To infer eigenvalues and eigenfunctions of the infinite-dimensional Koopman operator, we study the leading singular values of the Hankel matrix associated with a given observable of a dynamical system. We prove that for a discrete-time measure preserving ergodic dynamic system and a given square-integrable observable $f$, these leading singular values of the Hankel matrix have one-to-one correspondence with the energy of $f$ represented by the Koopman eigenfunctions. The proof is associated to the Birkhoff ergodic theorem, 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. 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 the leading temporal empirical orthogonal functions are indeed approximations of the Koopman eigenfunctions. A theorem-based practical methodology is then proposed to identify the Koopman eigenfunctions from a given time series. It builds on the fact that the convergence of the renormalized leading singular values of the Hankel matrix is a necessary and sufficient condition for the existence of Koopman eigenfunctions. 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 motivated 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] and [19]), the approximation of Koopman operator is performed by kernel methods. Recently, [19] showed the convergence of kernel methods for any measure preserving ergodic dynamical systems, the measure of which support lies on a compact manifold.

In this manuscript, we argue and prove that for any discrete-time measure preserving ergodic dynamical system and any square-integrable observable $f$, when the dimension of the Hankel matrix $A_{NM}(f)$ goes to infinity in the right order, the leading singular values renormal-
ized by the dimension of $A_{NM}(f)$ converge to the energy of $f$ that is represented by the Koopman eigenfunctions. All other renormalized singular values shall further converge to 0 uniformly. Despite its theoretical interests, the main theorem directly suggests a practical algorithm to explicitly identify the Koopman and non-Koopman eigenfunctions from given time series. As a by-product, it also shows that the leading temporal empirical orthogonal functions calculated by singular spectrum analysis (SSA, [20]) method are indeed Koopman eigenfunctions. Similarly, this theorem also sheds light on the theoretical foundation of data-adaptive harmonic decomposition (DAHD, [21]), and Hankel alternative view of Koopman analysis (HAVOK, [22]). Because all these methods are based either on Hankel matrix or on Gramian matrix.

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 present the detailed algorithm and compare it with another numerical method based on Yosida’s mean ergodic theorem ([23]). 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

We consider a discrete-time dynamical system $(X,T)$ which is ergodic and equipped with a finite invariant measure $\mu$. In this case, the Birkhoff ergodic theorem holds.

Theorem 1 (Birkhoff Ergodic Theorem). Let $(X,T)$ be an ergodic discrete-time dynamical system with a finite invariant measure $\mu$. Then for every $h \in L^1(X,d\mu)$, and for $\mu$–almost every $x \in X$,

$$\mu(h) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n} h(T^i(x)).$$

(1)

The (discrete) Koopman operator is defined as

$$K : L^2(X,\mu) \to L^2(X,\mu)$$

$$h \mapsto h \circ T.$$  

(2)

(3)

For any two observables $h_1, h_2 \in L^2(X,d\mu)$,

$$\langle h_1, h_2 \rangle = \mu(h_1 \bar{h_2}) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n} h_1 \bar{h_2}(T^i(x))$$

(4)

$$= \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n} h_1 \bar{h_2}(T^{i+1}(x)) = \mu(h_1 \bar{h_2} \circ T)$$

(5)

$$= \langle Kh_1, Kh_2 \rangle.$$  

Hence $K$ is an isometry. Since the dynamical system is ergodic, the eigen space of $\mathcal{K}$ of the same frequency must have dimension 1. Indeed, if $h \in L^2(X,\mu)$ such that $K^h = h$, let $A^{+(or -)} = \{x : h(x) > (or < )\mu(h)/\mu(X)\}$. $A^{+/−}$ are invariant sets as $T^{−1}(A^{+/−}) = A^{+/−}$. $T$ being ergodic implies that $A^{+/−} = X \setminus \emptyset$. But $\mu(h - \mu(h)/\mu(X)1_X) = 0$. Hence $A^{+/−} = \emptyset$. This shows that the eigenfunctions with eigenvalue 1 must be constant functions. Now let $h_1, h_2$ be two eigenfunctions associated with the same eigenvalue. Then $K(h_1 \bar{h}_2) = K_1 \bar{K}_2 = h_1 h_2$. Therefore by the previous implication $h_1 = c h_2$ for some constant $c$.

If we further assume that the system is invertible, then $K$ must be an isomorphism, hence unitary. Without this assumption, $K$ is not unitary. Still, we have the following useful result.

Theorem 2 (Wold decomposition). Let $\mathcal{H}$ be a Hilbert space and $\mathcal{K}$ an isometry of $\mathcal{H}$. Then we have an orthogonal decomposition $\mathcal{H} = \mathcal{H}_NU \oplus \mathcal{H}_U$, and $\mathcal{K} = \mathcal{K}_NU \oplus \mathcal{K}_U$, such that $\mathcal{H}_NU \oplus \mathcal{K}_NU$ acts on $\mathcal{H}_U$ for some index set $I$ as a unilateral shift, i.e. $\mathcal{K}_NU(v_0,v_1,...) = (0,v_0,v_1,...)$. And $\mathcal{K}_U$ acts on $\mathcal{H}_U$ and is unitary. $\mathcal{H}_NU$ is called the completely non unitary part of $\mathcal{H}$ as it does not contain closed subspaces of $\mathcal{H}$ on which $\mathcal{K}$ acts as a unitary operator.

