient of the Koopman eigenfunction of frequency  $\omega$  in Eq.(18):

$$a_\omega = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \exp(-2\pi i \omega t) f(t). \quad (55)$$

$a_\omega = 0$  if  $\omega$  is not a Koopman eigenfrequency. Under the assumption of ergodicity and finite invariant measure, this formula can be proved by combining Theorem 1, lemma 2, and Von-Neumann ergodic theorem. In the more general situation where only the existence of autocovariances is assumed, we do not know whether the limit in Eq.(55) always exists. Nor do we know if the output of Yosida's formula is strictly related to Koopman theory. This formula was first introduced to the fluid dynamics' community by [28, 29]. Eq.(55) is easy to compute for finite  $T$  and a given  $\omega$ . In the case for which the Koopman eigenfrequencies are unknown, numerically one still has the chance to identify some Koopman eigenfrequencies by calculating Eq.(55) for all  $\omega \in \{k\Delta\omega : k = 1, \dots, n\}$  and then finding the peak value.

On the other hand, from the theoretical point of view, when the system is ergodic and has a finite invariant measure, our result allows us to identify the Koopman eigenfunctions without having prior knowledge about the Koopman eigenfrequencies.

## V. NUMERICAL EXPERIMENTS

### A. Lorenz63 system

To first test the theorem-based methodology, we consider the Lorenz63 system. We integrate Lorenz system using the Runge-Kutta 4th order scheme with  $\Delta t = 0.01$ . As already mentioned in [17], this system is ergodic and has finite invariant measure. Hence the autocovariance always exist and  $\mathcal{H}_f$  can be identified with a subspace of  $L^2(X, d\mu)$  and  $\mathcal{K}$  coincides with the classical Koopman operator on  $\mathcal{H}_f$ . Due to its weakly mixing nature, the only Koopman eigenfunction of Lorenz 63 is the constant function which has frequency 0. Let  $f(t) = x - \bar{x}$ , where  $x$  is the first component of Lorenz system and  $\bar{x}$  is the temporal mean of  $x$ . We use  $E_{H,1}(N, M)$  to denote the leading renormalized singular value  $\frac{d_{NM,1}^2}{NM}$ . Then the decomposition  $f = \sum_i a_i v_i + f_{NU} + f_{sc} + f_{ac}$  can be reduced to  $f = f_{NU} + f_{sc} + f_{ac}$ . As expected, Fig.1 does not display the tendency that  $E_H(N, M_{\max})$  converge to some nonzero value as  $N \rightarrow \infty$ .

### B. A simple 4-dimensional system

Following one of the numerical examples in [17], we next consider a coupled system  $(X, T, \mu)$ , which consists of the discrete-time Lorenz system  $(X_{63}, T_{63}, \mu_{63})$  and a rotation on the unit circle  $(S^1, T_1, \mu_1)$ , i.e.  $X = X_{63} \times S^1$ ,  $T = T_{63} \times T_1$  and  $\mu = \mu_{63} \times \mu_1$ . It is outlined in [17] that  $\mu$  is an invariant measure. Still, the Lorenz system does not have non-trivial Koopman eigenfrequency and  $(X, T, \mu)$  is ergodic. Therefore the autocovariance exists and  $\mathcal{K}$  coincides with the classical Koopman operator..

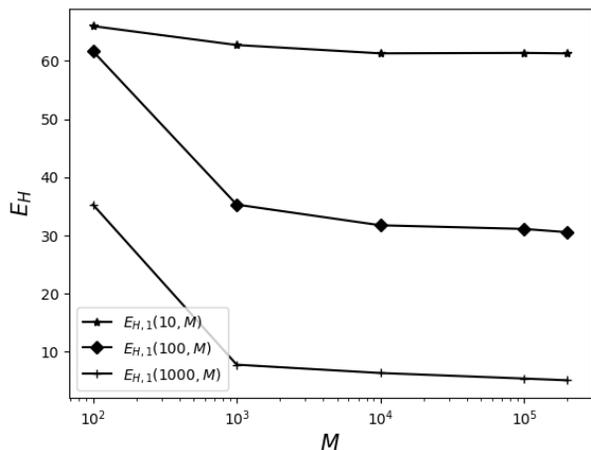


FIG. 1. The results about Lorenz 63 system. The behavior of  $E_{H,1}(N, M)$  agrees with the theoretical fact that  $E_H(N, M_{\max})$  does not converge to any nonzero value as  $N \rightarrow \infty$ .

