Abstract

Every invertible, measure-preserving dynamical system induces a Koopman operator, which is a linear, unitary evolution operator acting on the $L^2$ space of observables associated with the invariant measure. Koopman eigenfunctions represent the quasiperiodic, or non-mixing, component of the dynamics. The extraction of these eigenfunctions and their associated eigenfrequencies from a given time series is a non-trivial problem when the underlying system has a continuous spectrum or a dense point spectrum, behaving similarly to a noisy component of the signal. This paper describes methods for identifying Koopman eigenfrequencies and eigenfunctions from a discretely sampled time series generated by such a system with unknown dynamics. Our main result gives necessary and sufficient conditions for a Fourier function, defined on $\mathbb{N}$ states sampled along an orbit of the dynamics, to be extensible to a Koopman eigenfunction on the whole state space, lying in a reproducing kernel Hilbert space (RKHS). In particular, we show that such an extension exists if and only if the RKHS norm of the Fourier function does not diverge as $N \to \infty$, in which case the corresponding Fourier frequency is also a Koopman eigenfrequency. Using these results, we establish data-driven criteria for identifying Koopman eigenfrequencies from a set of candidate frequencies. Numerical experiments on mixed-spectrum systems with weak periodic components demonstrate that this approach has significantly higher skill in identifying Koopman eigenfrequencies compared to conventional spectral estimation techniques based on the discrete Fourier transform.

Keywords: Koopman operator, spectral theory, reproducing kernel Hilbert space, ergodic dynamical systems

2010 MSC: 37A05, 37A10, 37A30, 37A45

1. Introduction

A common scenario in the analysis of data generated by dynamical systems is that the underlying discrete or continuous time flow is unknown, and the system is observed through some observation map $F$ taking values in a vector space (the data space). The challenge then is to infer various properties of the system from the time series $\{F(x_0), F(x_1), \ldots\}$, where $\{x_0, x_1, \ldots\}$ is an orbit in the system’s state space with a fixed sampling interval $\Delta t$. The focus of this paper is on the identification of eigenfunctions of an operator called the Koopman operator [1], which governs the evolution of observables under the dynamics.

Koopman eigenfunctions form a distinguished class of observables that evolve by multiplication by a time-periodic factor $e^{i\omega t}$, even if the dynamics is aperiodic. They extract temporally coherent, and thus highly predictable, temporal patterns from complex dynamics, and these patterns have high physical interpretability by virtue of being associated with an operator intrinsic to the dynamical system generating the data. In addition, observables lying in the span of Koopman eigenfunctions have an integrable time evolution, and are useful for reduced order modeling and forecasting. Due to these properties, Koopman eigenfunctions warrant identification from data.

There have been many approaches to the identification of Koopman spectra (and the spectra of the related Peron-Frobenius operators, which are duals to Koopman operators), such as methods based on state space partitions [2], harmonic averaging [3, 4], Krylov subspace iteration [5, 6], dictionary-based approximation [7, 8, 9], Galerkin approximation [10, 11], delay-coordinate embeddings [10, 11, 12, 13], and spectral moment estimation methods [14]. Among these, the methods based on harmonic averaging are...
Figure 1: A mixed-spectrum signal (top row), and the results of spectral analysis via the DFT (bottom left) and the RKHS-based approach using a Gaussian kernel (e). The input signal $F(t) = (F_1(t), F_2(t), F_3(t)) \in \mathbb{R}^3$ is a superposition of a periodic signal and a chaotic signal, generated by the product dynamical system in (35) exhibiting both discrete and continuous spectral components. The Koopman eigenfrequencies present in $F(t)$ are 1 and 2. The DFT-based spectral analysis was performed on datasets consisting of $N = 50,000$ and $70,000$ samples. In either case, it fails to provide a clear demarcation of the true frequencies in the point spectrum of the system. In particular, note that the $\omega = 1$ peak in the DFT spectrum for $N = 50,000$ disappears at $N = 70,000$. On the other hand, the RKHS-based approach described in this paper correctly identifies the $\omega = 1$ and 2 eigenfrequencies of the Koopman operator. The method constructs a data-driven eigenbasis of an RKHS $H$ of functions on state space, whose elements can be nonlinear functions of the input data (depending on the kernel employed). Then, for each Fourier function $F_\omega$, sampled at $N$ time-ordered points along a dynamical orbit, it computes the squared RKHS norm $w_{N,l}(F_\omega)$ of its projection onto the RKHS subspace spanned by the leading $l$ basis functions (here, $l = 1240$). Theorem 1 shows that as $N \to \infty$, and for sufficiently large fixed $l$, this quantity converges to a nonzero finite number if and only if there exists an RKHS extension of $F_\omega$ to a Koopman eigenfunction at eigenfrequency $\omega$, and otherwise converges to 0.

closely related to spectral estimation techniques via the Fourier transform (DFT). Given the time series $\{F(x_0), F(x_1), \ldots\}$, harmonic averaging approaches compute the quantities

$$\mathcal{F}_{\omega,N}F := \frac{1}{N} \sum_{n=0}^{N-1} e^{-i\omega n \Delta t} F(x_n)$$

for a set of candidate frequencies $\omega \in \mathbb{R}$. If a frequency $\omega$ in the candidate set lies in the point spectrum of the dynamical system, then $f$ acquires a discrete spectral component $a_\omega$ at that frequency, and as the sample size $N$ increases, $\mathcal{F}_{\omega,N}f$ converges to $a_\omega$. Thus, the quantities in (1) can in principle reveal the point spectrum of the dynamics based on empirical time-ordered measurements of observables.

Despite this useful property, a direct application of harmonic averaging has a number of limitations. For example, signals are often noisy, and even in a deterministic system without noise, if there is a non-empty continuous spectrum [see the definitions preceding (7)], then the signal will be spectrally similar to one generated by a noisy source. In such cases, it would be difficult to distinguish the true discrete spectral components from those due to noise and/or the continuous spectrum. Moreover, the magnitude of the discrete spectral components carried by the signal may rapidly decay with increasing frequency, making the task more difficult. Figure 1 illustrates some of the shortcomings of spectral estimation via (1) with an application to a chaotic signal.

In this work, we approach the problem of estimating the point spectra of Koopman operators as an extrapolation problem, in which we seek to extend a candidate eigenfunction from its values on the sample
trajectory to the entire space, in a reproducing kernel Hilbert space (RKHS) of functions. Our approach is based on the observation that along an orbit of the dynamics a continuous Koopman eigenfunctions behaves like a Fourier function, evolving as $e^{i\omega t}$ for a real frequency $\omega$. On the other hand, it is not the case that every Fourier function on an orbit extends to a Koopman eigenfunction lying in a space of observables of sufficient regularity. Choosing the class of RKHSs as Hilbert spaces naturally encapsulating a notion of regularity of observables, we will show that the Fourier functions on orbits admitting RKHS extensions to the entire state space are precisely those having RKHS representatives constructed from $N$ samples on the orbit with convergent squared RKHS norm $w_N$ as $N \to \infty$. Moreover these extensions will be shown to be Koopman eigenfunctions with eigenfrequencies equal to the corresponding Fourier frequencies. The RKHS framework also allows for stable evaluation of the approximate Koopman eigenfunctions determined from $N$ samples at arbitrary points on state space, in contrast to approximation in $L^2$ spaces which only yields estimates for the eigenfunction values at the sampled states.

In Section 9, we will revisit the example shown in Figure 1, and discuss how the application of our methods using RKHSs associated with covariance kernels makes the analysis related to harmonic averaging. As shown in Figure 1, the RKHS-based analysis utilizing a Gaussian kernel, which results in an infinite-dimensional RKHS despite the fact that the data measurement function takes values in a finite-dimensional space, $\mathbb{R}^3$, correctly identifies two Koopman eigenfrequencies, which in this case is enough to recover all eigenfrequencies.

### 2. Assumptions and statement of the main results

In this work, we are interested in the spectral analysis of continuous-time flows, and the basic assumptions on the system are stated below.

**Assumption 1.** $\Phi^t : X \to X, \ t \in \mathbb{R}$, is a continuous flow on a compact metric space $X$, possessing an ergodic Borel probability measure $\mu$. The system is sampled at a fixed interval $\Delta t > 0$, such that the discrete-time evolution map $\Phi^{n\Delta t} : X \mapsto X, \ n \in \mathbb{Z}$, is also ergodic for the probability measure $\mu$.

Given an initial point $x_0 \in X$, we let $O$ denote the orbit $\{x_n = \Phi^{n\Delta t}(x_0) : n \in \mathbb{Z}\}$, and $X_N$ the finite trajectory $\{x_0, \ldots, x_{N-1}\} \subset O$. In addition, we shall be concerned with data-driven approximation of spectral quantities based on measurements of the system in a data space $Y$. The measurement function, $F : X \mapsto Y$, will be assumed to have the following properties.

**Assumption 2.** $F : X \mapsto Y$ is a continuous, injective map, taking values in a metric space $Y$.

Note that the injectivity requirement on $F$ can be generically relaxed through the use of delay-coordinate maps [15]; we will discuss this point further in Section 9.

*The Koopman operator.* Under Assumption 1, a natural space of observables (functions of the state) of the dynamical system is the Hilbert space $L^2(X, \mu)$ of square-integrable, complex-valued functions on $X$, henceforth abbreviated $L^2(\mu)$. This space is a Hilbert space equipped with the inner product $(f, g)_\mu = \int_X f \ast g \, d\mu$. The ergodic dynamical flow $\Phi^t$ induces a flow on $L^2(\mu)$ through a strongly continuous, 1-parameter group of unitary operators, $U^t : L^2(\mu) \mapsto L^2(\mu)$, $t \in \mathbb{R}$, called Koopman operators, which act on observables by composition with the flow map, $U^t f = f \circ \Phi^t$. The unitarity of $U^t$ stems from the fact that $\Phi^t$ is an invertible, $\mu$-preserving map, and it implies that all of its eigenvalues lie on the unit circle of the complex plane. In addition, it follows from the continuity of the map $t \mapsto \Phi^t$ that every eigenvalue of $U^t$ has the form $e^{i\omega t}$, where $\omega$ is a real eigenfrequency. As a result, an eigenfunction $z$ corresponding to that eigenvalue satisfies the eigenvalue equation

$$U^t z = e^{i\omega t} z.$$  \hspace{1cm} (2)

In fact, it follows from Stone’s theorem for strongly continuous unitary groups that $i\omega$ is an eigenvalue of the generator $V : D(V) \mapsto L^2(\mu)$ of the Koopman group; a skew-adjoint unbounded operator with a
dense domain $D(V) \subseteq L^2(\mu)$, acting on observables as a “time derivative”, viz. $Vf = \lim_{t \to 0} (U^t f - f)/t$, $f \in D(V)$. By ergodicity of the flow, all eigenvalues of $V$ are simple.

As stated above, our approach for identifying Koopman eigenfrequencies and their corresponding eigenfunctions involves testing for extensibility of Fourier functions $f_\omega$ defined on orbits of the dynamics. Specifically, for each frequency $\omega \in \mathbb{R}$ we define

$$f_\omega : \mathcal{O} \to \mathbb{C}, \quad f_\omega(x_n) = e^{in\omega \Delta t}, \quad (3)$$

and seek to determine whether $f_\omega$ extends to a function $\tilde{f}_\omega : X \to \mathbb{C}$ of appropriate regularity. Observe, in particular, that if $\omega$ were a Koopman eigenvalue corresponding to a continuous eigenfunction $z_\omega$, then by (2), the restriction $\tilde{z}_\omega|\mathcal{O}$ to the orbit coincides with $f_\omega$ in (3), and thus with $f_\omega|\mathcal{X}_N$, the restriction of $f_\omega$ on the finite trajectory $\mathcal{X}_N$. Here, we are interested in this question in the reverse direction, i.e., our objective is to identify for which $\omega \in \mathbb{R}$ the Fourier function $f_\omega$ can be extended to a Koopman eigenfunction $\tilde{f}_\omega$, satisfying (2) at eigenfrequency $\omega$. This is essentially an extrapolation problem of a candidate function from a countable set $\mathcal{O}$ to the entire space $X$. In addition, we are interested in performing this extrapolation in a data-driven manner; that is, from measurements of the system taken on the finite trajectory $\mathcal{X}_N$ without prior knowledge of the dynamics or the structure of state space.

While natural from a theoretical point of view, the space of continuous functions on $X$ is arguably not well suited for meeting our objectives, for it lacks the inner product structure that significantly facilitates the implementation of data-driven techniques. Instead, following the widely adopted paradigm in statistical learning theory, we will focus on identification of Koopman eigenfunctions lying in an RKHS, which, in addition to the pointwise approximation guarantees provided by the $C^0$ (uniform) norm on the space of continuous functions, its Hilbert space structure allows the construction of data-driven algorithms based on standard linear algebra tools. Throughout this work, we will restrict attention to RKHSs with continuous reproducing kernels, so that convergence in RKHS norm implies convergence in $C^0$ norm. Before proceeding, we briefly review some of the main properties of RKHSs.