Wold theorem is a particular case of (Szőkefalvi-Nagy–Foiaș’s) theorem for contraction operator.

Theorem 3 (Szőkefalvi-Nagy–Foiaș). Let $T$ be a contraction operator (i.e. $\|T\| \leq 1$) on a Hilbert space $\mathcal{H}$ then

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

is the largest space among all closed $T$-invariant and $T^*$-invariant subspaces of $\mathcal{H}$ on which $T$ 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 fix($A$) refers to the subspace spanned by all the invariant vectors of operator $A$.

Our goal is to analyze the Hankel matrix $A_{NM}(f)$:

$$A_{NM} = A_{NM}(f)$$

$$= \begin{pmatrix} f(x(0)), & f(x(1)), & \ldots & f(x(M)) \\ f(x(1)), & f(x(2)), & \ldots & f(x(M+1)) \\ \vdots & \vdots & \ddots & \vdots \\ f(x(N)), & f(x(N+1)), & \ldots & f(x(N+M)) \end{pmatrix},$$

$$\text{Span}_{c} \{f, Kf, \ldots \} \text{, where the closure is taken with respect to the strong topology of } L^2(X, d\mu).$$

Note that in principle $L^2(X, d\mu)$ could have a basis the cardinality
of which is uncountable. But $\mathcal{H}_f$ always has countably many or finite basis. The implication of the basis being uncountably many is that any finite dimensional approximation of the Koopman operator on the whole $L^2(X,d\mu)$ shall fail to converge in theory.

Theorem 2 implies an orthogonal decomposition $f = f_{NU} + f_U$, and it can be proved that $\mathcal{H}_{f,U} = \text{Span}_\mathbb{C}\{f_U, Kf_U, \ldots\}$. Note that for any Koopman eigenfunction $h$ of eigenvalue $\lambda$, we have an orthogonal decomposition $h = h_{f,U} + h_{f,NU} + h_0$, where $h_{f,U} \in \mathcal{H}_{f,U}$, $h_{f,NU} \in \mathcal{H}_{f,NU}$, $h_0 \in \mathcal{H}_f$. Then $K = \lambda h = \lambda h_{f,U} + \lambda h_{f,NU} + \lambda h_0$. Because $\mathcal{H}_{f,U}, \mathcal{H}_{f,NU}$ and $\mathcal{H}_f$ are invariant subspaces, $h = Kh_{f,U}/\lambda + Kh_{f,NU}/\lambda + h_0/\lambda$ is an orthogonal decomposition for which $Kh_{f,U}/\lambda \in \mathcal{H}_{f,U}$, $Kh_{f,NU}/\lambda \in \mathcal{H}_{f,NU}$, $h_0/\lambda \in \mathcal{H}_f$, implying that $h_{f,U}, h_{f,NU}$ and $h_0$ are all Koopman eigenfunctions of the same eigenvalue. $h_{f,NU}$ must be zero because $K$ acts on $\mathcal{H}_{f,NU}$ as the unilateral shift operator. The system is ergodic, and the eigen subspace corresponding to $\lambda$ has dimension 1. Hence a Koopman eigenfunction is either inside $\mathcal{H}_{f,U}$ or orthogonal to $\mathcal{H}_f$. It is possible for a Koopman eigenfunction to be orthogonal to $\mathcal{H}_f$. However, those Koopman eigenfunctions would not reveal themselves in time series analysis generated by the single observable $f$, which are done inside $\mathcal{H}_f$.

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

**Theorem 4** (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 $\phi: \mathcal{H} \rightarrow L^2(S^1, d\nu_f)$

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

See lemma 5.4 in [24] for a mathematical proof. Note that lemma 5.4 in [24] assumes that $\mathcal{H} = \text{Span}_\mathbb{C}\{\ldots U^{-1}f, Uf, \ldots\}$, which is a weaker assumption than $\mathcal{H}$ being $f$-cyclic in the sense of Definition 1. Therefore Theorem 4 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 two 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 $\nu_d(g_d) = \nu(g_d), \nu_{sc}(g_{sc}) = \nu_{ac}(g_{ac}) = 0$, similarly for $\nu_{sc}$ and $\nu_{ac}$. Together with theorem 2, this suggests the orthogonal decomposition of $f$

$$f = f_{NU} + f_d + f_{sc} + f_{ac}.$$

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

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

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

As such, for every $\xi_i \in \text{Supp}(\nu_d)$, $\phi^{-1}(1_{\xi_i})$ is an eigenfunction of $\mathcal{K}$. On the other hand, let $h \in \mathcal{H}_{f,U}$ be an eigenfunction of $\mathcal{K}$, i.e. $Kh = \xi h$ 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 $\nu(||\mathbb{A}\phi(h)||^2) = 0$, meaning that $1_{A}\phi(h) = 0$ in $L^2(S^1, d\nu)$. Hence $\phi(h) = 1_{\{\xi\}}(h)$, 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 Koopman eigenfunctions inside $\mathcal{H}_{f,U}$.