We choose the rotation  $T_1$  to have period  $p = \pi/5$  and define the observable

$$f(x, y, z, \xi) = \sin(\xi + x/10). \quad (56)$$

For simplicity, we also use  $f(t)$  to denote  $f(x(t), y(t), z(t), \xi(t))$ . Then  $f = \sum_i a_i v_i + f_{sc} + f_{ac} + f_{NU}$  as in Eq.(18).

Anticipated by our main theorem, the renormalized singular value of the trajectory matrix should then converge to the same quantity as the one calculated by Eq.(55). Hence it is worth to make a numerical comparison about  $|a_\omega|^2$  obtained from Yosida's formula and that from the singular values of  $A_{NM}$ .

The integration time step for Lorenz system is  $\Delta t = 0.01$ . The Runge-Kutta 4th order scheme is applied for integrating Lorenz system. The frequency  $\omega$  we investigate is exactly the inverse of the period of  $(S^1, T_1)$ , i.e.  $\omega = 5\Delta t/\pi$  for Eq.(55). We use  $E_Y(T)$  to denote the value of  $|a_\omega|^2$  computed by Eq.(55), and  $E_{H,i}(N, M)$  to denote the value of  $\frac{d_{NM,i}^2}{NM}$  which is computed from the singular value decomposition of  $A_{NM}$ . Note that  $t, T, N, M$  are all integers which refer to the number of time steps instead of the exact time.

Fig.2 shows the numerical results of  $E_Y(T)$  and  $E_H(N, M)$ . The value  $E_Y(T)$  does not converge to 0, showing that  $\omega$  is indeed a Koopman eigen frequency. We also computed  $\|f\|^2 \approx \frac{1}{T} \sum_{t=1}^T |f(t)|^2 \approx 0.5002$ .  $E_Y(2 \times 10^5) \approx 0.1303$ , meaning that the fraction of energy in  $f$  represented by the Koopman eigenvector  $v_\omega$  is about 26%. Note that  $E_H(10^3, 2 \times 10^5)$  is close to  $E_Y(2 \times 10^5)$ , meaning that the leading singular value of the trajectory matrix indeed corresponds to the eigenfrequency  $\omega$ .

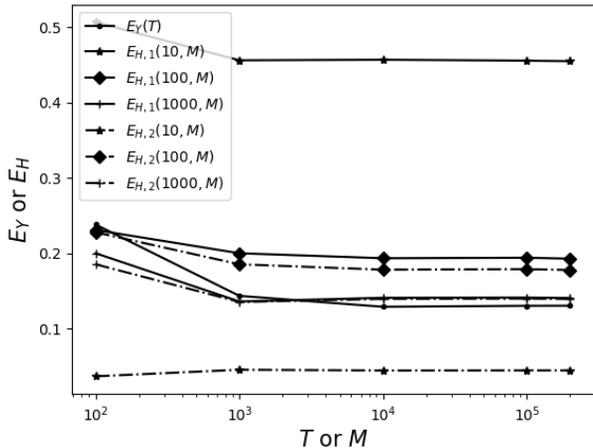


FIG. 2. The numerical results of  $E_Y(T)$  and  $E_{H,i}(N, M)$  for  $(X_{63} \times S^1, T_{63} \times T_1, \mu_{63} \times \mu_1)$ .