Reproducing kernel Hilbert spaces. An RKHS $\mathcal{H}$ on $X$ is a subspace of the linear space of complex-valued functions $f : X \to \mathbb{C}$, equipped with an inner product $\langle \cdot , \cdot \rangle_\mathcal{H}$ making it a Hilbert space. This Hilbert space structure allows operations such as orthogonal projections, needed for most numerical procedures. A distinguishing feature of $\mathcal{H}$ is that for every $x \in X$, the point-evaluation map $\delta_x : \mathcal{H} \to \mathbb{C}$, $\delta_x f = f(x)$, is a bounded, and thus continuous, linear functional. By the Moore-Aronsajn theorem [16], every RKHS is uniquely determined through its reproducing kernel; a bivariate function $k : X \times X \to \mathbb{C}$ with the following properties:

(i) $k$ is conjugate symmetric, i.e., $k(x, y) = k(y, x)^*$ for all $x, y \in X$.
(ii) $k$ is positive-semidefinite, i.e., for every $x_1, \ldots, x_n \in X$ and $a_1, \ldots, a_n \in \mathbb{C}$, $\sum_{i,j=1}^n a_i^* a_j k(x_i, x_j) \geq 0$, with equality occurring iff each of the $a_1, \ldots, a_n$ equals zero.
(iii) For every $x \in X$, the kernel sections $k(x, \cdot)$ lie in $\mathcal{H}$.
(iv) The reproducing property $f(x) = \delta_x f = \langle k(x, \cdot), f \rangle_\mathcal{H}$ holds for every $f \in \mathcal{H}$ and $x \in X$.

These properties imply that $\langle k(x, \cdot), k(y, \cdot) \rangle_\mathcal{H} = k(x, y)$ for all $x, y \in X$, and $\mathcal{H}$ is the closure (in the $\mathcal{H}$ norm) of finite sums $f = \sum_{i=1}^n a_i k(x_i, \cdot)$ with $a_1, \ldots, a_n \in \mathbb{C}$ and $x_1, \ldots, x_n \in X$. If $k$ is continuous and strictly positive definite, $\mathcal{H}$ is a dense subspace of the space of continuous functions on $X$; a property oftentimes referred to as universality [17]. In fact, if $k$ is a smooth kernel on a compact manifold, $\mathcal{H}$ is a subset of the $C^r$ spaces on the manifold for all $r \geq 0$ [18]. For our purposes, RKHSs have the key property that they allow out of sample extensions of $L^2$ equivalence classes of functions with respect to any finite measure $\nu$ on $X$ to everywhere-defined functions in $\mathcal{H}$. In light of the above, we will require that the following conditions on kernels be satisfied:

Assumption 3. $k : X \times X \to \mathbb{R}$ is a reproducing kernel for an RKHS $\mathcal{H}$. Moreover,

(i) $k$ is continuous.
(ii) $k$ is strictly positive-definite.
Consider now the finite trajectory $X_N$ and the probability measure $\mu_N = N^{-1} \sum_{n=0}^{N-1} \delta_{x_n}$, where $\delta_{x_n}$ is the Dirac delta-measure supported on the point $x_n$. $\mu_N$ is called the sampling measure supported on the finite trajectory $X_N$. By ergodicity, for $\mu$-a.e. $x \in X$, as $N \to \infty$, $\mu_N$ converges strongly to the invariant measure $\mu$, i.e., for every Borel-measurable set $A$, $\lim_{N \to \infty} \mu_N(A) = \mu(A)$. The data-driven analog of $L^2(\mu)$ is the Hilbert space $L^2(\mu_N) = L^2(X_N, \mu_N)$, consisting of equivalence classes of functions on $X$ taking the same values on the finite set $X_N$, while taking arbitrary values on $X \setminus X_N$. This space is an $N$-dimensional Hilbert space isomorphic to $\mathbb{C}^N$ equipped with a normalized Euclidean inner product. It is also isomorphic to the Hilbert space of complex-valued functions on $X_N$, equipped again with a normalized Euclidean product. In what follows, we will identify elements of the latter space with elements of $L^2(\mu_N)$ without the use of additional notation.

Next, as a data-driven analog of $\mathcal{H}$, we consider the $N$-dimensional subspace $\mathcal{H}_N \subset \mathcal{H}$, defined as the linear span of the set of kernel sections $\{k(x_0, \cdot), \ldots, k(x_{N-1}, \cdot)\}$ on $X_N$. In Section 4, the fact that $\bigcup_{N \in \mathbb{N}} X_N = \mathcal{O}$ and Assumption 3(ii) will be used to show that the union of the spaces $\mathcal{H}_N$ is dense in $\mathcal{H}$.

Fitting RKHS functions. As will be discussed in more detail in Section 4 under Assumption 3(ii), there exists an extension operator $T_N : L^2(\mu_N) \mapsto \mathcal{H}$, mapping the equivalence class $f \in L^2(\mu_N)$ to a pointwise defined function $h = T_N f \in \mathcal{H}$, such that $f(x_n) = h(x_n)$ for all $x_n \in X_N$. Moreover, $h$ is the unique element of $\mathcal{H}$ with this property, and lies in fact in $\mathcal{H}_N$ (i.e., $h$ is equal to a linear combination of $k(x_0, \cdot), \ldots, k(x_{N-1}, \cdot)$).

Using this operator, we define the non-negative functional

$$w_N : L^2(\mu_N) \mapsto \mathbb{R}, \quad w_N(f) = \|T_N f\|_{H_N}^2. \quad (4)$$

The functional $w_N$ induces a norm on equivalence classes of functions with respect to $\mu_N$, distinct from the $L^2(\mu_N)$ norm. Intuitively, the ratio

$$r_N(f) = w_N(f)/\|f\|_{\mu_N} \quad (5)$$

can be thought of as a measure of “roughness” of $f$; that is, the larger that quantity is, the stronger the degree of spatial variability of its RKHS extension $T_N f$ becomes. Our main result below establishes a necessary and sufficient condition in terms of $w_N(f_\omega|X_N)$ for $\omega$ to be a Koopman eigenvalue corresponding to a Koopman eigenfunction in $\mathcal{H}$.

**Theorem 1.** Let $f_\omega|X_N, \omega \in \mathbb{R}$, be a Fourier function on the finite trajectory $X_N$ as in (3). Then, under Assumptions 4 and 5,

(i) $\lim_{N \to \infty} w_N(f_\omega|X_N) = \infty \iff f_\omega$ does not have an extension $\tilde{f}_\omega \in \mathcal{H}.$

Moreover, if $f_\omega$ exists,

(ii) $\lim_{N \to \infty} w_N(f_\omega|X_N) = \|\tilde{f}_\omega\|_H^2$,

(iii) $f_\omega$ is a Koopman eigenfunction at eigenfrequency $\omega$.

Remark. The elements of the sequence $w_N(f_\omega|X_N)$ in Theorem 1 are the squared norms of the vectors $T_N(f_\omega|X_N)$ lying in $\mathcal{H}$. For a general sequence of vectors in a Hilbert space with bounded norm, the norms may not converge, and even if they did, the vectors themselves may not be convergent. For the vectors $f_\omega|X_N$, however, the boundedness of their norm is in fact equivalent to them forming a convergent sequence.

Theorem 1 gives necessary and sufficient conditions under which a Fourier function sampled on a countable set can be extended to a Koopman eigenfunction in an RKHS. In other words, it establishes necessary and sufficient conditions under which the following set of special frequencies is non-empty:

$$\Omega := \{\omega \in \mathbb{R} : \omega \text{ is a Koopman eigenfrequency corresponding to a Koopman eigenfunction in } \mathcal{H}\}. \quad (6)$$

Note that the result combines properties of both the dynamics and the RKHS, namely that $\omega$ is a Koopman eigenfrequency, and $f_\omega$ an RKHS function. That is, it establishes a connection between the dynamics, which is determined solely by the flow $\Phi^t$ and the RKHS, which is determined solely by the kernel $k$.

Theorem 1 is in fact a consequence of Theorem 2 below, which is a general RKHS result that only depends on the dynamical orbit $\mathcal{O}$ lying dense in $X$. 

5
Theorem 2. Let $k : X \to X$ be a kernel satisfying Assumption $\mathbb{H}$, the corresponding RKHS, and $\mathcal{O} \subset X$ a dense, countable set whose first $N$ points are labeled $X_N$. Then,

(i) For every $f \in \mathcal{H}$, $\lim_{N \to \infty} \|T_N(f|X_N)\|_{\mathcal{H}} = \|f\|_{\mathcal{H}}$.

(ii) If $f : \mathcal{O} \to \mathbb{C}$ is such that $\|T_N(f|X_N)\|_{\mathcal{H}}$ does not diverge as $N \to \infty$, then $f$ has a unique extension in $\mathcal{H}$.

Remark. Theorem 2 holds in a general RKHS, and does not require an underlying time-flow or dynamics. The result is of significantly broader applicability than dynamical systems, as it gives necessary and sufficient conditions for functions on dense, countable sets to have RKHS extensions. We have stated it as one of the main results because we have found no similar result in the literature.

Spectral decomposition. Let $\mathcal{D}$ be the closed subspace of $L^2(\mu)$ spanned by the eigenfunctions of $U^t$, and $\mathcal{D}^\perp$ its orthogonal complement. Systems in which $\mathcal{D}$ contains non-constant functions and $\mathcal{D}^\perp$ is non-empty are called mixed-spectrum systems. The $L^2(\mu)$ space of a general measure-preserving system admits the $U^t$-invariant decomposition

$$L^2(X, \mu) = \mathcal{D} \oplus \mathcal{D}^\perp.$$  

(7)

In the spectral study of dynamical systems, it is a classical approach to study the dynamics separately on $\mathcal{D}$ and $\mathcal{D}^\perp$; see, e.g., [19]. This is because not only are these spaces invariant under $U^t$, they also represent the quasiperiodic and chaotic component in the underlying dynamics. In particular, every observable $f = \sum_k c_k z_k \in \mathcal{D}$ can be expanded in an orthonormal basis $\{z_k\}$ consisting of Koopman eigenfunctions, and thus has integrable (quasiperiodic) time evolution, $U^t = \sum_k e^{i\omega_k t} c_k z_k$, whereas observables $g \in \mathcal{D}^\perp$ have an expansion associated with the continuous spectrum of $U^t$ and an associated weak-mixing property, $\lim_{t \to \infty} t^{-1} \int_0^t \langle |h(U^t g)\rangle \mu \, ds = 0$, for all $h \in L^2(\mu)$, characteristic of chaotic evolution. In a data-driven setting, the invariant splitting in (7) was introduced in [8] in the context of harmonic averaging techniques, and was also employed in [13] in Galerkin approximation techniques for the eigenvalues and eigenfunctions of the generator of the Koopman group.

High-dimensional complex systems are typically of mixed spectrum, and for such systems an important task at hand is to identify the Koopman eigenbasis of $\mathcal{D}$ and the associated eigenfrequencies from data. The following result establishes data-driven criteria to determine whether a candidate frequency $\omega \in \mathbb{R}$ is a Koopman eigenfrequency, depending on whether or not it lies in the set of eigenfrequencies $\Omega$ from (6) corresponding to RKHS-extensible eigenfunctions.

Theorem 3. Let Assumptions [1] and [3] hold, and $w_N$ be as in (18). Then for every $N \in \mathbb{N}$, there exists a sequence of approximations $(w_{N, l})_{l=1}^\infty$ of $w_N$ such that the following hold for $x_0$ in a set of $\mu$-measure 1.

(i) If $\omega \notin \Omega$, then $\lim_{N \to \infty} w_{N, l}(f_\omega) = 0$.