Let $v_i$ be the corresponding normalized Koopman eigenfunctions, then we have

$$f = \sum_i a_i v_i + f_{sc} + f_{ac} + f_{NU}.$$

And the Hankel matrix has the decomposition

$$A_{NM} = A_{NM}(f) = A_{NM}(f_{NU}) + A_{NM}(f_d) + A_{NM}(f_{sc}) + A_{NM}(f_{ac}).$$

Let $A_{NM} = U_{NM} \bar{D}_{NM} V_{NM}^\top$ be a singular value decomposition, where $D_{NM}$ is a real diagonal matrix with entries $d_{NM,1} \geq d_{NM,2} \geq \cdots > 0$. Our main result states that:

**Theorem 5** (Main result). Assume that the dynamical system is ergodic with a finite invariant measure. Let $\{v_i\}$ be the Koopman eigenfunctions of unit length. Let $f = \sum_i a_i v_i + f_{sc} + f_{ac} + f_{NU}$, 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 $|a_1| \geq |a_2| \geq \cdots \geq 0$. Then for any $i$:

$$\lim_{N \rightarrow \infty} \lim_{M \rightarrow \infty} \frac{d_{NM,1}^2}{NM} = |a_i|^2.$$

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

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

where $A^*$ refers to the conjugate transpose of $A$. Let $c_{NM,1} \geq c_{NM,2} \geq \cdots \geq 0$ be the eigenvalues of $G_{NM}(f)$. Then directly we have that $c_{NM,i} = d_{NM,i}^2/M$.

**Corollary 1** (Gramian matrix version). Assume that the dynamical system is ergodic with a finite invariant measure. Then

$$\lim_{N \rightarrow \infty} \lim_{M \rightarrow \infty} \frac{c_{NM,1}}{N} = |a_i|^2.$$
Remark 1. The Gramian matrix is used by singular spectrum analysis methods \cite{20} to construct temporal modes of the given time series. The eigenfunctions of $G_{NM}$ are called temporal empirical orthogonal functions (EoFs). Corollary \cite{2} implies that the leading temporal EoFs are theoretical Koopman eigenfunctions. Similarly, the data-adaptive harmonic decomposition (DAHD, \cite{21}) and Hankel alternative view of Koopman analysis (HAVOK, \cite{22}) are based on Hankel matrix and Gramian 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 Koopman eigenfunctions, and which features are not.

Remark 2. It is possible that some important systems are not ergodic. However, as long as

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} f(t) \tilde{f}(t+s)$$

exists for any $s \geq 0$, we can define an inner product on the complex valued linear space $\tilde{H}_f = \{ \sum_{i=1}^{n} c_i K^n f : c_i \in \mathbb{C}, n_i \geq 0, n \geq 1 \}$ by

$$\langle h, g \rangle = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} h(t) \tilde{g}(t),$$

where $h, g \in \tilde{H}_f$. Then we can define $\tilde{H}^{(2)}_f$ to be the completion of $\tilde{H}_f$. Thus we still can get a Hilbert space $\tilde{H}_f$ on which $K$ is an isometry. Then theorem 1 and theorem 2 still can be applied to $\tilde{H}^{(2)}_f$, Eq.\cite{8} still holds, and the identities Eq.\cite{12,14} are still valid. Note that the definition of $\tilde{H}^{(2)}_f$ only relies on the value of $f$ on the current trajectory. When the system is ergodic with a finite invariant measure, Birkhoff ergodic theorem implies that the inner product defined by the invariant measure coincides with the inner product for $\tilde{H}_f$. Hence $\tilde{H}_f$ and $\tilde{H}^{(2)}_f$ are the same. But there are non-ergodic systems that have finite invariant measure, while the limit in Eq.\cite{15} still exists. Hence the concept of $\tilde{H}^{(2)}_f$ is more general than that of $\tilde{H}_f$. In other words, as long as it is believed that the temporal mean of nice observables converges, identities Eq.\cite{12,14} shall hold.