$E_{H,1}(10^3, M)$  and  $E_{H,2}(10^3, M)$  seem to converge to the same value. This is because the Koopman eigenfrequencies always exist in pair, i.e.  $\exp(2\pi i\omega) \in \text{Supp}(\nu_d) \iff \exp(-2\pi i\omega) \in \text{Supp}(\nu_d)$ . Since the observable  $f$  is real, the coefficient  $a_\omega = \bar{a}_{-\omega}$ . Therefore, the total fraction of energy in  $f$  that is represented by signals of period  $p = \frac{\pi}{5}$  is about 52%.

### C. AVISO (DUACS) interpolated ocean topography data (1993-2019)

For final illustration, we consider sea surface height (SSH) estimates. The AVISO gridded products provide the global SSH interpolation since 1993, the year after the launch of the first satellite altimeter TOPEX/Poseidon. The SSH is interpolated daily at a grid resolution of  $0.25^\circ \times 0.25^\circ$ . In this subsection, we use the main theorem to possibly assess the use of Koopman analysis for this dataset.

The assumption of the main theorem is the existence of autocovariance, which implies that the system should be stationary. We thus process the data by removing the overall constant rising tendency of SSH at each grid point over the decades (see for instance Fig.2 in [30]). We can not assert that the whole Earth system is ergodic and has an invariant measure, which includes the Earth, the atmosphere, ocean, all celestial bodies, but also the biology and living animals, etc. Hence we can not claim that the quantities  $a_i, v_i$ , etc. are associated to the classical Koopman operator. But, as we already stated, all these quantities are well-defined mathematically as long as the autocovariances exist. Moreover, in order to apply Yosida's formula, we assume that the eigenvectors  $v_i \in \mathcal{H}_f$  can be represented as a time series of the form  $\{1, e^{i\omega}, e^{2i\omega}, \dots\}$  for some  $\omega \in [0, 2\pi]$ . In this case, the output of Yosida's formula  $a_\omega$  must be  $a_i$ . We renor-

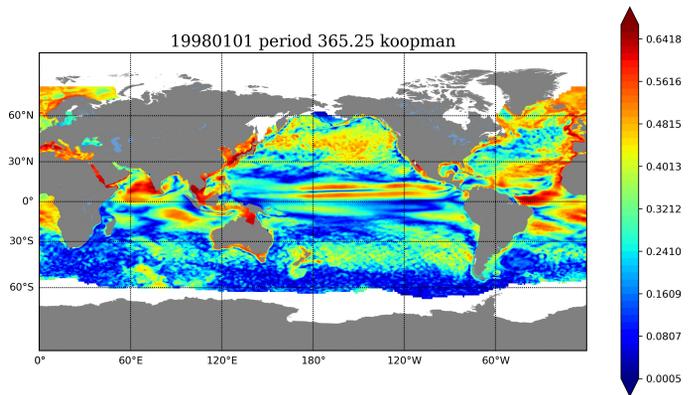


FIG. 3. The  $|a_\omega|$  value at each grid point computed at Jan. 1, 1998 using Eq.(55)

malize the SSH at each grid point, to simply ensure the data to have zero mean and unit variance at every grid point. We first apply Yosida's formula (Eq.55) to the global data to compute  $|a_\omega|$  at every grid point, where  $\omega = \exp(2\pi i/365.25)$ . This quantity is computed for January 1, 1998, i.e.  $f(1)$  refers to the SSH at Jan. 1, 1998. Note that in theory, i.e. assuming the autocovariance exists and the data set large enough, this quantity does not depend on time. Since the data now has unit variance,  $|a_\omega|^2$  can be interpreted as the fraction of energy in the SSH that is represented by the  $\mathcal{K}$ -eigenfrequency  $\omega$ . Similarly, since the SSH are real numbers,  $|a_\omega| = |a_{-\omega}|$  and  $|a_\omega|^2 + |a_{-\omega}|^2 = 2|a_\omega|^2$  represents the fraction of energy represented by the yearly signal.