(ii) If $\omega \in \Omega$, then by Theorem 1 there exists an RKHS extension $\tilde{f}_\omega$ of $f_\omega$, and moreover,

$$\lim_{t \to \infty} \lim_{N \to \infty} w_{N, l}(f_\omega) = \lim_{l \to \infty} w_{N, l}(f_\omega) = \lim_{N \to \infty} w_N(f_\omega) = \|\tilde{f}_\omega\|_\mathcal{H}^2.$$  

(iii) For any ordering $\omega_1, \omega_2, \ldots$ of the Koopman eigenvalues, one has for every $l \in \mathbb{N}$,

$$\lim_{k \to \infty} \lim_{N \to \infty} w_{N, l}(z_{\omega_k}) = 0.$$

Remark. The quantities $w_{N, l}$, which are explicitly defined in (18), Section 5, measure the square RKHS norm of $f_\omega|X_N$, projected onto an $l$-dimensional subspace of $\mathcal{H}_N$ containing functions of minimal “roughness”, as measured by the ratio $r_N$ from (5). Theorem 3 states that given $m$, there exists a fixed $l$ such that for large $N$, $w_{N, l}$ can be used to identify the frequencies corresponding to the first $m$ Koopman eigenfunctions ordered in order of increasing roughness. This is useful since, in many complex systems, the set of Koopman eigenfrequencies is dense on the real line, and there needs to be a robust way of separating frequencies. Under mild assumptions (see Section 8), the set of Koopman eigenfrequencies can be generated given knowledge of finitely many eigenfrequencies, and if such a generating set of frequencies is contained in $\Omega$, the recovered eigenfrequencies for large-enough $m$ are sufficient to determine all Koopman eigenfrequencies.
Uniformity of convergence. A natural question to ask is whether the convergence of \( w_{N,1}(f_\omega) \) to zero is uniform over \( \omega \notin \Omega \). A difficulty in addressing this question stems from the facts that the quantities \( w_{N,1}(f_\omega) \) depend on the initial point \( x_0 \), and \( \mathbb{R} \setminus \Omega \) is an uncountable set. As a result, the set of initial points \( x_0 \) for which \( \lim_{N \to \infty} w_{N,1}(f_\omega) \) is well-defined for all \( \omega \in \mathbb{R} \) is an uncountable intersection of sets of full measure, and may be empty (or non-measurable). One way of addressing this issue, is to consider the convergence of \( w_{N,1}(f_\omega) \) in expectation with respect to the starting point \( x_0 \in X \), which we assume to have distribution \( \mu \). Once the average is taken, one does not have to specify the appropriate set of initial points. The following result makes a statement in this setting about uniform convergence to zero, for systems which have an absolutely continuous Koopman spectrum with respect to Lebesgue measure (which means in particular, no eigenfrequencies). In what follows, \( \mathbb{E}_\mu \) will denote the expectation with respect to the initial point \( x_0 \in X \) distributed with measure \( \mu \).

**Theorem 4.** Let Assumptions 2 and 3 hold, and the quantities \( w_N, w_{N,1} \) be as defined in Theorem 3. Then if the Koopman group \( U^1 \) has an absolutely continuous spectrum, for every \( l \in \mathbb{N} \),

\[
\lim_{N \to \infty} \sup_{\omega \in \mathbb{R}} \mathbb{E}_\mu w_{N,1}(f_\omega) = 0.
\]

Remark. One of the consequences of having an absolutely continuous spectrum is that the discrete component \( D \) from (7) contains only constant functions. These systems are always weakly mixing (see Mixing, [19]), but the converse is not true, for example, there are weakly mixing systems whose spectral measure, besides having an absolutely continuous component, also has a singular continuous component (e.g., [20], p. 118).

An illustration of Theorem 4 based on the Lorenz 63 (L63) system is included in Section 9.

Note that in practice one always scans for eigenvalues over a countable subset \( \Omega' \subset \mathbb{R} \); for example, in the case of DFT,

\[
\Omega' = \{ j\pi/(N \Delta t) : N \in \{1, 3, \ldots\}, \, j \in \{- (N - 1)/2, \ldots, (N - 1)/2\} \},
\]

where we have assumed that the number of samples \( N \) is odd for simplicity. We will use the same set of trial frequencies in the numerical implementation of our techniques, presented in Section 8.

Outline of the paper. We first prove Theorem 1 in Section 3 by invoking Theorem 2. In Section 4, we review some important concepts from RKHS theory. Next, we prove Theorem 2 in Section 5, Theorem 3 in Section 6, and Theorem 4 in Section 7. In Section 8, we discuss the numerical realization of our methods. In Section 9, the methods are applied to various systems with different types of spectrum, and compared with regular Fourier analysis of signals.

3. Proof of Theorem 1

The proof of Theorem 1 requires two lemmas. The first shows that the projection of an \( L^2(\mu) \) function onto a Koopman eigenspace is also the limit of an exponentially weighted Birkhoff average of the function.

**Lemma 5.** The orthogonal projection \( \pi_\omega f \) of a function \( f \in L^2(\mu) \) onto the eigenspace of \( U^{\Delta t} \) corresponding to the eigenvalue \( e^{i\omega \Delta t} \) is given by the \( L^2(\mu) \)-limit

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} e^{-i\omega n \Delta t} U^n f = \pi_\omega f.
\]

Moreover, if \( \omega \) is not an eigenfrequency, then \( \pi_\omega \equiv 0 \).

**Proof.** Let \( L_\omega \) be the subspace of fixed points of the unitary operator \( W_\omega = e^{-i\omega \Delta t} U^{\Delta t} \), and \( \text{proj}_{L_\omega} : L^2(\mu) \to L^2(\mu) \) the corresponding orthogonal projection operator. This subspace is \( \{0\} \) if \( \omega \) is not a Koopman eigenfrequency; otherwise, it is the eigenspace of \( U^{\Delta t} \) corresponding to eigenvalue \( e^{i\omega \Delta t} \), so that
proj\_L_\omega = \pi_\omega. By the von Neumann mean ergodic theorem (e.g., [\ref{ref:19}]), \(N^{-1}\sum_{n=0}^{N-1} W^n\omega\) converges pointwise to proj\_L_\omega, and therefore, for any \(f \in L^2\),

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} e^{-in\omega\Delta t} U^n f = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} W^n f = \text{proj}\_L f = \pi_\omega f. \tag{\ref{eq:4}}
\]

The second result states that under mild conditions on the kernel, all RKHS functions are continuous.

**Lemma 6** ([\ref{ref:15}], Lemma 2.1). Given a closed, connected topological space \(X\) and a positive semi-definite kernel \(k\) on \(X\), the RKHS associated with \(k\) is a subspace of \(C^0(X)\).

Theorem [\ref{thm:1}] is now ready to be proved. Theorem [\ref{thm:2}] will be assumed to be true in the proof.

**Proof of Claim (i).** The “if” part of the claim is the negation of the claim in Theorem [\ref{thm:2}]ii). To verify the “only if” part, suppose that frequency \(\omega \in \mathbb{R}\) does not have a corresponding Koopman eigenfunction \(z_\omega \in \mathcal{H}\). If \(f_\omega\) had an extension \(\tilde{f}_\omega \in \mathcal{H}\), then that extension would be continuous by Lemma [\ref{lem:6}] and thus lie in \(L^2(\mu)\). Therefore, by Lemma [\ref{lem:5}] for \(\mu\)-a.e. \(x_0 \in X\), \(\sum_{n=0}^{N-1} e^{-in\omega\Delta t} f_\omega(x_n), x_n \in \mathcal{O}\), would converge to zero as \(N \to \infty\). However, each of the summands \(e^{-in\omega\Delta t} f_\omega(x_n)\) is identically equal to 1, so their average is equal to 1, leading contradiction. Therefore, \(f_\omega\) does not have an extension to \(\mathcal{H}\), and it follows from Theorem [\ref{thm:5}]ii) that \(\lim_{N \to \infty} w_N(f_\omega \mid X_N) = \infty\).

**Proof of Claim (ii).** The claim follows from Theorem [\ref{thm:2}]ii).

**Proof of Claim (iii).** First, observe that by strong continuity of the group \((U_t)_{t \in \mathbb{R}}\) it sufficient to show that any extension \(\tilde{f}_\omega \in \mathcal{H}\) of \(f_\omega\) is an eigenfunction of \(U^{\Delta t}\). Indeed, for any \(x_n \in \mathcal{O}\), we have

\[
U^{\Delta t} \tilde{f}_\omega(x_n) = \tilde{f}_\omega(x_{n+1}) = \tilde{f}_\omega(x_n + \Delta t) = e^{i\omega \Delta t} \tilde{f}_\omega(x_n) = e^{i\omega \Delta t} \tilde{f}_\omega(x_n) = e^{i\omega \Delta t} \tilde{f}_\omega(x_n),
\]

and since \(\tilde{f}_\omega\) is continuous (by Lemma [\ref{lem:5}]), and \(\mathcal{O}\) lies dense in \(X\), it follows that \(U^{\Delta t} \tilde{f}_\omega(x) = e^{i\omega \Delta t} \tilde{f}_\omega(x)\) for all \(x \in X\), proving the claim. This also completes our proof of Theorem [\ref{thm:1}].

4. Results from reproducing kernel Hilbert space theory

In this section, we briefly review a number of properties of RKHSs which will be employed in the proofs of Theorems [\ref{thm:2}] and [\ref{thm:3}]. For a more detailed exposition of this material we refer the reader to [\ref{ref:21}], or one of the many other references on RKHS theory.

**Convergence in RKHS and \(C^0\) norms.** If the reproducing kernel \(k : X \times X \to \mathbb{C}\) of an RKHS \(\mathcal{H}\) on a compact space \(X\) is continuous, then convergence in \(\mathcal{H}\) norm implies convergence in \(C^0(X)\) norm, which in turn implies pointwise convergence. In particular, given any \(f, g \in \mathcal{H}\) and \(x \in X\), it follows from the Cauchy-Schwartz inequality and the reproducing property that

\[
|f(x) - g(x)| = |\langle f - g, k(x, \cdot) \rangle_{\mathcal{H}}| \leq \|f - g\|_{\mathcal{H}} \|k(x, \cdot)\|_{\mathcal{H}}
\]

and

\[
\|k(x, \cdot)\|_{\mathcal{H}}^2 = \langle k(x, \cdot), k(x, \cdot) \rangle_{\mathcal{H}} = |k(x, x)| \leq \|k\|_{C^0},
\]

respectively, leading to

\[
\|f - g\|_{C^0}^2 = \max_{x \in X} |f(x) - g(x)| \leq \|f - g\|_{\mathcal{H}}^2 \max_{x \in X} |k(x, x)| \leq \|f - g\|_{\mathcal{H}}^2 \|k\|_{C^0}.
\]
**Kernel integral operators.** Kernel integral operators are compact operators on \( L^2(\mu) \) space, which provide a convenient way of realizing the RKHS associated with the kernel \( k \). Specifically, given a Borel probability measure \( \nu \) with compact support \( \text{supp}(\nu) \subseteq X \), the kernel integral operator \( S_\nu : L^2_X \mapsto C^0(X) \) associated with a continuous kernel \( k \) on \( X \) is defined by

\[
S_\nu := f \mapsto \int_X k(\cdot, x) f(x) \, d\nu(x).
\]

Let \( \mathcal{H}_\nu \) be the RKHS on \( \text{supp}(\nu) \) with kernel \( k_\nu = k \| \text{supp}(\nu) \times \text{supp}(\nu) \). It can be verified that \( \text{ran} \, S_\nu \) is a dense subspace of \( \mathcal{H}_\nu \), and \( S_\nu : L^2_X \mapsto \mathcal{H}_\nu \) is compact. Moreover, the adjoint map \( S_\nu^* : \mathcal{H}_\nu \mapsto L^2_X \) coincides with the inclusion map from \( C^0(X) \) to \( L^2_X \). As a result, for every \( f \in L^2(X, \nu) \), \( g \in \mathcal{H}_\nu \), and \( \nu \)-a.e. \( x \in X \),

\[
S_\nu^* f(x) = f(x), \quad \langle g, S_\nu f \rangle_{\mathcal{H}_\nu} = \langle g, f \rangle_\nu.
\]

Note that \( \mathcal{H}_\nu \) naturally embeds into \( \mathcal{H} \) through the linear isometry \( \sum_{n \in \mathbb{N}} a_n k(\cdot, x_n) \in \mathcal{H}_\nu \mapsto \sum_{n \in \mathbb{N}} a_n k(\cdot, x_n) \in \mathcal{H} \), which means that we can view \( \mathcal{H}_\nu \) as a subspace of \( \mathcal{H} \). Clearly, \( \mathcal{H}_\nu = \mathcal{H} \) if \( \text{supp}(\nu) = X \).

Consider now the operator \( G_\nu := S_\nu^* S_\nu \) on \( L^2_X \). This operator is a Hilbert-Schmidt, and therefore compact, self-adjoint, positive-semidefinite operator, and there exists an orthonormal basis \( \{ \phi_i \}_{i=0}^\infty \) of \( L^2(\mu) \) consisting of its eigenvectors. By convention, we order the basis elements \( \phi_i \) in order of decreasing corresponding eigenvalues, \( \lambda_i \geq 0 \), which converge monotonically to 0 as \( i \to \infty \) by compactness of \( G \).

Let now \( J = \{ i \in \mathbb{N} : \lambda_i \neq 0 \} \) be the index set for the nonzero eigenvalues of \( G_\nu \), and define the set \( \{ \psi_i \}_{i \in J} \),

\[
\psi_i = \lambda_i^{-1/2} S_\nu \phi_i.
\]

It follows from (8) that the \( \psi_i \) form an orthonormal set on \( \mathcal{H}_\nu \). If, in addition, \( k \) is continuous, then it follows from Mercer’s theorem that \( k(x, y) = \sum_{i \in J} \psi_i(x) \overline{\psi}_i(y) \), uniformly on \( \text{supp}(\nu) \times \text{supp}(\nu) \), which implies in turn that \( \{ \psi_i \}_{i \in J} \) is an orthonormal basis of \( \mathcal{H}_\nu \). The range of \( S_\nu \) can also be expressed in terms of the \( \psi_i \) as

\[
\text{ran} \, S_\nu = \left\{ f = \sum_{i \in J} b_i \psi_i : (\lambda_i^{-1/2} b_i)_{i \in J} \in l^2 \right\}.
\]

In particular, because \( \lambda_i \to 0 \) as \( i \to \infty \), this shows that the range of \( S_\nu \) is always a proper dense subspace of \( \mathcal{H}_\nu \), as stated above, unless \( \mathcal{H}_\nu \) is finite dimensional.