III. PROOF OF THE MAIN THEOREM

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

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

$$\lim_{n \to \infty} \langle T^n f, g \rangle_H = 0$$

Proof of lemma \cite{2} For every $f \in \mathcal{H}$ the sequence $(\|T^n f\|)_{n \in \mathbb{N}}$ is decreasing thus convergent. For any $k \in \mathbb{N}$, we have

$$\|T^{k}T^n f - T^n f\|^2 = \|T^{k}T^n f\|^2 - 2 \text{Re} \langle T^{k}T^n f, T^n f \rangle + \|T^n f\|^2$$

$$= \|T^{k}T^n f\|^2 - 2\|T^{k}T^n f\|^2 + \|T^n f\|^2$$

$$\leq \|T^n f\|^2 - 2\|T^{k}T^n f\|^2 + \|T^n f\|^2$$

$$= \|T^n f\|^2 - 2\|T^{k}T^n f\|^2 \to 0 \text{ as } n \to \infty$$

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

$$\langle T^n f, g \rangle_H \to 0 \text{ for every } g \in \text{ran}(I - T^{k}T^n)$$

as $n \to \infty$.

The same argument for $T^*$ yields

$$\langle T^n f, g \rangle_H = \langle f, T^n g \rangle_H \to 0 \text{ for every } g \in \text{ran}(I - T^n T^{k*})$$

as $n \to \infty$.

We obtain that $\langle T^n f, g \rangle_H \to 0$ as $n \to \infty$ for every $f, g \in \text{ran}(I - T^{k}T^n) + \text{ran}(I - T^n T^{k*})$

$$= (\cap_{k \geq 1} \text{ran}(I - T^{k}T^n) + \cap_{k \geq 1} \text{ran}(I - T^n T^{k*}))$$

$$= (\cap_{k \geq 1} \text{fix}(T^{k}T^n) + \text{fix}(T^n T^{k*}))^\perp = \tilde{H}_{\mathcal{U}} = \mathcal{H}_{NU}$$

\square

Lemma 2. Let $\mathcal{H}$ be a Hilbert space over $\mathbb{C}$ of finite dimension or infinite dimension. $a_0, a_1, \ldots \in \mathcal{H}$ such that

$$\sum_i \|a_i\|^2 < \infty.$$ Then

$$\lim_{k \to \infty} \sum_i \|a_i, a_{i+k}\| = 0$$

Proof of Lemma \cite{3} 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 $j \geq M$, $\|a_j\| < \frac{\epsilon}{2}\|a_i\|$. Now for any $k > M$,

$$\sum_i \|a_i, a_{i+k}\| = \sum_{i=1}^{N} \|a_i, a_{i+k}\| + \sum_{i > N} \|a_i, a_{i+k}\|$$

$$\leq \sum_{i \leq N} \|a_i\| \|a_{i+k}\| + \frac{1}{2} \sum_{i > N} (\|a_i\|^2 + \|a_{i+k}\|^2)$$

$$\leq \frac{\epsilon}{2} \sum_{i \leq N} \|a_i\|^2 + \sum_{i > N} \|a_i\|^2$$

$$\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

\square
(c) holds for $i = 1$, we can then recursively deduce Theorem 5 for all $i$ by removing $a_{i}v_{i}$ from $f$ at each step. It is thus sufficient to prove that:

$$
\lim_{N \to \infty} \frac{\|A_{NM}(f_{NU})\|^2}{NM} = \lim_{N \to \infty} \frac{\|A_{NM}(f_{ac})\|^2}{NM} = 0,
$$

and that

$$
\lim_{N \to \infty} \frac{\|A_{NM}(f_{d})\|^2}{NM} = |a_{1}|^{2}.
$$

Now fix $N$, for any $g \in L^{2}(X, \mu)$, by Theorem 1

$$
\lim_{M \to \infty} \frac{\|A_{NM}(g)\|^2}{M} = \max_{\alpha} \frac{\sum_{i=0}^{M} |\alpha_{i}g(\bar{\alpha}_{i}+j)|^{2}}{M(|\alpha_{1}|^{2} + \ldots + |\alpha_{N}|^{2})}
$$

Hence Eq. (25) and (26) are equivalent to the following:

$$
\lim_{N \to \infty} \max_{\alpha} \frac{\sum_{i=0}^{N} |\alpha_{i}K_{i}g|^{2}}{N(|\alpha_{1}|^{2} + \ldots + |\alpha_{N}|^{2})} = \begin{cases} 
|a_{1}|^{2} & \text{if } g = f_{d} \\
0 & \text{if } g = f_{ac}\text{ or } f_{NU}. 
\end{cases}
$$

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

**Proposition 1** (The case for $f_{NU}$).

$$
\lim_{N \to \infty} \max_{\alpha} \frac{\sum_{i=0}^{N} |\alpha_{i}K_{i}f_{NU}|^{2}}{N(|\alpha_{1}|^{2} + \ldots + |\alpha_{N}|^{2})} = 0
$$

