Embedding classical dynamics in a quantum computer

Dimitrios Giannakis
Department of Mathematics, Dartmouth College, Hanover, New Hampshire 03755, USA;
Department of Physics and Astronomy, Dartmouth College, Hanover, New Hampshire 03755, USA;
and Department of Mathematics, Courant Institute of Mathematical Sciences, New York University, New York, New York 10012, USA

Abbas Ourmazd
Department of Physics, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin 53211, USA

Philipp Pfeffer
Institut für Thermo- und Fluiddynamik, Technische Universität Ilmenau, D-98684 Ilmenau, Germany

Jörg Schumacher
Institut für Thermo- und Fluiddynamik, Technische Universität Ilmenau, D-98684 Ilmenau, Germany
and Tandon School of Engineering, New York University, New York, New York 11201, USA

Joanna Slawinska
Department of Mathematics, Dartmouth College, Hanover, New Hampshire 03755, USA;
Center for Climate Physics, Institute for Basic Science (IBS), Busan, South Korea;
and Pusan National University, Busan, South Korea

(Received 14 June 2021; accepted 30 March 2022; published 3 May 2022)

We develop a framework for simulating measure-preserving, ergodic dynamical systems on a quantum computer. Our approach provides an operator-theoretic representation of classical dynamics by combining ergodic theory with quantum information science. The resulting quantum embedding of classical dynamics (QECD) enables efficient simulation of spaces of classical observables with exponentially large dimension using a quadratic number of quantum gates. The QECD framework is based on a quantum feature map that we introduce for representing classical states by density operators on a reproducing kernel Hilbert space, $\mathcal{H}$. Furthermore, an embedding of classical observables into self-adjoint operators on $\mathcal{H}$ is established, such that quantum mechanical expectation values are consistent with pointwise function evaluation. In this scheme, quantum states and observables evolve unitarily under the lifted action of Koopman evolution operators of the classical system. Moreover, by virtue of the reproducing property of $\mathcal{H}$, the quantum system is pointwise-consistent with the underlying classical dynamics. To achieve a quantum computational advantage, we project the state of the quantum system onto a finite-rank density operator on a 2^n-dimensional tensor product Hilbert space associated with n qubits. By employing discrete Fourier-Walsh transforms of spectral functions, the evolution operator of the finite-dimensional quantum system is factorized into tensor product form, enabling implementation through an n-channel quantum circuit of size $O(n)$ and no interchannel communication. Furthermore, the circuit features a state preparation stage, also of size $O(n)$, and a quantum Fourier transform stage of size $O(n^2)$, which makes predictions of observables possible by measurement in the standard computational basis. We prove theoretical convergence results for these predictions in the large-qubit limit, $n \to \infty$. In light of these properties, QECD provides a consistent simulator of the evolution of classical observables, realized through projective quantum measurement, which is able to simulate spaces of classical observables of dimension 2^n using circuits of size $O(n^2)$. We demonstrate the consistency of the scheme in prototypical dynamical systems involving periodic and quasiperiodic oscillators on tori. These examples include simulated quantum circuit experiments in Qiskit Aer, as well as actual experiments on the IBM Quantum System One.

DOI: 10.1103/PhysRevA.105.052404

I. INTRODUCTION

Ever since a seminal paper of Feynman in 1982 [1], the problem of identifying physical systems that can faithfully and efficiently simulate large classes of other systems (performing, in Feynman's words, universal computation) has received considerable attention. Under the operating principle that nature is fundamentally quantum mechanical, and with the realization that simulating quantum systems by classical systems is exponentially hard, much effort has been focused on the design of universal simulators of quantum systems. Such efforts are based on the axioms of quantum mechanics, with gates connected in quantum circuits

performing unitary (and thus reversible) transformations of quantum states [2–7].

Over the past decades, several numerically hard problems have been identified, for which quantum algorithms are significantly faster than their classical counterparts. A prominent example is the Grover search algorithm, which results in a quadratic speedup over classical search [8]. In a few cases, such as random sampling, quantum computers have solved problems that would be effectively unsolvable with present-day classical supercomputing resources, thus opening the way to quantum supremacy [9]. See also Ref. [10] for a discussion of the result in Ref. [9].