In what follows, the measure \( \nu \) will be either the invariant ergodic measure \( \mu \) with full support on \( X \), or a sampling measure \( \mu_N \) with finite discrete support. We will use the abbreviated notations \( S_\mu = S \) and \( S_{\mu_N} = S_N \). Note in particular that the action \( S_N : L^2(\mu_N) \mapsto \mathcal{H} \) on \( L^2(\mu_N) \) equivalence classes corresponds to weighted averages of kernel sections, viz.

\[
S_N f \mapsto \frac{1}{N} \sum_{n=0}^{N-1} k(\cdot, x_n) f(x_n).
\]

We will denote the eigenvalues and eigenfunctions of \( G_N := S_N^* S_N \) by \( \lambda_{N,i} \) and \( \phi_{N,i} \), respectively.

**Nyström extension.** The map \( T_N \) used in [4] can be built using a procedure called Nyström extension. For a general measure \( \nu \) with compact support in \( X \), the Nyström extension operator \( T_\nu : D(T_\nu) \mapsto \mathcal{H}_\nu \) has domain

\[
D(T_\nu) = \left\{ f = \sum_{i \in J} a_i \phi_i \in L^2(X, \nu) : (\lambda_i^{-1/2} a_i)_{i \in J} \in l^2 \right\},
\]

and is action on every such \( f \) is given by

\[
T_\nu f = \sum_{i \in J} \lambda_i^{-1/2} a_i \psi_i.
\]
A key property of this operator, which is a consequence of (8), is $S_T^* T_{\nu} f = f$ for all $f \in L^2(X, \nu)$. Since $S_T^*$ is an inclusion map, this shows that $T_{\nu} f = f$ $\nu$-a.e., and thus that $T_{\nu}$ extends $L^2_{\nu}$ equivalence classes in its domain to $\mathcal{H}$ functions. If, in addition, $S_{\nu}$ is a positive operator (i.e., all eigenvalues $\lambda_i$ are positive), which occurs if and only if $k_{\nu}$ is strictly positive definite, then it follows from (10) that $D(T_{\nu}u)$ is dense in $L^2_{\nu}$.

The Nyström extension operator in (4) is given by $T_N = T_{\mu_N}$, and in summary, for every $f = \sum_{i \in J} a_i \phi_{N,i} \in D(T_N)$, we have:

$$T_N f = \sum_{i \in J} a_i \lambda_{N,i}^{-1/2} \psi_{i,k}, \quad \psi_{N,i} = \lambda_{N,i}^{-1/2} S_N \phi_{N,i}, \quad S_N^* \psi_{N,i} = \lambda_{N,i}^{1/2} \phi_{N,i}.$$  \hspace{1cm} (12)

Note that, in general, $J \subseteq \{1, \ldots, N\}$, and equality holds if Assumption 3 is satisfied. In that case, $D(T_N) = L^2(\mu_N)$.

5. Proof of Theorem 2

We will need the following lemma for the proof, which shows that the data-driven finite dimensional subspaces $\mathcal{H}_N$ “converge” to the RKHS $\mathcal{H}$.

**Lemma 7.** Let the assumptions of Theorem 2 hold. Then, for $\mu$-a.e. $x_0 \in X$ and every $N \in \mathbb{N}$, the subspace $\mathcal{H}_N \subset \mathcal{H}$ is $N$-dimensional, and $\mathcal{H}_1, \mathcal{H}_2, \ldots$ forms a nested sequence of subspaces of $\mathcal{H}$. Moreover, $\cup_{N \in \mathbb{N}} \mathcal{H}_N$ is dense in $\mathcal{H}$; that is, $\mathcal{H}$ is the closure of the span of $\{k(\cdot, y) : y \in \mathcal{O}\}$.

**Proof.** First, note that since $\mathcal{H}_N = \{k(x_0, \cdot), \ldots, k(x_{N-1}, \cdot)\}$, it is clear that $\mathcal{H}_N \subset \mathcal{H}_{N+1}$. Moreover, by Assumption 3(ii), the $N$ kernel sections $k(x_0, \cdot), \ldots, k(x_{N-1}, \cdot)$ are linearly independent, and therefore $\mathcal{H}_N$ is $N$-dimensional.

Next, consider the closed subspace $W = \bigcup_{N \in \mathbb{N}} \mathcal{H}_N = \text{span}\{k(\cdot, y) : y \in \mathcal{O}\}$, where closure is taken with respect to $\mathcal{H}$ norm. It has to be shown that $W = \mathcal{H}$, or equivalently that $W^\perp = \{0\}$. But $f \in W^\perp$ iff for every $y \in \mathcal{O}$, $(k(\cdot, y), f)_{\mathcal{H}} = 0$. Since $k$ is the reproducing kernel of $\mathcal{H}$, this is equivalent to saying that $f|\mathcal{O} \equiv 0$. Now, since $\text{supp}(\mu) = X$, for $\mu$-a.e. $x_0 \in X$, the orbit $\mathcal{O}$ is dense in $X$. Thus, for every such $x_0$, $f$ is zero on a dense subset of $X$. However, by Assumption 3(i) and Lemma 6, $\mathcal{H} \subset C^0(X)$, and thus $f$ is continuous and equal to $0$ on the entire space $X$. This shows that $W^\perp = \{0\}$, as claimed. \hfill $\Box$

Theorem 2 is now ready to be proved.

**Proof of Claim (i).** We will first express $f|X_N$ as the result of applying a combination of operators from $\mathcal{H}$ to $L^2(\mu_N)$. Let $P_N : C^0(X) \rightarrow L^2(\mu_N)$ be the restriction map, satisfying $P_N f(x_0) = f(x_0)$ . Let also $P : C^0(X) \rightarrow L^2(\mu)$ be the canonical inclusion map. As stated above, if the kernel $k$ is continuous $\mathcal{H}$ is a subspace of $C^0(X)$, and therefore there exist inclusion maps $j : \mathcal{H} \rightarrow C^0(X)$ and $j_N : \mathcal{H}_N \rightarrow C^0(X)$. The diagram below shows how $P_N, j, T_N$ are related:

$$\begin{array}{c}
L^2(\mu_N) \\
\downarrow P_N \\
\mathcal{H} \subset \mathcal{H} \\
\downarrow j_N \\
C^0(X)
\end{array}$$  \hspace{1cm} (13)

Since $\mathcal{H}_N$ is a finite-dimensional and hence closed subspace of $\mathcal{H}$, there exists an orthogonal projection $\pi_N : \mathcal{H} \rightarrow \mathcal{H}$ with ran $\pi_N = \mathcal{H}_N$. By Lemma 7, $\mathcal{H}_N$ is a sequence of nested subspaces whose union is dense in $\mathcal{H}$, hence

$$\lim_{N \rightarrow \infty} \|\pi_N f - f\|_{\mathcal{H}} = 0, \quad \forall f \in \mathcal{H}. \hspace{1cm} (14)$$

Next, observe that for every $f \in \mathcal{H}$, $f|X_N = P_N j f$, and therefore $T_N(f|X_N) = T_N P_N j f$. Thus, Claim (i) will be proved if it can be shown that the map $\pi_N = T_N P_N j : \mathcal{H} \rightarrow \mathcal{H}$ is the same as the orthogonal projection $\pi_N$, for, in that case, $\|\pi_N f\|_{\mathcal{H}}$ and thus $\|T_N f|X_N\|_{\mathcal{H}}$ will converge to $\|f\|_{\mathcal{H}}$ by (14).
To prove that $\pi'_N = \pi_N$, it will be first shown that $\pi'_N$ is an idempotent operator, i.e., $\pi'_N \pi'_N = \pi'_N$. Indeed, by (13), $P_N j T_N$ is the identity map on $L^2(\mu_N)$, and therefore

$$\pi'_N \pi'_N = T_N P_N j T_N P_N j = T_N (P_N j T_N) P_N j = T_N P_N j = \pi'_N.$$  

Second, it will be shown that the range of $\pi'_N - I$ is orthogonal to $H_N$. Again, by (13),

$$(P_N j) \circ \pi'_N = P_N j T_N P_N j \equiv P_N j,$$

which implies that $(\pi'_N g)(x_n) = g(x_n)$ and therefore

$$\langle \pi'_N g - g, k(\cdot, x_n) \rangle_H = (\pi'_N g)(x_n) - g(x_n) = 0,$$

for every $g \in H$ and $n \in \{0, \ldots, N - 1\}$. Now, since $H_N$ is spanned by $\{k(\cdot, x_0), \ldots, k(\cdot, x_{N-1})\}$, $\pi'_N g - g$ is orthogonal to $H_N$, and therefore $\pi'_N$ is an orthogonal projection, as claimed. This completes the proof of Claim (i).

**Proof of Claim (ii).** Under the assumptions of the claim, $\{T_N(f|x_N) : N \in \mathbb{N}\}$ is bounded in $H$. Therefore, because every bounded sequence in a Hilbert space has a weakly convergent subsequence, $T_N(f|x_N)$ has a weakly convergent subsequence. The proof will be completed by the following proposition.

**Proposition 8.** Under Assumption [7], if $f : \mathcal{O} \rightarrow C$ is such that the sequence of functions $h_N = T_N f|X_N$ has a weakly-convergent subsequence in $H$, then $f$ has a unique extension to $H$.

**Proof.** Let $(h_N)_{j=0}^{\infty}$ be such a weakly-convergent subsequence, and $h$ its weak limit. Note that by definition of $T_N$, for fixed $x_n \in \mathcal{O}$ and every $N > n$, $h_N(x_n)$ is constant and equal to $f(x_n)$. Therefore, by definition of weak convergence,

$$h(x_n) = \langle h, k(x_n, \cdot) \rangle_H = \lim_{j \rightarrow \infty} \langle h_N, k(x_n, \cdot) \rangle_H = \lim_{j \rightarrow \infty} h_N(x_n) = f(x_n), \quad \forall n \in \{0, 1, 2, \ldots\},$$

which shows that $h$ is an extension of $f$ to $H$. The uniqueness of $h$ follows from the fact that it is continuous (since $H \subseteq C^0(X)$), and $\mathcal{O}$ is dense.

As a side note, we will mention the following corollary of Lemma [7].

**Corollary 9.** For every $t \in \mathbb{R}$, the space $H \circ \Phi^t$ obtained by composing every element of $H$ by the flow $\Phi^t$, is an RKHS with reproducing kernel $k^{(t)} : X \times X \rightarrow \mathbb{R}$, $k^{(t)}(x, y) = k(\Phi^t(x), \Phi^t(y))$. In particular, if the kernel $k$ is invariant under the flow, i.e., if for every $t \in \mathbb{R}$, $k \equiv k^{(t)}$, then $H \circ \Phi^t = H$.

**Proof.** Let $H^{(t)}$ be the RKHS with reproducing kernel $k^{(t)}$, then for every $N \in \mathbb{N}$, the subspace $H^{(t)}_N$ can be defined analogously to $H_N$. The important observation is that

$$H_N \circ \Phi^t = H^{(t)}_N,$$

which, in conjunction with Lemma [7], implies that $H \circ \Phi^t = H^{(t)}$. Now assume that $k$ is flow invariant. Every $f \in H$ is a sum of the form $f = \sum_{i \in \mathbb{N}} a_i k(y_i, \cdot)$, and therefore, for any $t \in \mathbb{R},$

$$\|f\|^2_H = \sum_{i, j \in \mathbb{N}} a_i^* a_j k(y_i, y_j) = \sum_{i, j \in \mathbb{N}} \sum_{i, j \in \mathbb{N}} a_i a_j k(\Phi^{-t}(y_i), \Phi^{-t}(y_j)) = \sum_{i \in \mathbb{N}} a_i k(\Phi^{-t}(y_i), \cdot),$$

and we conclude that $\tilde{f} = \sum_{i \in \mathbb{N}} a_i k(\Phi^{-t}(y_i), \cdot)$ lies in $H$. However,

$$\tilde{f}(x) = \sum_{i \in \mathbb{N}} a_i k(\Phi^{-t}(y_i), x) = \sum_{i \in \mathbb{N}} a_i k(y_i, \Phi^t(x)) = f \circ \Phi^t(x),$$

which shows that $f \circ \Phi^t = \tilde{f}$ also lies in $H$, as claimed.
By Corollary 9, one of the consequences of having a flow-invariant kernel is that one can define a group of Koopman operators \( U^t : \mathcal{H} \mapsto \mathcal{H}, \ t \in \mathbb{R} \), acting on the corresponding RKHS, \( \mathcal{H} \). Flow-invariant kernels are a more special class of kernels, as they incorporate information about the underlying dynamics. One way of realizing them in a data-driven environment is as sequences of kernels operating on delay-embedded data with an increasing number of delays [13]. It can be shown that the spectra of kernel integral operators associated with such kernels have a close correspondence with those of the unitary Koopman operators on \( L^2(\mu) \); see [13].