Fig.3 shows that more than  $0.5^2 + 0.5^2 = 50\%$  of energy at the pacific ocean to the north of the equator (for instance at  $(114.875^\circ\text{W}, 6.125^\circ\text{N})$ ) is represented by the yearly signal. Constructing the trajectory matrix for SSH at  $(114.875^\circ\text{W}, 6.125^\circ\text{N})$ , i.e. we choose  $f = \text{SSH}(114.875^\circ\text{W}, 6.125^\circ\text{N})$ , we can then compare  $E_Y(T)$  and  $E_{H,i}(N, M)$ , for  $i = 0, 1, 2, 3$ ,  $N = p, 3p, 6p$ ,  $M = 3p, 6p, 20p$ , with  $p = 1$  year = 365.25(days). Fig.4 shows that the first two renormalized singular values apparently converge to the fraction of energy represented by  $\mathcal{K}$ -eigenfrequencies  $\omega$  and  $-\omega$ . However, the third and fourth renormalized singular values do now show a sign of convergence. As shown in Fig.5, this is likely due to the overall limited length of the present-day data set regarding the high dimensional state space of the dynamical system at stake.

## VI. CONCLUSION

The main objective of this study is to provide a rigorous and practical method to identify Koopman eigenfrequencies for discrete-time ergodic and (finite) measure preserving dynamical systems. In the more general situation where we only assume that the time series has well-defined autocovariances, we define the Hilbert space

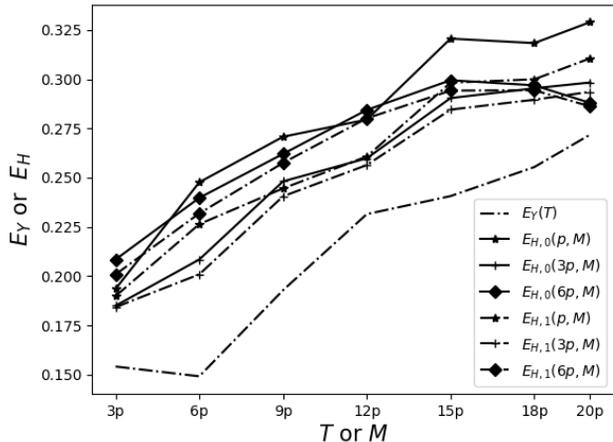


FIG. 4. The numerical results of  $E_Y(T)$  and  $E_{H,i}(N, M)$  ( $i=0,1$ ) for AVISO interpolated SSH at  $(114.875^\circ\text{W}, 6.125^\circ\text{N})$ . It shows that the first and the second singular value have possibly converged.

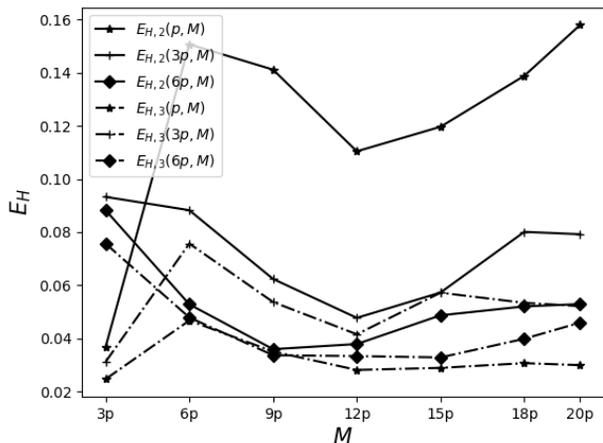


FIG. 5. The numerical results of  $E_Y(T)$  and  $E_{H,i}(N, M)$  ( $i=2,3$ ) for AVISO interpolated SSH at  $(114.875^\circ\text{W}, 6.125^\circ\text{N})$ . It shows that the third and the fourth singular value have not yet converged.