Proof of proposition 1. Without loss of generality, we may assume that $|\alpha_{1}| = 1$. Since $f_{NU} \in H_{NU} = \bigoplus_{s \geq 0} H_{s}$, we can write $f_{NU} = (a_{0}, a_{1}, \ldots)$, where $a_{i} \in H_{i}$. For $k > 0$, let $c_{k} = |(K^{i+k}f_{NU}, K^{i+k}f_{NU})| = |(K^{i+k}f_{NU}, K^{i+k}f_{NU})|$, which does not depend on $i$. Lemma 1 implies that $\lim_{|i-j| \to \infty} c_{i-j} = 0$. 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 $|\alpha_{1}|^{2} + \ldots + |\alpha_{N}|^{2} = 1$

$$
\sum_{i=0}^{N} |\alpha_{i}K_{i}f_{NU}|^{2} = \sum_{i,j} |\alpha_{i}\bar{\alpha}_{j}|(K_{i}^{j}f_{NU}, K_{i}^{j}f_{NU}) \\
\leq 2\sum_{k=0}^{N-k} |\alpha_{i}\bar{\alpha}_{i+k}|c_{k} \leq 2\sum_{k=0}^{N-k} (|\alpha_{i}|^{2} + |\alpha_{i+k}|^{2})c_{k} \\
\leq 2\sum_{k=0}^{N} c_{k} \leq 2\sum_{k=0}^{M_{\epsilon}} c_{k} + 2\sum_{k > M_{\epsilon}} c_{k} \\
\leq M_{\epsilon}/|f_{NU}|^{2} + (N - M_{\epsilon})\epsilon/2 \leq N\epsilon/2 + N\epsilon/2 = N\epsilon
$$

\[\blacksquare\]
Recall the notations in Theorem 4 for any $g \in \mathcal{H}_{f,U}$,
\[
\|\sum_{i=0}^{N} \alpha_i \mathcal{K}^i g\|_{L^2(X,dv)}^2 = \|\sum_{i=0}^{N} \alpha_i z^i \phi(g)\|_{L^2(S^1, dv_f)}^2 \\
= \int_{S^1} \left| \sum_{i=0}^{N} \alpha_i z^i \right|^2 |\phi(g)(z)|^2 dv_f(z)
\]
(32)
This proves that

**Proposition 2.** For any $g \in \mathcal{H}_{f,U}$,
\[
\lim_{N \to \infty} \lim_{M \to \infty} \frac{\|A_{NM}(g)\|^2}{NM} = \lim_{N \to \infty} \max_{\|\alpha\|=1} \frac{1}{N} \int_{S^1} \left| \sum_{i=0}^{N} \alpha_i z^i \right|^2 |\phi(g)(z)|^2 dv_f(z)
\]
(33)
To prove Eq. (29) for $g = f_d, f_{sc}$ and $f_{ac}$, we start with the following lemma.

**Lemma 5.** Let $f,h \in \mathcal{H}_{f,U}, \nu_f, \phi$ be the same as in Theorem 2. For simplicity, denote $\nu_f$ by $\nu$, $\nu = \nu_d + \nu_{ac} + \nu_{sc}$. Let $\nu_{d,1}$ be a purely discrete finite measure on $S^1$, such that $\{\xi_1, ..., \xi_L\} = \text{Supp}((\nu_{d,1})) \subset \text{Supp}(\nu_d)$ 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 $f_k = \phi^{-1}(1_{\{\xi_k\}})$, and set $h = \sum_{k=1}^{L} f_k$.

Let $d_{NM,1}(h)$ be the leading singular value of $A_{NM}(h)$. Then
\[
\lim_{N \to \infty} \lim_{M \to \infty} \frac{d_{NM,1}^2(h)}{NM} = \max_{k} c_k^2.
\]
(34)
**Proof of proposition 3** According to proposition 2
\[
\lim_{N \to \infty} \lim_{M \to \infty} \frac{d_{NM,1}^2(h)}{NM} = \lim_{N \to \infty} \max_{\|\alpha\|=1} \frac{1}{N} \int_{S^1} \left| \sum_{i=0}^{N} \alpha_i z^i \right|^2 |\phi(h)(z)|^2 dv_f(z)
\]
\[
= \lim_{N \to \infty} \max_{\|\alpha\|=1} \frac{1}{N} \int_{S^1} \sum_{i=0}^{N} \alpha_i z^i \phi(h)(z)^2 dv_f(z)
\]
(35)
Then lemma 4 implies what we want to prove.

**Proposition 3** (The case for $f_d$). Eq. (26) holds.
**Proof of proposition 3** For any $\epsilon > 0$, we choose a truncation $\nu_d = \nu_{d,1} + \nu_d(\epsilon)$, so that $\nu_d(\epsilon) < \epsilon$, $|\text{Supp}(\nu_{d,1})| < \infty$, and that $\nu_{d,1}(\{\xi_k\}) = \nu_d(\{\xi_k\})$ whenever $\xi_k \in \text{Supp}(\nu_{d,1})$. Then $\epsilon$ is small enough, $|\alpha_i|^2 = \max_{k} \nu_d(\{\xi_k\})$.