6. Proof of Theorem 3

Let \( \langle \cdot, \cdot \rangle_{\mu_N} \) denote the inner product of \( L^2(\mu_N) \). Under the assumptions of the theorem, \( \{\phi_{N,0}, \ldots, \phi_{N,N-1}\} \) is an orthonormal basis of \( L^2(\mu_N) \) consisting of eigenfunctions of \( G \) with nonzero corresponding eigenvalues, \( \lambda_{N,j} \). Therefore, the Fourier function \( f_\omega |X_N \) on the finite trajectory can be expressed as

\[
f_\omega |X_N = \sum_{j=0}^{N-1} \langle \phi_{N,j}, f_\omega |X_N \rangle_{\mu_N} \phi_{N,j}, \tag{15}\]

and by definition of the Nyström extension in [12],

\[
T_N(f_\omega |X_N) = \langle f_\omega |X_N, \phi_{N,j} \rangle_{\mu_N} \psi_{N,j} / \lambda_{N,j} \tag{16}\]

and

\[
w_N(f_\omega) = \|T_N(f_\omega |X_N)\|^2 = \sum_{j=0}^{N-1} |\langle f_\omega |X_N, \phi_{N,j} \rangle_{\mu_N}|^2 / \lambda_{N,j}. \tag{17}\]

It follows from the above that \( w_N(f_\omega) \) can be approximated by the sequence of spectrally truncated norms

\[
w_{N,l}(f_\omega) := \sum_{j=0}^{l-1} |\langle f_\omega |X_N, \phi_{N,j} \rangle_{\mu_N}|^2 / \lambda_{N,j}, \quad l \in \{1, \ldots, N\}, \tag{18}\]

with \( w_{N,N}(f_\omega) = w_N(f_\omega) \).

Remark. Theorems 1 and 3 establish criteria to determine whether a Fourier function \( f_\omega \) extends to a Koopman eigenfunction in RKHS based on the limiting behavior of either \( w_N(f_\omega) \) or \( w_{N,l}(f_\omega) \). In particular, Theorem 1 implies that if \( \lim_{N \to \infty} w_N(f_\omega) \) is finite, then \( \omega \) is a Koopman eigenfrequency, and similarly, by Theorem 3, \( \omega \) is a Koopman eigenfrequency if \( \lim_{N \to \infty} w_{N,l}(f_\omega) \neq 0 \). Between these two criteria, the one based on \( w_{N,l}(f_\omega) \) is stronger as it implies the one based on \( w_N(f_\omega) \). In practical applications, however, the latter is arguably more useful than the former. In particular, as is evident from (17), \( w_{N,l}(f_\omega) \) depends on all of the eigenpairs \( (\lambda_{N,j}, \phi_{N,j}) \) of \( G_N \), but for a given \( N \), as \( j \) increases the eigenvalues become increasingly sensitive to the particular trajectory \( X_N \); that is, \( w_{N,l}(f_\omega) \) has high sensitivity to sampling errors. On the other hand, \( w_{N,l}(f_\omega) \) depends on a fixed number \( l \) of eigenvalues and eigenfunctions, which converge as \( N \to \infty \) uniformly with respect to \( j \in \{0, \ldots, l-1\} \) for \( \mu \)-a.e. starting state.

Next, observe that the quantities \( \langle \phi_{N,j}, f_\omega |X_N \rangle_{\mu_N} \) in (15) are functions of the starting state \( x_0 \) in the sampled orbit. The following proposition establishes the limit as \( N \to \infty \) of these quantities, as \( L^2(\mu) \) functions of \( x_0 \).

Proposition 10. Let \( \phi_j \in L^2(\mu) \) be an eigenfunction of \( G \) corresponding to a nonzero eigenvalue. Then, under the assumptions of Theorem 3, there exists a sequence of eigenfunctions \( \phi_{N,j} \) of \( G_N \) such that, for every \( \omega \in \mathbb{R} \), every \( j \in \mathbb{N} \), and \( \mu \)-a.e. \( x_0 \in X \),

\[
\lim_{N \to \infty} \langle f_\omega |X_N, \phi_{N,j} \rangle_{\mu_N} = \pi_\omega \phi_j.
\]

In particular, \( \lim_{N \to \infty} \langle f_\omega |X_N, \phi_{N,j} \rangle_{\mu_N} \) is \( \mu \)-a.e. equal to zero if \( \omega \notin \Omega \), and otherwise \( \mu \)-a.e. equal to the projection of the eigenfunction \( \phi_j \) onto the the Koopman eigenspace corresponding to the eigenfrequency \( \omega \).
Proof. Only the case $\omega \notin \Omega$ will be proved, as the proof for the other case is analogous. We begin from a result based on Theorem 15 in [22] and Corollary 4 in [13], which states that given a eigenfunction $\phi_j$ corresponding to a nonzero eigenvalue and with a corresponding RKHS function $\psi_j$ from (9), there exists a sequence $\phi_{N,j}$ of eigenfunctions of $S_N$ (also at nonzero eigenvalue) such that the corresponding the RKHS functions $\psi_{N,j}$ converge to $\psi_j$ in $C^0$ norm, viz.

$$\lim_{N \to \infty} \|\psi_{N,j} - \psi_j\|_{C^0} = 0, \ \forall j \in \mathbb{N}. \quad (19)$$

Let now $\omega \notin \Omega$ and $j \in \mathbb{N}$ be fixed. Then,

$$\lim_{N \to \infty} \langle f_x | X_N, \phi_{N,j} \rangle_{\mu_N} = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} e^{-i \omega \Delta t} \phi_{N,j}(x_n) = \lim_{N \to \infty} \frac{\lambda_{N,j}}{N} \sum_{n=0}^{N-1} e^{-i \omega \Delta t} \psi_{N,j}(x_n).$$

Therefore, the claim will be proved if it can be shown that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} e^{-i \omega \Delta t} \psi_{N,j}(x_n) = 0, \ \forall \omega \notin \Omega, \ \forall j \in \mathbb{N}. \quad (20)$$

To show that (20) holds, let $\epsilon > 0$ be arbitrary, and note that

$$1 \sum_{n=0}^{N-1} e^{-i \omega \Delta t} \psi_{N,j}(x_n) = \frac{1}{N} \sum_{n=0}^{N-1} e^{-i \omega \Delta t} \psi_j(x_n) + \frac{1}{N} \sum_{n=0}^{N-1} e^{-i \omega \Delta t} [\psi_{N,j}(x_n) - \psi_j(x_n)],$$

and. By Lemma 5 for $\mu$-a.e. $x_0 \in X$, there exists $N_1 \in \mathbb{N}$ such that for $N > N_1$, the first sum in the right-hand side is less than $\epsilon$ in magnitude. Moreover, by (19), there exists $N_2 \in \mathbb{N}$ such that for $N > N_2$, $\|\psi_{N,j} - \psi_j\|_{C^0} < \epsilon$, and thus the second sum in the right-hand side is less than $\epsilon$ in magnitude. We therefore conclude that for $N > \max(N_1, N_2)$, the right-hand side is less than $2\epsilon$, and since $\epsilon$ was arbitrarily chosen, it converges to 0 as $N \to \infty$. Therefore, (20) holds, and Proposition 10 is proved.

Remark. If the eigenvalue $\lambda_j$ is simple, then the eigenfunctions $\phi_{0,j}, \phi_{1,j}, \ldots$ underlying the sequence $\psi_{N,j}$ converging to $\psi_j$ are only arbitrary up to phase. If $\lambda_j$ is not simple, then $T_N W_{N,j}$ converges as a subspace of $C^0(X)$ to $TW_j$, where $W_{N,j}$ and $W_j$ are the eigenspaces of $G_N$ and $G$ corresponding to eigenvalues $\lambda_{N,j}$ and $\lambda_j$, respectively. See [13], Appendix A, for further details. As a result, for every $\psi_j \in TW_j$, there exists a sequence of eigenfunctions $\phi_{N,j} \in W_{N,j}$ whose RKHS extensions $\psi_{N,j}$ converge in $C^0$ norm to $\psi_j$.

The claims of Theorem 3 can now be proved.

Proof of Claims (i), (ii). Since the right-hand side of (18) has a fixed, finite number of summands,

$$\lim_{N \to \infty} w_{N,l}(f_x) = \lim_{N \to \infty} \sum_{j=0}^{l-1} |\langle f_x | X_N, \phi_{N,j} \rangle_{\mu_N}|^2 / \lambda_{N,j} = \lim_{N \to \infty} \sum_{j=0}^{l-1} \frac{|\langle f_x | X_N, \phi_{N,j} \rangle_{\mu_N}|^2}{\lambda_{N,j}}.$$

If $\omega \notin \Omega$, then by Proposition 10, each of the $l$ limits is equal to 0, and Claim (i) is proved. If $\omega \in \Omega$, then again by Proposition 10, $\lim_{N \to \infty} w_{N,l}(f_x) = \sum_{j=0}^{l-1} |(z_0, \phi_j)_\mu|^2 / \lambda_{N,j}$, and because $\{(\phi_0, \phi_1, \ldots)\}$ is an orthonormal basis, the sum converges to $\|z_0\|^2_H$ as $l \to \infty$, proving Claim (ii).

Proof of Claim (iii). For any ordering $\omega_k$ of the Koopman eigenfrequencies, the corresponding eigenfunctions $(z_{\omega_k})_{k \in \mathbb{N}}$ form an orthonormal sequence in $L^2(\mu)$, and therefore the sequence $z_{\omega_k}$ converges weakly to 0. By Proposition 10 for $\mu$-a.e. $x_0 \in X$,

$$\lim_{k \to \infty} \lim_{N \to \infty} w_{N,l}(z_{\omega_k}) = \lim_{k \to \infty} \lim_{N \to \infty} \sum_{j=0}^{l-1} |\langle f_{\omega_k} | X_N, \phi_{N,j} \rangle_{\mu_N}|^2 / \lambda_{N,j} = \lim_{k \to \infty} \sum_{j=0}^{l-1} \frac{|\langle f_{\omega_k}, \phi_j \rangle_{\mu}|^2}{\lambda_j}.$$
Since \( t < \infty \), the limit can be brought inside the sum, and by the weak convergence of the \( z_{\omega k} \) to zero, we obtain

\[
\lim_{k \to \infty} \lim_{N \to \infty} w_{N,j}(z_{\omega k}) = \sum_{j=0}^{\infty} \| (f_{\omega_k}, \phi_j)_{\mu} \|^2 / \lambda_j = 0.
\]

This proves Claim (iii), concluding the proof of Theorem 3.

\[ \square \]

7. Proof of Theorem 4

In what follows, \( \Var_{\mu} \) will denote the variance with respect to the initial point \( x_0 \in X \) distributed with measure \( \mu \). Note that the inner product \( (f_{\omega}, \phi)_{\mu N} \) may not be well defined for the \( L^2(\mu) \) functions \( f_{\omega} \) and \( \phi_j \), but the quantity \( \Var_{\mu}(f_{\omega}, \phi)_{\mu N} \) is well defined. Moreover, while the eigenfunctions \( \phi_j \) are \( L^2(\mu) \) functions, they have continuous representatives and henceforth, will be considered as continuous functions.