$\mathcal{H}_f$  purely based on the time series and the time shift operator  $\mathcal{K}$  that acts on  $\mathcal{H}_f$ . When the system is ergodic and has finite invariant measure,  $\mathcal{H}_f$  can be identified with the closure of the Krylov subspace generated by an observable  $f$ , and the time-shift operator  $\mathcal{K}$  coincides with the classical Koopman operator on the observable space  $\mathcal{H}_f$ . This work follows the result in [7], but further extend the applicability of the Hankel-DMD. It provides a theorem-based practical way to help assess the results of the decomposition in terms of Koopman eigenfrequencies. It shows that the leading temporal EoFs, which are calculated from the eigen decomposition of the Gramian

matrix, are indeed due to intrinsic eigenfrequencies. The main theorem provides partial theoretical foundation to several existing empirical methods including SSA, DAHD, and HAVOK. The main result shows that the discrete spectrum of  $\mathcal{K}$  can be characterized by the singular values (eigenvalues) of trajectory (Gramian, respectively) matrix. It remains to study whether the continuous spectrum can also be characterized by these matrices.

The numerical illustrations demonstrate the applicability of the theorem-based methodology for low dimensional systems. Yet, using sea surface height observables to inform about a very large dimension dynamical planet system, it is also apparent that one major difficulty of applying the main theorem might be the length of the data-set. An heuristic solution is to possibly associate the observables at different grid points, and/or to consider multiple observables, i.e. sea surface temperature. We reserve these investigations for future studies.

## ACKNOWLEDGEMENT

The authors acknowledge the support of the ERC EU project 856408-STUOD, the support of the ANR Melody project, the support from China Scholarship Council, and the support from the National Natural Science Foundation of China (Grant No. 42030406).

## Appendix A: An alternative proof of the weakly mixing property

In this appendix we provide an alternative proof of the weakly mixing property. Note that the proof of mixing theorem on page 39 of [27] implies that the weakly mixing property is equivalent to

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^N \left| \int_{S^1} z^i d\mu(z) \right|^2 = 0, \quad (\text{A1})$$

for any continuous measure  $\mu$ , which apparently is equivalent to

$$\lim_{n \rightarrow \infty} \frac{\#\{0 < k < n : \left| \int_{S^1} z^k d\mu(z) \right| > \epsilon\}}{n} = 0, \quad (\text{A2})$$

for any  $\epsilon > 0$  and continuous measure  $\mu$ . We shall provide an alternative proof for proposition 5 and then derive Eq.(A2) from proposition 5. To do this, we need the following lemma.

**Lemma 6.** *Let  $\nu_c$  be a continuous finite measure on  $S^1$ , and  $\{A^{(i)}\}_{i=1}^\infty$  a sequence of subsets of  $S^1$  such that  $\nu_c(A^{(i)}) > \epsilon$  for some fixed  $\epsilon > 0$  and for any  $i$ . Then for any  $L > 0$ , there exists  $\xi_1, \dots, \xi_L \in S^1$ , and a subsequence  $\{A^{(i_k)}\}_{k=1}^\infty$ , such that  $\xi_j \in A^{(i_k)}$  for any  $j$  and  $k$ .*

*Proof of lemma 6.* The idea of the proof is that we first show that there exists a point  $\xi_1 \in S^1$  and  $\Lambda_1 \subset \mathbb{N}$  such

that  $|\Lambda_1| = \infty$  and  $\xi_1 \in A^{(i)}$  for any  $i \in \Lambda_1$ . Then we choose a small neighborhood  $I_1$  of  $\xi_1$  so that  $\nu_c(I_1) < \epsilon/2$ . This can be done merely because  $\nu_c$  is a continuous measure. Let  $A^{(i),1} = A^{(i)} - I_1$  for  $i \in \Lambda_1$ , we have  $\nu_c(A^{(i),1}) > \epsilon/2$ . Then we apply the same analysis to  $\{A^{(i),1}\}_{i \in \Lambda_1}$  to find  $\xi_2$ , etc. After doing the same analysis for  $L$  times, we get  $\xi_1, \dots, \xi_L$  and  $\Lambda_L$  such that  $|\Lambda_L| = \infty$  and  $\xi_j \in A^{(i)}$  for any  $j$  and  $i \in \Lambda_L$ .