Apply lemma 5 to $\nu_{d,1}$, and let $h$ be defined as in lemma 5. Then $f = h + f_d$ and
\[
A_{NM}(f) = A_{NM}(h) + A_{NM}(f_d),
\]
(36)
and
\[
\lim_{N \to \infty} \lim_{M \to \infty} \frac{\|A_{NM}(h)\|^2}{NM} = c_1^2 = \max_{\xi} \nu_d(\{\xi\}) = |\alpha_1|^2
\]
(37)
And note that, applying Cauchy-Schwartz inequality,
\[
\lim_{N \to \infty} \lim_{M \to \infty} \frac{\|A_{NM}(f_d)\|^2}{NM} = \lim_{N \to \infty} \frac{\max_{\|\alpha\|=1} \frac{1}{N} \int_{S^1} \sum_{i=0}^{N} |\alpha_i z^i|^2 dv_{d,e}}{NM} \leq \nu_{d,e}(\{S^1\}) < \epsilon
\]
(38)
Eq. (37) (38) implies Eq. (26) by letting $\epsilon \to 0$.

**Proposition 4.** Let $\nu_c$ be a continuous finite measure on $S^1$. Then
\[
\lim_{N \to \infty} \max_{\|\alpha\|=1} \frac{1}{N} \int_{S^1} \sum_{i=0}^{N} |\alpha_i z^i|^2 dv_c = 0.
\]
(39)
**Proof of proposition 4** Let $c_k = |\nu_c(z^k)|$. In lemma 3 let $f = g = 1$, it implies that
\[
\lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N} c_k^2 = 0.
\]
(40)
Therefore for $\|\alpha\| = 1,$
\[
\int_{S^1} \frac{1}{N} \sum_{i=0}^{N} |\alpha_i z^i|^2 dv_c \leq \frac{2}{N} \sum_{k=0}^{N} \sum_{i=0}^{N-k} |\alpha_{i+k} \alpha_i| c_k
\]
\[
\leq \frac{1}{N} \sum_{k=0}^{N-k} (|\alpha_{i+k}|^2 + |\alpha_i|^2) c_k \leq \frac{1}{N} \sum_{k=0}^{N} 2c_k
\]
\[
\leq \frac{2}{N} \sqrt{(N+1) \sum_{k=0}^{N} c_k^2} = 2 \sqrt{\frac{N+1}{N} \sum_{k=0}^{N} c_k^2} \to 0,
\]
(41)
as $N \to \infty$.

**Corollary 2** (The case for $f_{sc}$ and $f_{ac}$). Eq. (26) holds for $f_{sc}$ and $f_{ac}$.
**Proof.** This is the direct consequence of proposition 2 and 4.

**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 constructing the $i$-th Koopman eigenfunction using Hankel matrix. For this purpose, we provide the following algorithm.

- Given a time series data \{f(t)\}_{0 \leq t \leq T}, choose $N_k, M_{k,j}$ where $1 \leq j \leq L_k$, such that $N_k + M_{k,j} \leq T$, $M_{k,1} < M_{k,2} < ... < M_{k,L_k} \gg N_k$. 

• For each $N_k, M_{k,j}$, compute the renormalized singular values of $A_{N_k M_{k,j}}$, denoted by $\sigma_{k,j,i} = \frac{d_{k,j,i}^2}{N_k M_{k,j}}$.

• 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 Koopman eigenfunctions is not well represented by this dataset.

• Given $i$, for all $k$, $\sigma_{k,j,i}$ shows good convergence, then check if $\sigma_{k,j,i}$ converges as $k$ increases. If $\sigma_{k,j,i}$ converges to some nonzero number, then the $i$–th Koopman eigenfunction is well represented by this data set. Otherwise, the $i$–th Koopman eigenfunction is not well-represented by this data set.

• Given $i$, if $\sigma_{k,j,i}$ passes the convergence tests, then it shows that $i$–th left singular vector $v_i$ of $A_{NM}$ is the realization of some Koopman eigenfunction at time $t = 0, 1, 2, \ldots, N$. The absolute value of the frequency of this Koopman eigenfunction can be approximated by $n_p/N$, where $n_p$ is the number of local maximums of the real part of $v_i$. The sign of the frequency of this Koopman eigenfunction is positive (negative) if the complex singular vector $v_i$ rotates counter-clockwise (clockwise, respectively).

### A. Implication to Hankel DMD

In [7] a Hankel DMD algorithm has been proposed and the authors showed that $\lim_{M \to \infty} d_{N M,i}^2/M$ can be used to identify Koopman and non-Koopman eigenfunctions for fixed $N$ under the conditions that 1), the Hilbert space $H_f$ is finite dimensional and 2), $N$ is larger than the dimension of $H_f$. More precisely, they showed that $\lim_{M \to \infty} d_{N M,i}^2/M > 0$ if and only if $d_{N M,i}$ corresponds to a Koopman eigenfunctions. 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 $H_f$ is infinite, their method unfortunately fails. Because $\lim_{M \to \infty} d_{N M,i}^2/M = \lim_{M \to \infty} c_{NM,i} = c_{N \infty,i}$. And $c_{N \infty,i}$ can be positive even if there is no Koopman eigenfunctions. Therefore Theorem [6] can be thought of as a completion of the method posed in [7] under a much weaker assumption, by letting $N \to \infty$.