We begin by computing the bound

\[
|\langle f_{\omega}, \phi_{N,j} \rangle_{\mu N} - \langle f_{\omega}, \phi_j \rangle_{\mu N}| = \left| \frac{1}{N} \sum_{n=0}^{N-1} f_{\omega}(x_n) [\phi_{N,j}(x_n) - \phi_j(x_n)] \right|
\]

\[
\leq \frac{1}{N} \sum_{n=0}^{N-1} |f_{\omega}(x_n)| |\phi_{N,j}(x_n) - \phi_j(x_n)|
\]

\[
= \frac{1}{N} \sum_{n=0}^{N-1} |\psi_{N,j}(x_n) - \psi_j(x_n)| \leq \| \psi_{N,j} - \psi_j \|_{C^0}.
\]

As in (19), for every \( j \in \{0, 1, \ldots, \| \psi_{N,j} - \psi_j \|_{C^0} \) vanishes as \( N \to \infty \); therefore,

\[
\lim_{N \to \infty} \sup_{\omega \in \mathbb{R}} |\langle f_{\omega}, \phi_{N,j} \rangle_{\mu N} - \langle f_{\omega}, \phi_j \rangle_{\mu N}| = 0. \tag{21}
\]

Based on these results, Theorem 4 can be proved if it can be shown that for every \( \phi \in L^2(\mu) \),

\[
\lim_{N \to \infty} \sup_{\omega \in \mathbb{R}} \Var_{\mu}(f_{\omega}, \phi)_{\mu N} = 0. \tag{22}
\]

Indeed, since \( t \) is fixed and finite, (22) and Proposition 10 imply the claim of the theorem,

\[
\lim_{N \to \infty} \sup_{\omega \in \mathbb{R}} \Ex_{\mu} w_{N,t}(f_{\omega}) = \lim_{N \to \infty} \sup_{\omega \in \mathbb{R}} \Ex_{\mu} \sum_{j=0}^{\infty} |\langle f_{\omega}, \phi_{N,j} \rangle_{\mu N}|^2 / \lambda_{N,j}
\]

\[
= \lim_{N \to \infty} \sum_{j=0}^{\infty} \frac{1}{\lambda_{N,j}} \sup_{\omega \in \mathbb{R}} \Ex_{\mu} |\langle f_{\omega}, \phi_j \rangle_{\mu N}|^2 = \lim_{N \to \infty} \sup_{\omega \in \mathbb{R}} \Var_{\mu}(f_{\omega}, \phi_j)_{\mu N} = 0,
\]

where the second equality above follows from (21). It thus remains to prove (22).

To that end, let \( F_{\omega,N} : L^2(\mu) \to L^2(\mu) \) be the discrete Fourier operator, defined for every \( \omega \in \mathbb{R} \) and \( N \in \mathbb{N} \) as the unitary operator

\[
F_{\omega,N} := \frac{1}{N} \sum_{n=0}^{N-1} e^{-i\omega n} f_n \Delta t.
\]

Note that \( F_{\omega,N} \) maps continuous functions to continuous functions. Using the Fourier operator, we can express the inner products \( \langle f_{\omega}, \phi \rangle_{\mu N} \) and their variance as

\[
\langle f_{\omega}, \phi \rangle_{\mu N} = F_{\omega,N} \phi(x_0), \quad \Var_{\mu}(f_{\omega}, \phi)_{\mu N} = \| F_{\omega,N} \phi \|^2_{\mu}.
\]

\[ 14 \]
Thus, it is sufficient to show that
\[
\lim_{N \to \infty} \sup_{\omega \in \mathbb{R}} \| \mathcal{F}_\omega \phi \|_\mu^2 = 0, \quad \forall \phi \in L^2(\mu).
\] (23)

For this purpose, the spectral representation of unitary operators will be used. Let \( H_\phi \) be the cyclic subspace generated by \( \phi \), i.e., the closure of \( \text{span}\{U^n \Delta_t \phi : n \in \mathbb{Z}\} \). By Herglotz’s theorem, there exists a Borel probability measure \( \mu_\phi \) on \( S^1 \) such that, for every \( n \in \mathbb{Z} \),
\[
c_n := \langle U^n \phi, \phi \rangle_\mu = \int_{S^1} z^{-n} d\mu_\phi.
\]
As a result, the linear map \( \Psi_\phi : H_\phi \to L^2(S^1, \mu_\phi) \), defined uniquely through the requirement that \( \Psi_\phi : U^n \phi \mapsto z^n \), for all \( n \in \mathbb{Z} \), is an isometry. Moreover,
\[
\Psi_\phi (\mathcal{F}_\omega \phi)(e^{i\theta}) = \frac{1}{N} \sum_{n=0}^{N-1} e^{-i\omega n} \Psi_\phi (U^n \phi)(e^{i\theta}) = \frac{1}{N} \sum_{n=0}^{N-1} e^{i(\theta-n) \omega} = \frac{1}{N} e^{i(\theta-\omega)N} - 1.
\]
The term in the right-hand side of the last inequality can be succinctly expressed in terms of a function
\[
S_N(\omega) := \left| \frac{\sin(Nu/2)}{N \sin(u/2)} \right| = \left| \frac{1}{N} e^{i\omega} - 1 \right|.
\]
Using this function, we obtain
\[
\| \mathcal{F}_\omega \phi \|_\mu^2 = \int_0^{2\pi} |\Psi_\phi (\mathcal{F}_\omega \phi)(e^{i\theta})|^2 d\mu_\phi(\theta) = \int_0^{2\pi} S_N^2(\theta - \omega) d\mu_\phi(\theta).
\] (24)

Now, by assumption, the measure \( \mu_\phi \) is absolutely continuous with respect to the Lebesgue measure \( \mu_{Leb} \) on \( S^1 \), and hence it has a density \( \rho_\phi = d\mu_\phi / d\mu_{Leb} \in L^1(S^1, \mu_{Leb}) \). Therefore, (23) will be proved if it can be shown that
\[
\lim_{N \to \infty} \sup_{\omega \in \mathbb{R}} \int_0^{2\pi} S_N^2(\theta - \omega) \rho_\phi(\theta) d\theta = 0.
\] (25)

Given \( \epsilon > 0 \), it will be shown that for \( N \) sufficiently large, \( \sup_{\omega \in \mathbb{R}} \| \mathcal{F}_\omega \phi \|_\mu^2 < 2\epsilon \). Since \( \rho_\phi \) is integrable with respect to the Lebesgue measure, by the continuity of integrals, there exists \( \delta > 0 \) such that for any Borel set \( E \subset \mathbb{R} \) with \( \mu_{Leb}(E) < \delta \), \( \int_E \rho d\mu_{Leb} < \epsilon \). For \( \omega \in \mathbb{R} \), let \( I_\omega \subset S^1 \) be an interval of length \( \delta \) centered around \( \omega \). By (24),
\[
\| \mathcal{F}_\omega \phi \|_\mu^2 = \int_{S^1} S_N^2(\theta - \omega) \rho_\phi(\theta) d\theta = \int_{I_\omega} S_N^2(\theta - \omega) \rho_\phi(\theta) d\theta + \int_{S^1 \setminus I_\omega} S_N^2(\theta - \omega) \rho_\phi(\theta) d\theta.
\]
The function \( S_N \) satisfies \( 0 \leq S_N(u) \leq 1 \) for every \( u \in \mathbb{R} \), and therefore by our choice of \( \delta \), the first integral is less than \( \epsilon \). To bound the second integral, note that for any \( \delta > 0 \) and as \( N \to \infty \),
\[
\sup_{u \in \mathbb{R} : |u| > \delta} S_N(u) = O(1/N),
\]
which implies that the second integral is \( O(N^{-2}) \), and thus less than \( \epsilon \) for sufficiently large \( N \). This proves (25), which proves (22), and thus Theorem 4. \( \square \)
8. Numerical computation of RKHS norms

When numerically implementing the results of Theorems 1 and 3 to find Koopman eigenfrequencies and eigenfunctions, one has to deal with two limitations, namely: (i) instead of obtaining $L^2(\mu)$ functions defined over the entire space $X$, one can only obtain functions defined on the finite orbit $X_N$; (ii) one can practically only compute a finite subset of the eigenfrequencies. Regarding (i), note that the Nyström extension $h_N = T_N(f_0 | X_N)$ is a continuous function on $X$, and can be calculated even at points $x$ lying outside the orbit. By Theorem 1 if $\omega \in \Omega$, then $h_N$ converges in $H$ to a Koopman eigenfunction, so even though $N$ is finite, if it is large enough, then $h_N$ is a good approximation of that eigenfunction. To address (ii), we rely on finding a smaller set of frequencies, called a generating set, as described below.

Generating frequencies. First, note that by definition of the Koopman operator, the product of any two Koopman eigenfunctions $z_{\omega_1}, z_{\omega_2} \in L^\infty(X, \mu)$ corresponding to eigenfrequencies $\omega_1$ and $\omega_2$, respectively, is also an $L^\infty$ eigenfunction corresponding to the eigenfrequency $\omega_1 + \omega_2$. Next, suppose that there exist non-constant Koopman eigenfunctions, and all Koopman eigenfunctions are continuous. Then, by ergodicity (which, as stated in Section 1 is equivalent to the eigenvalues of $V$ being simple), every Koopman eigenfunction with unit $L^2(\mu)$ norm takes values on the unit circle in the complex plane (notice the similarity with the Fourier functions $f_\omega$ in that regard), and for every collection $z_1, \ldots, z_m$ of eigenfunctions corresponding to rationally independent eigenfrequencies, $x \in X \mapsto (z_1(x), \ldots, z_m(x)) \in \mathbb{T}^m$ is a surjective map taking values on the $m$-torus. There exists a minimal $d \in \mathbb{N}$, such that every eigenfrequency can be expressed as a linear combination of $d$ rationally independent eigenfrequencies, $\omega_1, \ldots, \omega_d$, with integer coefficients. In other words, the set of Koopman eigenfrequencies $\{\omega = \sum_{j=1}^d c_j \omega_j, \, c_j \in \mathbb{Z}\}$ forms a $d$-dimensional vector space over the integers, having $\{\omega_1, \ldots, \omega_d\}$ as a basis. Moreover, if $z_1, \ldots, z_d$ are Koopman eigenfunctions of unit $L^2(\mu)$ norm, corresponding to $\omega_1, \ldots, \omega_d$, respectively, then every unit-norm Koopman eigenfunction $z$ is of the form $\sum_{j=1}^d c_j z_j$, $c_j \in \mathbb{Z}$. The set of all such eigenfunctions forms an orthonormal basis of $\mathcal{D}$, which also has the structure of an Abelian group under multiplication of continuous functions. Overall, it is enough to identify $d$ generating eigenpairs $(\omega_j, z_{\omega_j})_{j=1}^d$ to recover all eigenfrequencies and their corresponding eigenfunctions.

Kernels on data space. In data-driven applications, the sampled dynamical states $x_0, \ldots, x_{N-1}$ in $X$ are usually unknown, and one only has access to the values $F(x_0), \ldots, F(x_{N-1})$ of an observation map $F : X \mapsto Y$ taking values in a data space $Y$. This places a restriction on the type of kernel $k : X \times X \mapsto \mathbb{C}$ used in numerical implementations, as it must be computable from the values of $F$ alone. That is, $k$ must have the structure of a “pullback kernel”, $k(x, y) := \kappa(F(x), F(y))$, where $\kappa : Y \times Y \mapsto \mathbb{C}$ is a kernel on $Y$ designed according to the requirements of the application at hand. Here, we require that Assumption 3 be satisfied, which implies that $F$ be injective and continuous (i.e., Assumption 2 is satisfied), and the restriction of $\kappa$ on $F(X) \times F(X) \subseteq Y \times Y$ be continuous and strictly positive.

As a guideline for choosing $\kappa$ so as to satisfy the strict positivity condition, we note that the reproducing kernel $k : X \times X \mapsto \mathbb{C}$ of an RKHS on $X$ is strictly positive if and only if the kernel sections $k(x_1, \cdot), \ldots, k(x_n, \cdot)$ are linearly independent for all $x_1, \ldots, x_n$ in $X$. When one does not have a priori knowledge of the image $F(X)$ of state space in data space, it is generally preferable to define $k$ through a $\kappa$ which is positive-definite on the whole of $Y$. For example, in the case $Y = \mathbb{R}^m$, the radial Gaussian kernels,

$$\kappa(y_1, y_2) = \exp\left(-\frac{d^2(y_1, y_2)}{\epsilon}\right),$$

(26)

where $d : Y \times Y \mapsto \mathbb{R}$ is the Euclidean metric and $\epsilon$ a positive bandwidth parameter, are positive definite. On the other hand, the covariance kernel,

$$\kappa(y_1, y_2) = \langle y_1, y_2 \rangle_Y = [d^2(y_1, y_2) - d^2(y_1, -y_2)] / 4,$$

(27)

does not lead to a positive-definite kernel $k$ on $X$, as in this case $k(x, \cdot)$ depends linearly on $F(x)$. We will return to a discussion of the behavior of our methods implemented with covariance kernels and their relationship to DFT approaches in Section 9. There, we will also discuss methods for constructing positive-definite
kernels from non-injective observation maps through the use of delay-coordinate embeddings. Additional examples of commonly used kernels in machine learning and signal processing can be found in [24].