Yet, at least at the level of effective theories, a great variety of phenomena are well described by classical dynamical systems, generally formulated as systems of ordinary or partial differential equations. Since simulating a quantum system by a classical system can be exponentially hard, it is natural to ask whether simulation of a classical system by a quantum system is an exponentially "easy" problem, enabling a substantial increase in the complexity and range of computationally amenable classical phenomena.

The possibility to simulate classical dynamical systems on a quantum computer has attracted growing attention, on par with research on fundamental new quantum algorithms and their practical implementation [11]. Already 20 years ago, for example, Benenti *et al.* [12] studied the sawtooth map generating rich and complex dynamics. The implementation of an Euler method to solve systems of coupled nonlinear ordinary differential equations (ODEs) was addressed by Leyton and Osborne [13]. A framework for sequential data assimilation (filtering) of partially observed classical systems based on the Dirac–von Neumann formalism of quantum dynamics and measurement was proposed in Ref. [14]. The simulation of classical Hamiltonian systems using a Koopman–von Neumann approach was studied by Joseph [15]. This quantum computational framework was shown to be exponentially faster than a classical simulation when the Hamiltonian is represented by a sparse matrix. More recently, the potential of quantum computing for fluid dynamics, in particular turbulence, was explored in Refs. [16,17]. This includes, for example, transport simulators for fluid flows in which the formal analogy between the lattice Boltzmann method and Dirac equation is used [18]. Lubasch *et al.* [19] took a different path inspired by the success of quantum computing in solving optimization problems, modeling the one-dimensional Burgers equation by a variational quantum computing method, made possible by its correspondence with the nonlinear Schrödinger equation. Quantum systems have also been employed in the modeling of classical stochastic processes, where they have shown a superior memory compression [20,21].

Here, we present a procedure for simulating a classical, measure-preserving, ergodic dynamical system by means of a finite-dimensional quantum mechanical system amenable to quantum computation. Combining operator-theoretic techniques for classical dynamical systems with the theory of quantum dynamics and measurement, our framework leads to exponentially scalable quantum algorithms, enabling the simulation of classical systems with otherwise intractably high-dimensional spaces of observables. Our work thus opens a route to the full realization of quantum advantage in the computation of classical dynamical systems.

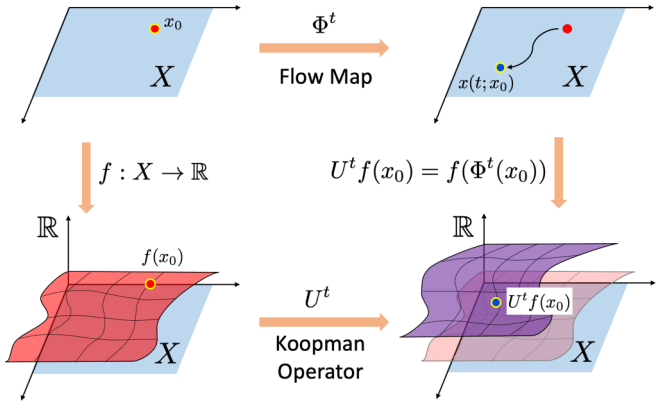

FIG. 1. Schematic of the relationship between the flow map Φ^t that advances the (nonlinear) dynamical system in a state space X in time and the linear Koopman operator U^t that advances observables f on X in an infinite-dimensional Hilbert space.

Another noteworthy aspect of our approach is that it interfaces between classical [22–29] and quantum [30–36] machine learning techniques based on kernel methods. Connections with data-driven, operator-theoretic techniques for classical dynamics [37–43] are also prevalent. Building on our previous work on quantum mechanical approaches to data assimilation [14], the framework presented here offers a mathematically rigorous route to representing complex, high-dimensional classical dynamics on a quantum computer. The primary contributions of this work are as follows.

(1) We present a generic pipeline that casts classical dynamical systems in terms amenable to quantum computation. This approach consists of four steps: (1) a dynamically consistent embedding of the classical state space X into the state space of an infinite-dimensional quantum system with a diagonalizable Hamiltonian; (2) eigenspace projection of the infinite-dimensional quantum system onto a finite-dimensional system, whose dynamics are representable by composition of basic commuting unitary transformations, realizable via quantum gates; (3) a preparation process, encoding the classical initial state in X to a quantum computational state; and (4) a quantum measurement process in the standard basis of the quantum computer to yield predictions for observables. These four steps result in simulations of a 2^n-dimensional space of classical observables using n qubits and a circuit of size (i.e., number of quantum gates) $O(n^2)$ and depth $O(n)$. We call this framework for encoding a classical dynamical system in terms of a quantum computational system quantum embedding of classical dynamics (QECD).