### B. Comparison with Yosida’s formula

Yosida mean ergodic theorem [23] provides a formula to calculate $a_\omega$, the coefficient of the Koopman eigenfunction of frequency $\omega$ in Eq. [10]:

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

$a_\omega = 0$ if $\omega$ is not a Koopman eigenfrequency. Under the assumption of ergodicity and finite invariant measure, this formula can be proved by combining Theorem [2] lemma [2] and Von-Neumann ergodic theorem. This formula was first introduced to the fluid dynamics’ community by [26, 27]. Eq. [42] is easy to compute for 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. [42] for all $\omega \in \{k \Delta \omega : k = 1, \ldots, n\}$ and then finding the peak value.

On the other hand, from the theoretical point of view, 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 [16], 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 $\frac{d_{N M,1}^2}{N M}$.

Then the decomposition $f = \sum_i a_i f_i + f_{NU} + f_{sc} + f_{ac}$ can be reduced to $f = f_{NU} + f_{sc} + f_{ac}$. As expected, Fig. [1] does not display the tendency that $E_{H}(N, \infty)$ converge to some nonzero value as $N \to \infty$.

#### B. A simple 4-dimensional system

Following one of the numerical examples in [10], 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 [10] that $\mu$ is an invariant measure. Still, the Lorenz system does not have non-trivial Koopman eigenfunctions and $(X, T, \mu)$ is ergodic.

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

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

(43)
Hence it is worth to make a numerical comparison about the same quantity as the one calculated by Eq.(42). The singular value of the Hankel matrix should then converge as in Eq.(10).

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.\( \text{(10)} \).

Anticipated by our main theorem, the renormalized singular value of the Koopman eigenfunction is exactly the inverse of the period of \( (\omega_S, \omega_S) \). The integration time step for Lorenz system is \( \Delta t = 0.01 \). The Runge-Kutta 4th order scheme is applied for the numerical integration. We also computed \( \| \sum_i a_i v_i + f_{sc} + f_{ac} + f_{NU} \| \) obtained from Yosida’s formula and that from the numerical results of \( (N, M) \) for \( (X_{63} \times S^1, T_{63} \times T_1, \mu_{63} \times \mu_1) \).

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 eigenfunction \( v_\omega \) 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 Hankel matrix indeed corresponds to the same eigenfrequency \( \omega \). \( E_{H,1}(10^3, M) \) and \( E_{H,2}(10^3, M) \) seem to converge to the same value. This is because the Koopman eigenfunctions always exist in pair, i.e. \( \exp(2\pi i \omega) \in \text{Supp}(\nu_i) \iff \exp(-2\pi i \omega) \in \text{Supp}(\nu_{\bar{i}}) \). Since the observable \( f \) is real, the coefficient \( a_\omega = a_{\bar{\omega}} \). 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 through Hankel DMD.

A main assumption is the ergodicity of the system, 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. Another assumption is that the system should be measure-preserving. It certainly cannot be assert if there exists an invariant measure for the whole Earth system, which includes the Earth, the atmosphere, ocean, all celestial bodies, but also the biology and living animals, etc. However, remark \( 2 \) applies to this situation since the overall constant rising tendency has already been removed. We simply renormalize 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.\( \text{(42)} \)) 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 system is ergodic and measure preserving 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 Koopman eigenfunction of frequency $\omega$. Similarly, since the SSH are real numbers, $|a_\omega| = |a_{\bar{\omega}}|$ and $|a_\omega|^2 + |a_{\bar{\omega}}|^2 = 2|a_\omega|^2$ 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 $(114.875^\circ W, 6.125^\circ N)$) is represented by the yearly signal. Constructing the Hankel matrix for SSH at $(114.875^\circ W, 6.125^\circ N)$, i.e. we choose $f = \text{SSH}(114.875^\circ W, 6.125^\circ N)$, 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 Koopman eigenfunctions for $\omega$ and $\bar{\omega}$. However, the third and fourth renormalized singular values do not 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 size of the dynamical system.