Kernels from delay-coordinate mapped data. The condition that the observation map \( F : X \rightarrow Y \) is injective may not be satisfied in a number of real-world applications, for example when \( F \) is a low-dimensional observation map that only provides partial information about the underlying dynamical states. In such cases, it is possible to utilize the dynamical flow to construct an empirically accessible observation map \( F_Q : X \rightarrow Y^Q, Q \in \mathbb{N} \), using the data obtained from Theorem 3. Equivalently, the quantity \( \omega / \bar{\omega} \)

By Theorem 1, if \( m \) is chosen large enough, then the RKHS subspace spanned by \( \psi_0, \ldots, \psi_{l_0-1} \) approximates well the first \( m \) Koopman eigenfunctions for some \( m \geq 1 \). If \( m \) is large enough, then the subspace would contain a generating set of Koopman eigenfunctions. By Theorem 1 if \( \omega \notin \Omega \), for large \( N \), \( w_{N,l_0}(f_\omega) \) should be smaller than a threshold \( \delta > 0 \) (Step 2). The parameter \( l_1 \) needs to be large enough so that for \( \omega \notin \Omega \), \( w_{N,l_1}(f_\omega) \) is large for large \( l_1 \), as established in Theorem 3. Equivalently, the quantity

\[
 r(\omega; l_0, l_1, N) := \left| \frac{w_{N,l_1}(f_\omega)}{w_{N,l_0}(f_\omega)} - 1 \right| 
\]

should also be large, and we empirically require (in Step 3) it to be larger than the same threshold \( \delta \) as above. Finally, note that data-driven approaches can only recover the eigenfrequencies of the discrete time map \( \Phi^{\Delta t} \), rather than the actual flow \( \Phi \). Therefore, it is sufficient to look for frequencies only in the interval \( [-\pi/\Delta t, \pi/\Delta t] \), in accordance with the Nyquist sampling theorem. Algorithm 1 is meant to be used in conjunction with Proposition 11 stated below.

Algorithm 1. The algorithm assumes that there is an underlying ergodic flow \((X, \Phi^t)\) satisfying Assumption 1 and the kernel \( k : X \times X \rightarrow \mathbb{C} \) satisfies Assumption 3.

- **Input:** sampling interval \( \Delta t \); the values of a kernel \( k \) on a trajectory \( \{x_0, \ldots, x_{N-1}\} \) of length \( N \); a threshold \( \delta > 0 \); integers \( l_0, l_1 \) such that \( 0 < l_0 < l_1 \leq N \).
- **Output:** A collection of approximate Koopman eigenpairs \( (\omega_1, f_{\omega_1}), \ldots, (\omega_d, f_{\omega_d}) \) with \( \omega_j \in \mathbb{R} \) and \( f_{\omega_j} \in \mathcal{H} \).
- **Steps**
  1. Choose a subset \( \Omega_N \) of frequencies contained in \( [-\pi/\Delta t, \pi/\Delta t] \). See Proposition 11 on how to choose this set. Calculate the quantities \( w_{N, l}(f_\omega) \) for \( \omega \in \Omega_N \) and \( l \in \{l_0, l_1\} \) using the DFT-based method described in Proposition 11.
  2. Discard the \( \omega \) in \( \Omega_N \) for which the quantity \( r(\omega; l_0, l_1, N) \) in (28) is greater than \( \delta \).
  3. Of the remaining \( \omega \) in \( \Omega_N \), select the ones for which \( w_{N, l_0}(f_\omega) > \delta \).
  4. Collect the non-discarded frequencies \( \omega_1, \ldots, \omega_d \) from Step 3, and compute the Nyström extensions \( f_{\omega_j} = T_N(f_{\omega_j}|X_N) \) of the corresponding Fourier functions via (16).
Scanning for the eigenvalues. Although the goal is to scan for Koopman eigenfrequences in the interval \([-\pi/N\Delta t, \pi/N\Delta t]\), in any practical situation, \(w_{N,l}(f_{\omega_r})\) can only be computed for some finite collection of \(\omega_r\).

The rest of this section describes how to choose this collection so that it both gets denser in the interval \([0, 2\pi/N\Delta t]\) as \(N \to \infty\), and also allows fast efficient computation through the use of fast Fourier transform (FFTs). For a fixed \(N\), assumed odd for simplicity, define

\[
\omega_r := 2\pi r/N\Delta t, \quad r \in \{-{(N-1)/2}, \ldots, (N-1)/2\}.
\] (29)

Moreover, let \(Z_N : \mathbb{C}^N \to \mathbb{C}^N\) be the discrete Fourier transform with \(\vec{b} = Z_N\vec{a}, \vec{a} = (a_0, \ldots, a_{N-1})\), \(\vec{b} = (-b_{-(N-1)/2}, \ldots, b_{(N-1)/2})\), and

\[
b_j := \frac{1}{N} \sum_{n=0}^{N-1} e^{-2\pi inj/N} a_n.
\]

The following proposition shows how one can utilize the speed of the FFT to compute \(w_{N,l}(f_{\omega_r})\) for the frequencies \(\omega_r\) in (29).

**Proposition 11.** Let \(\Phi_N\) be the \(N \times N\) matrix whose \((j, n)\)-th element is \(\lambda_{N,n}^{-1/2} \phi_{N,n}(x_j)\). Let also \(A_N := Z_N\Phi_N\), where \(Z_N\) operates on each column of \(\Phi_N\). Then, \(w_N(f_{\omega_r})\) is equal to the squared \(l^2\) norm of the \(r\)-th row of \(A_N\), indexed such that \(r \in \{-{(N-1)/2}, \ldots, (N-1)/2\}\). Moreover, \(w_{N,l}(f_{\omega_r})\) is equal to the squared \(l^2\) norm of the \(r\)-th row of \(A_N\), truncated to the first \(l\) entries.

**Proof.** It follows from (18) that \(w_{N,l}(f_{\omega_r}) = \sum_{j=0}^{l-1} \left| a_{r,N,n}^{\ast} \right|^2\), where \(a_{r,N,n} = \langle \phi_{N,n}, f_{\omega_r}X_N \rangle_{\mu_N}\), and \(l \in \{0, \ldots, N-1\}\). The claim of the Proposition for \(w_N(f_{\omega_r})\) follows from the fact that \(\left(\lambda_{N,n}^{-1/2} a_{r,N,n}^{\ast}\right)_{r=-N/2}^{N/2}\) is the DFT of the sequence \(\left(\lambda_{N,n}^{-1/2} \phi_{N,n}(x_j)\right)_{j=0}^{N-1}\), which forms the \(n\)-th column of \(\Phi_N\). The claim for \(w_{N,l}(f_{\omega_r})\) follows in an analogous manner. \(\square\)

9. Examples and discussion

In this section, we apply the methods described in Sections 8 to ergodic dynamical systems with different types of spectra. The goal is to demonstrate that the results of Theorems 1 and 3 hold, and the procedure described by Algorithm 1 is effective in identifying Koopman eigenfunctions and eigenfrequencies.

We consider the following three systems, whose spectra are respectively pure point, absolutely continuous, and mixed, respectively:

1. A linear quasiperiodic flow \(R_{\alpha_1,\alpha_2}\) on \(\mathbb{T}^2\), defined as

\[
dR_{\alpha_1,\alpha_2}'(\theta)/dt = (\alpha_1, \alpha_2), \quad \theta = (\theta_1, \theta_2) \in \mathbb{T}^2, \quad \alpha_1 = 1, \quad \alpha_2 = \sqrt{2},
\] (30)

and observed through the observation map \(F: \mathbb{T}^2 \to \mathbb{R}^1\) with

\[
F(\theta_1, \theta_2) = \sin(\theta_1) \cos(\theta_2).
\] (31)

This system has a pure point Koopman spectrum, consisting of eigenfrequencies of the form \(n_1 \alpha_1 + n_2 \alpha_2\) with \(n_1, n_2 \in \mathbb{Z}\). Because \(\alpha_1\) and \(\alpha_2\) are rationally independent, the set of eigenfrequencies lies dense in \(\mathbb{R}\), which makes the problem of numerically distinguishing eigenfrequencies from non-eigenfrequencies non-trivial despite the simplicity of the underlying dynamics.

2. The Lorenz 63 (L63) flow \(\Phi_{l63}^t : \mathbb{R}^3 \to \mathbb{R}^3\), generated by the \(C^\infty\) vector field \(\vec{V}\) with components \((V^x, V^y, V^z)\) at \((x, y, z) \in \mathbb{R}^3\) given by

\[
V^x(x^y - x), \quad V^y(x^y - z) - y, \quad V^z = x^y - \beta z,
\] (32)
where $\beta = 8/3$, $\rho = 28$, and $\sigma = 10$. The system is sampled through the observation map $F : \mathbb{R}^3 \to \mathbb{R}^3$ with
\[
F(x, y, z) := (x, y, z).
\] (33)

The L63 flow is known to have a chaotic attractor $X_{l63} \subset \mathbb{R}^3$ with fractal dimension $\approx 2.06$ [20], supporting a physical invariant measure [27] with a corresponding purely continuous spectrum of the generator [28]. That is, there exist no nonzero Koopman eigenfrequencies for this system.

3. A Cartesian product $\Phi_{l63} \times R^\alpha$ of the L63 flow with a linear flow $R^\alpha$ on $S^1$ defined by
\[
dR^\alpha_{\theta}(\theta)/dt = \alpha, \quad \theta \in S^1, \quad \alpha = 1.
\] (34)

This system is observed through a map $F : \mathbb{R}^3 \times S^1 \to \mathbb{R}^3$, which combines the coordinates of the continuous-spectrum subsystem with the rotation, viz.
\[
F(x, \theta) = x + c \left( \sin(\theta), \cos(2\theta), \sin(2\theta) \right), \quad x \in \mathbb{R}^3, \quad \theta \in S^1, \quad c = 0.2.
\] (35)

This system has a mixed spectrum system, containing the discrete spectrum $\{ j\omega : j \in \mathbb{Z} \}$ of the rotation and the continuous spectrum of the L63 flow as subsets. Due to the smallness of the constant $c$ in (35), the L63 signal dominates the contribution from the rotation in the observation map $F$.

Methodology. The following steps describe sequentially the entire numerical procedure carried out.

1. Numerical trajectories $x_0, x_1, \ldots, x_{N-1}$, with $x_n = \Phi^{n\Delta t}(x_0)$ of length $N$ are generated, using a sampling interval $\Delta t = 0.01$ in all cases. In the L63 experiments, we let the system relax toward the attractor, and set $x_0$ to a state sampled after a long spinup time (4000 time units); that is, we formally assume that $x_0$ has converged to the ergodic attractor. We use the ode45 solver of Matlab to compute the trajectories. For the three systems $R^\alpha_{\omega_1, \alpha_2}, \Phi^\omega_{l63}$, and $\Phi^\omega_{l63} \times R^\alpha$, the sample numbers are $N = 40,000, 60,000$, and $70,000$, and the initial points in state space are $x_0 = (0,0)$, $(0,1,1.05)$, and $(0,1,1.05,0)$, respectively.

2. The observation map $F$ described for each system is used to generate the respective time series $F(x_0), F(x_1), \ldots, F(x_{N-1})$. This dataset forms the basis of all subsequent computations. We include delays in the observation map to obtain an injective observation map, they were 5, 2 and 10 respectively for the three systems.

3. The kernel $k$ is obtained by starting with the Gaussian kernel from [26], and then performing two normalizations so as to make it a Markov kernel as in the diffusion maps algorithm [29]. The kernel bandwidth $\epsilon$ is determined automatically via the procedure described in [30, 31, 32], which yields $\epsilon = 0.0459, 0.015,$ and 0.0804, for the $R^\alpha_{\omega_1, \omega_2}, \Phi^\omega_{l63}$, and $\Phi^\omega_{l63} \times R^\alpha$, systems, respectively.

4. The eigenpairs $(\phi_{N,j}, \lambda_{N,j})$ are computed for $j \in \{0, \ldots, l_1\}$ using Matlab’s eigs iterative solver.

5. The quantity $w_{N,l}(f_\omega)$ is computed for $l = l_0, l_1$ and the frequencies $\omega$, from [29]. The values of $(l_0, l_1)$ for the $R^\alpha_{\omega_1, \omega_2}, \Phi^\omega_{l63}$, and $\Phi^\omega_{l63} \times R^\alpha$, systems are $(100, 1000), (100, 1000)$, and $(1240, 1500)$, respectively. Algorithm [3] is then applied with these parameters to determine Koopman eigenfrequencies. The values of $\delta$ are 1.0, 1.0, and 0.1 respectively.

Results. The results shown in Figs. [4] verify the statements of Theorems [1] and [3] in particular, recall that $w_N$ is the squared RKHS norm of $T_N(f_\omega | X_N)$, while $w_{N,l}$ is the squared RKHS norm of $T_N(f_\omega | X_N)$ projected to the subspace spanned by $\psi_{N,0}, \ldots, \psi_{N,l-1}$. Hence, $l$ behaves like a resolution parameter. According to Theorem [3] at fixed finite resolution $l = l_0$, $w_{N,l_0}(f_\omega)$ converges to 0 if $\omega \notin \Omega$, whereas Theorem [3] states that at variable resolution $l = N$, $w_N(f_\omega) = w_{N,N}(f_\omega)$ diverges to $\infty$. Since in practice we are not at a liberty to increase $N$ to test for the asymptotic behavior of $w_{N,l_0}(f_\omega)$ and $w_{N}(f_\omega)$, we take advantage of the different nature of these criteria to identify Koopman eigenfrequencies via a two-step approach (Steps 2 and 3 in Algorithm [3]).