(2) We develop the principal mathematical tools employed in this construction using Koopman and transfer operator techniques [44,45] and the theory of reproducing kernel Hilbert spaces (RKHSs) [46,47] and Banach function algebras on locally compact Abelian groups [48–50]. The connection between the dynamical system and the Koopman operator is illustrated in Fig. 1. Using RKHSs as the foundation to build quantum mechanical models (as opposed to the L^2 spaces

employed in Ref. [25]) leads to pointwise consistency with the underlying classical dynamical system; that is, consistency for every classical initial condition, rather than in the sense of expectations over initial conditions. This result should be of independent interest in the broader context of representations of classical dynamics in terms of quantum systems, which has received significant attention [51–55].

(3) In the particular setting of quantum computation, we establish theoretical convergence results for the finite-dimensional systems generated by QECD, including asymptotic convergence rates in the large-qubit limit, $n \to \infty$. The time evolution of the quantum computational systems leverages discrete Fourier-Walsh techniques [56] to efficiently represent the Koopman operator using a circuit of size $O(n)$ and depth $O(1)$. The state preparation step, which is a major challenge in quantum computing [57,58], is also carried with a circuit of size $O(n)$ and depth $O(1)$. In particular, we take advantage of the fact that every quantum state associated with a classical initial state in X can be reached to any desired accuracy by efficient unitary transformations applied to a uniform-superposition state constructed using Hadamard gates. Meanwhile, the measurement process employs the quantum Fourier transform (QFT) to perform efficient approximate diagonalization of observables with a circuit of size $O(n^2)$ and depth $O(n)$ [59,60].

(4) We demonstrate the QECD framework in simple, analytically solvable examples of classical dynamics, so that all steps of the procedure are fully reproducible. Specifically, we use QECD to simulate the evolution of observables of periodic and quasiperiodic dynamical systems in a one- and two-dimensional phase space, respectively. We employ the gate-based, universal quantum computing toolkit Qiskit Aer [61], using up to $n = 8$ qubits. Results from simulated quantum circuit experiments (see Figs. 6 and 8) are found to be in good agreement with the true classical dynamics. In addition, we perform experiments for the periodic system on an actual quantum computer, the IBM Quantum System One, demonstrating the ability of QECD to simulate a classical system on a noisy intermediate-scale quantum (NISQ) device.

We note that the two-dimensional quasiperiodic dynamics in our examples can be straightforwardly extended to higher dimensions, where the dynamics becomes increasingly indistinguishable from a chaotic system. For quasiperiodic dynamics, no interchannel communication is necessary. Circuits of higher complexity that create interqubit entanglement may need to be explored for treatment of chaotic dynamics.

The outline of the paper is as follows. First, in Sec. II we give a high-level description of the methodological framework underlying the quantum embedding. In Sec. III, we introduce the class of dynamical systems under study, along with the corresponding RKHSs of classical observables. This is followed in Secs. IV–VIII by a detailed description of the construction of the QECD for this class of systems. In Sec. IX, we discuss aspects of the computational complexity of our approach in relation to classical simulators of quantum computational systems [62–64]. In Secs. X and XI, we present our results from simulated and actual quantum computation experiments, respectively. Our primary conclusions are summarized in Sec. XII. The paper contains Appendixes on RKHS-based quantum mechanical representations of classical systems (Appendix A), Fourier-Walsh factorization of the Koopman generator (Appendix B), and QFT-based approximate diagonalization of observables (Appendix C). In addition, we provide an overview of elements of Koopman operator theory related to this work and associated numerical techniques in Appendixes D and E, respectively.