To prove that there exists a point  $\xi_1 \in S^1$  and  $\Lambda_1 \subset \mathbb{N}$  such that  $|\Lambda_1| = \infty$  and  $\xi_1 \in A^{(i)}$  for any  $i \in \Lambda_1$ . We prove by contradiction. Suppose that this is not true, i.e. for any  $\xi \in S^1$  there exists  $N$  such that  $\xi \notin A^{(i)}$  for any  $i > N$ . Let  $B_N = \cup_{i \geq N} A^{(i)}$ . Then  $B_1 \supset B_2 \supset \dots$  and  $\cap_{N \geq 1} B_N = \emptyset$ . It means that  $\nu_c(B_N) \rightarrow 0$  as  $N \rightarrow \infty$ . This is apparently not true because  $\nu_c(B_N) \geq \nu_c(A^{(N+1)}) > \epsilon$ .  $\square$

Now we give another proof of proposition 5 based on lemma 6.

*An alternative proof of proposition 5.* We prove by contradiction. We assume that proposition 5 does not hold. Then there exists  $\epsilon > 0$  and a sequence of  $\alpha^{(i)} = \{\alpha_1^{(i)}, \dots, \alpha_{n_i}^{(i)}\}$ , such that  $\|\alpha^{(i)}\|^2 = 1$  and

$$\nu_c(|P^{(i)}|^2/n_i) > \epsilon, \quad (\text{A3})$$

where  $P^{(i)}(z) = \sum_{j=0}^{n_i} \alpha_j^{(i)} z^j$ . Note that  $|P^{(i)}(z)|^2/n_i \leq 1$  for any  $z \in S^1$ . Let

$$A^{(i)} = \{z \in S^1 \mid |P^{(i)}(z)|^2 > \frac{\epsilon}{2} n_i\}. \quad (\text{A4})$$

Then  $\nu_s(A^{(i)}) \geq \frac{\epsilon}{2}$ .

Now we pick any  $L > 4/\epsilon$ . Lemma 6 implies that we can find  $\xi_1, \dots, \xi_L \in S^1$ ,  $\xi_i \neq \xi_j$  for  $i \neq j$ , and a subsequence  $A^{(i_j)}$  of  $A^{(i)}$ , such that  $\xi_k \in A^{(i_j)}$  for any  $k, j$ . Then for any  $i_j$ , set  $N = n_{i_j}$  and  $\alpha = \alpha^{(i_j)}$ ,

$$\sum_{k=1}^L \frac{1}{N} \left| \sum_{i=0}^N \alpha_i \xi_k^i \right|^2 \geq \sum_{k=1}^L \frac{1}{N} \frac{\epsilon}{2} N \geq 2 \quad (\text{A5})$$

This contradicts with lemma 4 by setting  $c_1 = c_2 = \dots = c_L = 1$ .  $\square$

Now we derive Eq.(A2) from proposition 5.

**Corollary 3.** *Let  $\nu_c$  be a continuous finite measure on  $S^1$ . Then for any  $\epsilon > 0$ ,*

$$\lim_{n \rightarrow \infty} \frac{\#\{0 < k < n : |\int_{S^1} z^k d\nu_c(z)| > \epsilon\}}{n} = 0 \quad (\text{A6})$$

*Proof.* For any  $n$ , let  $M_n = \#\{0 < k < n : |\int_{S^1} z^k d\nu_c(z)| > \epsilon\}$ . For any  $k \leq n$ , pick  $\beta_k \in S^1$  so that

$\int_{S^1} \beta_k z^k d\nu_c(z) > 0$ . Let  $\alpha_k = \beta_k/\sqrt{n}$ . Then  $\|\alpha\| = 1$  and

$$\int_{S^1} \left| \sum_{i=0}^n \alpha_i z^i \right|^2 d\nu_c(z) \geq \frac{1}{n} \left| \int_{S^1} \sum_{i=0}^n \beta_i z^i d\nu_c(z) \right|^2 \geq \frac{M_n^2 \epsilon^2}{n}$$