As can be seen in Figs. [2] and [4], this procedure leads to accurate identification of Koopman eigenfrequencies in all three systems described above. In the case of the quasiperiodic rotation on $T^2$ (Fig. [2]) and the mixed-spectrum system on $\mathbb{R}^3 \times S^1$ (Fig. [4]), the method identifies a sufficient number of eigenfrequencies.
Figure 2: Results of Algorithm 1 applied to the linear quasiperiodic flow (30) on the 2-torus, using $N = 40,000$ samples. (a) Observable time series used for the analysis, obtained from the map $F$ in (31). (b) Spectrally truncated squared RKHS norm $w_{N,l_0}(f_{\omega})$ from (18) as a function of frequency $\omega_r \in \Omega_N$ for $l_0 = 100$. (c) Ratio $r(\omega_r; l_0, l_1, N)$ from (28) as a function of $\omega_r$, computed for $l_1 = 1000$ and $l_0$ as in (b). (d) Dependence of $w_{N,l}(f_{\omega})$ on $l$ and $\omega_r$ for the Koopman eigenfrequencies estimated by discarding all frequencies in $\Omega_N$ for which $w_{N,l_0}(f_{\omega}) < \delta$ and $r(\omega_r; l_0, l_1, N) > \delta$, for a threshold $\delta = 1.0$ and $(l_0, l_1)$ as in (c). The selected frequencies include 2 and $1 + \sqrt{2}$, which are integer linear combinations of the two basic frequencies of 1 and $\sqrt{2}$. The plot shows that for these frequencies $w_{N,l}(f_{\omega})$ does not change significantly as $l$ is changed from $l_0$ to $l_1$. 

20
Figure 3: As in Fig. 2, but for the L63 system in (32) observed through the observation map in (33). The number of samples is $N = 60,000$, while the parameters of Algorithm 1 are set to $(l_0, l_1) = (100, 1000)$ and $\delta = 1$. As shown in (d), the algorithm identifies $\omega_r = 0$ as the only Koopman eigenfrequency (corresponding to a constant eigenfunction), consistent with the fact that L63 system has a purely continuous spectrum.
Figure 4: As in Fig. 2, but for the mixed-spectrum system on $\mathbb{R}^3 \times S^1$ from (34), observed via the map $F$ from (35). The number of samples is $N = 70,000$, while Algorithm 1 is executed using the parameters $(l_0, l_1) = (1240, 1500)$ and $\delta = 0.1$. Panels (a), (b) are reproduced from Fig. 4 for convenience. The Koopman eigenfrequencies identified by the algorithm in (d) are $\omega_r = 0$, 1, and 2, consistent with the discrete spectrum of the rotation $R^\alpha_t$ with $\alpha = 1$. 

\[\text{Figure 4: As in Fig. 2, but for the mixed-spectrum system on } \mathbb{R}^3 \times S^1 \text{ from (34), observed via the map } F \text{ from (35). The number of samples is } N = 70,000, \text{ while Algorithm 1 is executed using the parameters } (l_0, l_1) = (1240, 1500) \text{ and } \delta = 0.1. \text{ Panels (a), (b) are reproduced from Fig. 4 for convenience. The Koopman eigenfrequencies identified by the algorithm in (d) are } \omega_r = 0, 1, \text{ and } 2, \text{ consistent with the discrete spectrum of the rotation } R^\alpha_t \text{ with } \alpha = 1.\]
to generate the full set of eigenfrequencies via linear combinations. Specifically, in Fig. 2(d) we find the frequencies $0.419, 2.0, 2.419, 2.827, 5.246, \text{and} 5.655$, which agree with the theoretically expected eigenfrequencies $\sqrt{2} - 1, 2, 1 + \sqrt{2}, 2\sqrt{2}, 3\sqrt{2} - 1, 3\sqrt{2} + 1, \text{and} 4\sqrt{2}$ respectively, of the $T^2$ rotation to within two significant digits. In Fig. 3(d), we recover the two eigenfrequencies of the circle rotation, 1 and 2, present in the observation map. In the case of the L63 flow (Fig. 3), the method only identifies the trivial (zero) eigenfrequency, consistent with the fact that the L63 flow is mixing and its associated Koopman group does have nonconstant eigenfunctions.

Note that the frequencies identified by Algorithm 1 are fairly insensitive to the input parameters, $l_0, l_1, \text{and} \delta$, and the two-step approach for selecting eigenfrequencies contributes at least partly to that robustness. That is, in the case of the quasiperiodic rotation, Step 2 does not reject any candidate frequencies in $\Omega_\delta$ (Fig. 2(b)), which suggests that the $\delta$ satisfies the requirement that the $\delta N$ frequencies which have nonconstant eigenfunctions.

In the case of the mixed-spectrum system, Step 2 with $\delta N$ does not lead to rejection of any candidate frequency (Fig. 4(c)).

Covariance kernels. One of the requirements for Theorems 1–4 to hold is that the kernel is strictly positive definite. As stated in Section 8, this requirement is not satisfied when using a covariance kernel $k(x, x') = \langle F(x), F(x') \rangle Y$ from (27) associated with an observation map $F: X \rightarrow Y$ taking values in $\mathbb{R}^m$. In particular, the rank of the kernel integral operators $S_N: L^2(\mu_N) \rightarrow H$ associated with such a kernel (and also the rank of $S: L^2(\mu) \rightarrow H$) is never more than $m$, meaning that $f_0|X_N \in L^2(\mu_N)$ may fail to have a Nyström extension in $H$ (for the domain $D(T_N)$ of the extension operator $T_N$ will be a strict subspace of $L^2(\mu_N)$ for $N > \text{rank } S_N$). In effect, a covariance kernel on a finite-dimensional data space significantly limits the richness of observables in the corresponding RKHS, thus decreasing the likelihood that Koopman eigenfunctions can be found in this space.

To examine the relationship between our approach and conventional harmonic averaging techniques, observe that even if $f_0|X_N$ does not lie in $D(T_N)$ (taking $H$ to be the RKHS associated with a covariance kernel as above), it is still possible to compute the $H$ extension of the orthogonal projection $g_{\omega,N} \in D(T_N)$ of $f_0|X_N$ onto the domain of the Nyström extension operator $T_N$, and evaluate the squared RKHS norm

$$\|T_N g_{\omega,N}\|^2 = w_N l_N(f_\omega) = \sum_{j=0}^{l_N-1} \|\phi_{j,N}|X_N\|_{\mu_N}^2 / \lambda_{j,N},$$

where $l_N = \text{rank } S_N \leq m$. Note also that the kernel integral operator $G_N = S_N^* S_N: L^2(\mu_N) \rightarrow L^2(\mu_N)$ associated with the covariance kernel takes the form

$$G_N = A_N^* A_N,$$

where $A_N: L^2(\mu_N) \rightarrow \mathbb{R}^m$ is the rank-$l_N$ operator acting on $f \in L^2(\mu_N)$ by component-wise integration against the observation map, viz.

$$A_N f = \int_X F(x) f(x) d\mu_N(x).$$

It then follows from (37) that the $L^2(\mu_N)$ basis vectors $\phi_{0,N}, \ldots, \phi_{l_N-1,N}$ are also right singular vectors of $A_N$ corresponding to the (strictly positive) singular values $\lambda_{0,N}, \ldots, \lambda_{l_N-1,N}^{1/2}$, respectively. This property, in conjunction with (38), leads to

$$F(x_n) = \sum_{j=0}^{l_N-1} c_j \lambda_{j,N}^{1/2} \phi_{j,N}(x_n), \quad \forall x_n \in X_N,$$
where \( \epsilon_{0,N}, \ldots, \epsilon_{N-1,N} \) are orthonormal left singular vectors of \( A_N \) in \( \mathbb{R}^m \). Inserting this representation of \( F(x_n) \) in the harmonic averaging formula in (1), we obtain

\[
\| \mathcal{F}_{\omega,N} F \|_{\mathbb{R}^m}^2 = \sum_{j=0}^{l_N-1} | \langle \phi_{j,N}, f_\omega | X_N \rangle |^2 \frac{\lambda_j}{\lambda_j}.
\]

A comparison of (36) and (39) then shows that the power spectral density \( \| \mathcal{F}_{\omega,N} F \|_{\mathbb{R}^m}^2 \) from harmonic averaging has a structurally similar representation to the squared RKHS norm \( \| T_{N,N} g_{\omega,N} \|_{H}^2 \) in terms of the \( (\lambda_j, \phi_j) \) eigenpairs, apart from the fact that the former involves multiplication by \( \lambda_j \) (thus being dominated by the projections of \( f_\omega | X_N \) along the most energetic signal components), whereas the latter involves division by \( \lambda_j \) (thus being dominated by the projections of \( f_\omega | X_N \) along the most irregular signal components in the sense of the covariance kernel). In applications where the ratio \( \lambda_0/N/\lambda_{1,N-1,N} \) is not too large, \( \| \mathcal{F}_{\omega,N}(F) \|_{\mathbb{R}^m}^2 \) and \( w_{N-1,N}(f_\omega) \) will thus be comparable. If, however, \( \lambda_0/N/\lambda_{1,N-1,N} \) is large (as will typically be the case in high data space dimensions), then, depending on the frequency \( \omega \), the two quantities can be vastly different. As illustrated in Fig. 1(d), the limitation \( l_N \leq m \) may be inadequate for estimating eigenfrequencies from observables of mixed-spectrum systems dominated by the continuous spectrum.

Comparison with harmonic averaging. When interpreting the relationship between harmonic averaging and the RKHS-based approach proposed here, it should be kept in mind that for a fixed finite spectral truncation \( l \leq m \), Theorem 3 shows convergence results similar to harmonic averaging (that is, if either of \( \| \mathcal{F}_{\omega,N} \|_{\mathbb{R}^m}^2 \) or \( w_{N-1,N}(F) \) do not converge to zero as \( N \to \infty \), then \( \omega \) is an eigenfrequency), while taking \( l = N \), Theorem 4 shows convergence properties of a fundamentally different nature. In effect, Theorem 4 states that if the projection of \( f_\omega | X_N \) onto the eigenspaces of \( G_N \) corresponding to \( H \) subspaces of low regularity decays rapidly enough as \( N \) increases, then \( \omega \) is an eigenfrequency. In order for this result to hold, the RKHS \( H \) must lie dense in \( L^2(\mu) \) (which implies that the rank of \( G_N \) increases without bound as \( N \to \infty \), and its smallest eigenvalue \( \lambda_N \) is strictly positive and converges to zero), and this will not occur with non-positive-definite kernels such as covariance kernels associated with data in \( \mathbb{R}^m \). Nevertheless, it is possible that for sufficiently large \( l_N = \text{rank} G_N \) and eigenvalue ratio \( \lambda_0/N/\lambda_{1,N-1,N} \) the squared RKHS norms \( \| T_{N,N} g_{\omega,N} \|_{H}^2 \) from (36) may approximate the behavior established in Theorem 1.

Summary. By working with positive-definite, universal kernels, our approach overcomes the limitation of the dimensionality of the observation map, while also providing two distinct criteria to identify Koopman eigenfrequencies. The truncated RKHS norm \( w_{N,l}(f) \) effectively computes harmonic averages of \( l \) observables \( \phi_{0,N}, \ldots, \phi_{l,N} \), with an important weighting by the inverse of the corresponding eigenvalues \( \lambda_j \), and as \( N \) grows this collection provides infinitely many observables to carry out harmonic averaging on. This results in estimates of the RKHS norms of candidate Koopman eigenfunctions, which is a measure of their regularity and an important criterion in identifying eigenfrequencies. In addition, the method computes approximations of \( \| \mathcal{F}_{\omega,N} F \|_{\mathbb{R}^m}^2 \) (as opposed to \( L^2 \)) Koopman eigenfunctions \( T_N(f_\omega | X_N) \in H \), which can be evaluated at arbitrary points in \( X \), and converge pointwise as \( N \to \infty \).

Acknowledgments. Dimitrios Giannakis received support from ONR YIP grant N00014-16-1-2649, NSF grant DMS-1521775, and DARPA grant HR0011-16-C-0116. Suddhasattwa Das is supported as a post-doctoral research fellow from the first grant.

References

[1] B. O. Koopman, Hamiltonian systems and transformation in Hilbert space, Proc. Natl. Acad. Sci. 17 (5) (1931) 315–318.
[2] M. Dellnitz, O. Junge, On the approximation of complicated dynamical behavior, SIAM J. Numer. Anal. 36 (1999) 491. doi:10.1137/S0036142996313002
[3] I. Mezić, A. Banaszuk, Comparison of systems with complex behavior, Phys. D. 197 (2004) 101–133. doi:10.1016/j.physd.2004.06.015
[4] I. Mezić, Spectral properties of dynamical systems, model reduction and decompositions, Nonlinear Dyn. 41 (2005) 309–325. doi:10.1007/s11071-005-2824-x