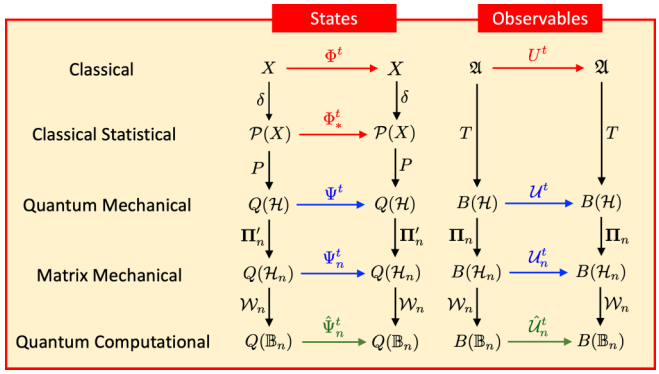

FIG. 2. Schematic representation of the QECD framework applied to states and observables of a classical dynamical system in five successive levels, leading to an n-qubit quantum computational system. These are the classical, classical statistical, quantum mechanical, matrix mechanical, and quantum computational levels. The horizontal arrows from top to bottom in the left- and right-hand columns represent the time evolution maps of states and observables, respectively. These are the flow map Φ^t on the classical state space X, the transfer operator Φ^t_* on the space of probability measures $\mathcal{P}(X)$, and the Koopman operator U^t on the algebra of classical observables $\mathfrak{A} \subseteq C(X)$. They are followed by the unitary evolution map Ψ^t and the Heisenberg operator $\mathcal{U}^t$ on the space of density operators $Q(\mathcal{H})$ and bounded linear operators $B(\mathcal{H})$, respectively, on the reproducing kernel Hilbert space $\mathcal{H}$. The maps at the matrix mechanical level, Ψ^t_n and $\mathcal{U}^t_n$, are finite-rank projections of Ψ^t and $\mathcal{U}^t$, respectively, acting on operators on 2^n-dimensional subspaces $\mathcal{H}_n$ of $\mathcal{H}$. The corresponding maps $\hat{\Psi}^t_n$ and $\hat{\mathcal{U}}^t_n$, respectively, at the quantum computational level act on operators on the 2^n-dimensional tensor product Hilbert space $\mathbb{B}_n$, which forms the basis of an n-qubit quantum computer. The vertical arrows correspond to maps that translate states (left-hand column) and observables (right-hand column) to the next representation level. Under the combined action of these maps, a classical state $x \in X$ is mapped to an n-qubit density matrix $\hat{\rho}_{x,n} \in Q(\mathbb{B}_n)$, and a classical observable $f \in \mathfrak{A}$ is mapped to a self-adjoint operator $\hat{S}_n \in B(\mathbb{B}_n)$. A loop of arrows represents a commutative diagram.

II. A ROUTE TO QUANTUM EMBEDDING OF CLASSICAL DYNAMICS

We begin by describing the main components of the QECD framework for representing classical dynamics on a quantum computer. Figure 2 schematically summarizes the successive levels used in the procedure, passing through classical, classical statistical, infinite-dimensional quantum mechanical, finite-dimensional quantum mechanical (referred to as matrix mechanical), and quantum computational levels. This diagram juxtaposes the steps for states and observables side by side for easy comparison. In the following subsections, we discuss the individual horizontal and vertical connections (which are maps) on each of the five levels of this diagram.

A. Classical and classical statistical levels

Consider a classical dynamical system on a compact metric space X, described by a dynamical flow map

$$\Phi^t : X \to X \quad \text{with} \quad t \in \mathbb{R}, \tag{1}$$

as indicated by a horizontal arrow in the left-hand column of Fig. 2. The classical state space X is embedded into the space of Borel probability measures $\mathcal{P}(X)$ (i.e., the classical statistical space) by means of the map δ sending $x \in X$ to the Dirac measure $\delta_x \in \mathcal{P}(X)$ supported at x. The dynamics acts naturally on the classical statistical space by the pushforward map on measures,

$$\Phi^t_* : \mathcal{P}(X) \to \mathcal{P}(X) \quad \text{with} \quad \Phi^t_*(\nu) = \nu \circ \Phi^{-t}, \tag{2}$$

also known as the transfer or Perron-Frobenius operator [44,45]. The map δ has the equivariance property $\Phi^t_* \circ \delta = \delta \circ \Phi^t$, represented by the top loop in the the left-hand column in Fig. 2.