Then proposition 5 implies that

$$\lim_{n \rightarrow \infty} \frac{M_n^2 \epsilon^2}{n^2} = 0. \quad (\text{A7})$$

Hence

$$\lim_{n \rightarrow \infty} \frac{M_n}{n} = 0. \quad (\text{A8})$$

$\square$

## Appendix B: Proof of proposition 1

Proposition 1: Assume that the curve  $\mathcal{K}^s : [0, \infty) \rightarrow \mathcal{H}_f^{\text{cont}}$  is continuous in  $s$ . Let  $q$  be an eigenfrequency of the discrete-time operator  $\mathcal{K}^{\Delta t}$ , i.e. there exists a time-series  $h \in \mathcal{H}_f \leftrightarrow \mathcal{H}_f^{\text{cont}}$  so that  $\mathcal{K}^{\Delta t} h = e^{iq} h$ . Then there exists at least an integer  $k$ , and  $0 \neq h_k \in \mathcal{H}_f^{\text{cont}}$ , so that

$$\mathcal{K}^s h_k = e^{i \frac{q+2k\pi}{\Delta t} s} h_k \quad (\text{B1})$$

for all  $s \geq 0$ .

*Proof of proposition 1.* Consider  $g_s \in \mathcal{H}_f^{\text{cont}}$  where  $g_s = e^{-is \frac{q}{\Delta t}} \mathcal{K}^s h$ . Since  $\mathcal{K}^{\Delta t} h = e^{iq} h$ ,  $g_{\Delta t} = g_0 = h$ . It is easy to show that  $\|g_s\| = \|h\|$  for any  $s \geq 0$ . So we have a closed loop in  $\mathcal{H}_f^{\text{cont}}$ :

$$g : [0, \Delta t] \rightarrow \mathcal{H}_f^{\text{cont}}. \quad (\text{B2})$$

Now we do Fourier decomposition to this circle, i.e. for every integer  $k$ , we define

$$h_k = \frac{1}{\sqrt{\Delta t}} \int_0^{\Delta t} e^{-i \frac{2k\pi}{\Delta t} s} g_s ds. \quad (\text{B3})$$

Parseval's theorem implies that

$$\Delta t \|h\|^2 = \int_0^{\Delta t} \|g_s\|^2 ds = \sum_{k=-\infty}^{+\infty} \|h_k\|^2. \quad (\text{B4})$$

Therefore there exists an integer  $l$ , such that  $h_l \neq 0$ . Now

for any  $t \geq 0$ ,

$$\begin{aligned}
\mathcal{K}^s h_l &= \frac{1}{\sqrt{\Delta t}} \mathcal{K}^t \int_0^{\Delta t} e^{-i\frac{2l\pi}{\Delta t}s} e^{-is\frac{q}{\Delta t}} \mathcal{K}^s h ds \\
&= \frac{1}{\sqrt{\Delta t}} \int_0^{\Delta t} e^{-i\frac{2l\pi+q}{\Delta t}s} s \mathcal{K}^{t+s} h ds \\
&= e^{i\frac{2l\pi+q}{\Delta t}t} \frac{1}{\sqrt{\Delta t}} \int_0^{\Delta t} e^{-i\frac{2l\pi+q}{\Delta t}(t+s)} \mathcal{K}^{t+s} h ds \\
&= e^{i\frac{2l\pi+q}{\Delta t}t} \frac{1}{\sqrt{\Delta t}} \int_0^{\Delta t} e^{-i\frac{2l\pi}{\Delta t}(t+s)} g_{t+s} ds \\
&= e^{i\frac{2l\pi+q}{\Delta t}t} \frac{1}{\sqrt{\Delta t}} \int_t^{t+\Delta t} e^{-i\frac{2l\pi}{\Delta t}s} g_s ds \\
&= e^{i\frac{2l\pi+q}{\Delta t}t} h_l.
\end{aligned} \tag{B5}$$