VI. CONCLUSION

The main objective of this study is to provide a rigorous and practical method to identify Koopman eigenfunctions for discrete-time ergodic and (finite) measure preserving dynamical systems. 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 eigenfunctions. As a direct corollary of the main theorem, the leading temporal EOFs, which are calculated from the eigen decomposition of the Gramian matrix, are showed to be approximations of theoretical Koopman eigenfunctions. The main theorem provides partial theoretical foundation to several existing empirical methods including SSA, DAHD, and HAVOK. For real application purposes, the dataset of which is on a single trajectory and the system of which is infinite dimensional, it remains to fully assess whether the alternative assumption in Remark 2 is satisfied. The main result shows that the discrete spectrum can be characterized by the singular values (eigenvalues) of Hankel (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 system.
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 [2] implies that the weakly mixing property is equivalent to

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N-1} | \int_{S^1} z^i d\mu(z) |^2 = 0,$$

(A1)

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

$$\lim_{n \to \infty} \frac{\# \{0 < k < n : | \int_{S^1} z^k d\mu(z) | > \epsilon \}}{n} = 0,$$

(A2)

for any $\epsilon > 0$ and continuous measure $\mu$. We shall provide an alternative proof for proposition [4] and then derive Eq. (A2) from proposition [4]. 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 $\xi_1, \ldots, \xi_L \in S^1$, 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 $A_1 \subset \mathbb{N}$ such that $|A_1| = \infty$ and $\xi_1 \in A(i)$ for any $i \in A_1$. 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 $A(i_1, 1) = A(i) - I_1$ for $i \in A_1$, we have $\nu_c(A(i_1, 1)) > \epsilon/2$. Then we apply the same analysis to $\{A(i_1, 1)\}_{i \in A_1}$, to find $\xi_2$, etc. After doing the same analysis for $L$ times, we get $\xi_1, \ldots, \xi_L$ and $A_L$, such that $|A_L| = \infty$ and $\xi_j \in A(i)$ for any $j$ and $i \in A_L$.

To prove that there exists a point $\xi_1 \in S^1$ and $A_1 \subset \mathbb{N}$ such that $|A_1| = \infty$ and $\xi_1 \in A(i)$ for any $i \in A_1$. 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 $B_1 \supset B_2 \supset \ldots$ and $\cap_{N \geq 1} B_N = \emptyset$. It means that $\nu_c(B_N) \to 0$ as $N \to \infty$. This is apparently not true because $\nu_c(B_N) \geq \nu_c(A(N+1)) > \epsilon$.

Now we give another proof of proposition 4 based on lemma 6.

An alternative proof of proposition 4. We prove by contradiction. We assume that proposition 4 does not hold. Then there exists $\epsilon > 0$ and a sequence of $\alpha(i) = \{\alpha_{i,1}, \ldots, \alpha_{in}\}$, such that $\|\alpha(i)\|^2 = 1$ and

$$\nu_c(|P(i)|^2/n_i) > \epsilon,$$

(A3)

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

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

(A4)

Then $\nu_c(A(i)) \geq \frac{\delta}{2}$. Now we pick any $L > 4/\epsilon$. Lemma 6 implies that we can find $\xi_1, \ldots, \xi_L \in S^1$, $\xi_i \neq \xi_j$ for $i \neq j$, and a subsequence $A(i_{j_k})$ of $A(i)$, such that $\xi_k \in A(i_{j_k})$ for any $k, j$. Then for any $i, j$, set $N = n_{i,j}$ and $\alpha = \alpha(i_{j_k})$,

$$\sum_{k=1}^L \frac{1}{N} \sum_{i=0}^N |\alpha_i \xi_k|^2 \geq \sum_{k=1}^L \frac{1}{N} \frac{\epsilon}{2} N \geq 2$$

(A5)

This contradicts with lemma 6 by setting $c_1 = c_2 = \ldots = c_L = 1$.

Now we derive Eq. (A2) from proposition 4.

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

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

(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

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

Then proposition 4 implies that

$$\lim_{n \to \infty} \frac{M_n \epsilon^2}{n^2} = 0.$$  

(A7)

Hence

$$\lim_{n \to \infty} \frac{M_n}{n} = 0.$$  

(A8)
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, arXiv: Dynamical Systems (2017).
16. S. Das and D. Giannakis, Journal of Statistical Physics 175, 1107 (2017).
17. S. Das and D. Giannakis, arXiv: Dynamical Systems (2018).
18. D. Giannakis, S. Das, and J. Slawinska, arXiv: Dynamical Systems (2018).
19. D. Giannakis, Research in the Mathematical Sciences 8, 1 (2020).
20. 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).
21. D. Kondrashov, E. Ryzhov, and P. Berloff, Chaos 30 6, 061105 (2020).
22. S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, and J. N. Kutz, Nature Communications 8 (2017).
23. K. Yosida, Functional Analysis (Springer, Berlin, Heidelberg, 1995).
24. D. Borthwick, Spectral Theory: Basic Concepts and Applications, 1st ed., Graduate Texts in Mathematics №284 (Springer International Publishing;Springer, 2020).
25. P. R. Halmos, Lectures on Ergodic Theory (Chelsea Publishing Company, New York, N.Y, 1956).
26. I. Mezić and A. Banaszuk, Physica D: Nonlinear Phenomena 197, 101 (2004).
27. I. Mezić, Nonlinear Dynamics 41, 309 (2005).
28. A. Cazenave and W. Llovel, Annual review of marine science 2, 145 (2010).