In other words,  $h_l$  is an eigen-vector of the continuous-time operator  $\mathcal{K}^s$  for any  $s \geq 0$ .  $\square$

- 
- [1] P. Schmid, *Journal of Fluid Mechanics* **656**, 5 (2010).
- [2] C. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. Henningson, *Journal of Fluid Mechanics* **641**, 115 (2009).
- [3] Y. Saad, *Linear Algebra and its Applications* **34**, 269 (1980).
- [4] K. K. Chen, J. Tu, and C. Rowley, *Journal of Nonlinear Science* **22**, 887 (2012).
- [5] A. Wynn, D. S. Pearson, B. Ganapathisubramani, and P. J. Goulart, *Journal of Fluid Mechanics* **733**, 473–503 (2013).
- [6] J. Tu, C. Rowley, D. M. Luchtenburg, S. Brunton, and J. Kutz, *ACM Journal of Computer Documentation* **1**, 391 (2014).
- [7] H. Arbabi and I. Mezić, *SIAM J. Appl. Dyn. Syst.* **16**, 2096 (2017).
- [8] A. Kusaba, T. Kuboyama, and S. Inagaki, *Plasma and Fusion Research* **15**, 1301001 (2020).
- [9] J. Kutz, X. Fu, and S. Brunton, arXiv: Dynamical Systems (2015).
- [10] M. Williams, I. Kevrekidis, and C. Rowley, *Journal of Nonlinear Science* **25**, 1307 (2015).
- [11] J. Proctor, S. Brunton, and J. Kutz, *SIAM J. Appl. Dyn. Syst.* **15**, 142 (2016).
- [12] M. Hemati, C. Rowley, E. A. Deem, and L. Cattafesta, *Theoretical and Computational Fluid Dynamics* **31**, 349 (2017).
- [13] J. P. Taylor-King, A. N. Riseth, W. Macnair, and M. Claassen, *PLoS Computational Biology* **16** (2020).
- [14] M. Korda and I. Mezić, *Journal of Nonlinear Science* **28**, 687 (2018).
- [15] M. Korda, M. Putinar, and I. Mezić, *Applied and Computational Harmonic Analysis* **48**, 599 (2020).
- [16] M. O. Williams, C. W. Rowley, and I. G. Kevrekidis, *Journal of Computational Dynamics* **2**, 247 (2015).
- [17] S. Das and D. Giannakis, *Journal of Statistical Physics* **175**, 1107 (2017).
- [18] S. Das and D. Giannakis, arXiv: Dynamical Systems (2018).
- [19] D. Giannakis, S. Das, and J. Slawinska, arXiv: Dynamical Systems (2018).
- [20] D. Giannakis, *Research in the Mathematical Sciences* **8**, 1 (2020).
- [21] F. Nuske, S. Peitz, F. Philipp, M. Schaller, and K. Worthmann, arXiv (2021).
- [22] M. Ghil, M. R. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou, *Reviews of Geophysics* **40**, 3 (2002).
- [23] D. Kondrashov, E. Ryzhov, and P. Berloff, *Chaos* **30** **6**, 061105 (2020).
- [24] S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, and J. N. Kutz, *Nature Communications* **8** (2017).
- [25] K. Yosida, *Functional Analysis* (Springer, Berlin, Heidelberg, 1995).
- [26] D. Borthwick, *Spectral Theory: Basic Concepts and Applications*, 1st ed., Graduate Texts in Mathematics №284 (Springer International Publishing; Springer, 2020).
- [27] P. R. Halmos, *Lectures on Ergodic Theory* (Chelsea Publishing Company, New York, N.Y., 1956).
- [28] I. Mezić and A. Banaszuk, *Physica D: Nonlinear Phenomena* **197**, 101 (2004).
- [29] I. Mezić, *Nonlinear Dynamics* **41**, 309 (2005).
- [30] A. Cazenave and W. Llovel, *Annual review of marine science* **2**, 145 (2010).