Associated with the dynamical system are spaces of classical observables, which we take here to be spaces of complex-valued functions on X. A natural example is the space of continuous functions, denoted as $C(X)$, which also forms an (Abelian) algebra with respect to the pointwise product of functions. The Koopman operator [65,66], U^t, acts on observables in $C(X)$ by composition with the flow map, i.e.,

$$U^t : C(X) \to C(X) \quad \text{with} \quad U^t f = f \circ \Phi^t; \tag{3}$$

see also Fig. 1. The horizontal arrow in the first line of the right-hand column in Fig. 2 represents the action of the Koopman operator on a subalgebra $\mathfrak{A} \subseteq C(X)$ that will be described in Sec. II B below.

In this context, a simulator of the system can be described as a procedure which takes as an input an observable $f \in C(X)$ and an initial condition $x \in X$, and produces as an output a function $\hat{f}^{(t)}(x)$ approximating the evolution $f(\Phi^t(x))$ of the observable under the dynamics. For instance, if Φ^t is the flow generated by a system of ODEs $\dot{x} = \vec{V}(x)$ on $\mathcal{X} = \mathbb{R}^m$, and $X \subset \mathcal{X}$ is an invariant subset of this flow (e.g., an attractor), a standard simulation approach is to construct a finite-difference approximation $\hat{\Phi}^t : \mathcal{X} \to \mathcal{X}$ of the dynamical flow based on a timestep Δt (using interpolation to generate a continuous-time trajectory), and obtain $\hat{f}^{(t)}(x) = f(\hat{\Phi}^t(x))$ by evaluating the observable of interest f on the approximate trajectory. The scheme then converges in a limit of $\Delta t \to 0$ by standard results in ODE theory and numerical analysis for observables f of sufficient regularity.

From an observable-centric standpoint, a simulator of the system corresponds to a linear operator $\hat{U}^t$ approximating the Koopman operator U^t, giving $\hat{f}^{(t)}(x) = \hat{U}^t f(x)$. For instance, the ODE-based approximation just mentioned can be described in this way for $\hat{U}^t f = f \circ \hat{\Phi}^t$, but note that not every approximation of U^t has to be of the form of a composition operator by a flow. Indeed, "lifting" the task of simulation from states to (classical) observables opens the possibility of using new approximation techniques, which in some cases can resolve computational bottlenecks, e.g., due to high dimensionality (m) of the ambient state space $\mathcal{X}$ [67]. Invariably, every practical simulator $\hat{U}^t$ is restricted to act on a space of observables of finite dimension, N (e.g., a subspace of $C(X)$ or L^2). In general, the computation cost of acting with $\hat{U}^t$ on elements of this space scales as N^2, but can be reduced to $O(N)$ if $\hat{U}^t$ is efficiently represented by a diagonal matrix. The evaluation cost of observables, which corresponds to summation of an N-term basis expansion such as a Fourier series, is typically $O(N)$.

In what follows, rather than employing an approximation $\hat{U}^t$ acting on classical observables, our goal is to simulate the action of U^t using a quantum mechanical system. As we will see, this can be achieved at a logarithmic cost of elementary quantum operations (gates); specifically, QECD allows simulation of spaces of classical observables of dimension $N = 2^n$ using $O(n^2)$ gates.

B. Quantum computational representation

The QECD framework effecting the representation of the classical system by a quantum mechanical system employs the following key spaces:

(1) The classical state space X
(2) A Banach *-algebra $\mathfrak{A} \subseteq C(X)$ of classical observables
(3) An infinite-dimensional RKHS $\mathcal{H} \subset \mathfrak{A}$
(4) A finite-dimensional Hilbert space $\mathbb{B}_n$ associated with the quantum computer.

The Hilbert spaces $\mathcal{H}$ and $\mathbb{B}_n$ have corresponding (non-Abelian) algebras of bounded linear operators, $B(\mathcal{H})$ and $B(\mathbb{B}_n)$, respectively, acting as quantum mechanical observables. Moreover, states on these algebras are represented by density operators, i.e., trace-class, positive operators of unit trace, acting on the respective Hilbert space. We denote the spaces of density operators on $\mathcal{H}$ and $\mathbb{B}_n$ by $Q(\mathcal{H})$ and $Q(\mathbb{B}_n)$, respectively. Below, n represents the number of qubits, thus the dimension of $\mathbb{B}_n$ is 2^n.

The spaces of classical states and observables X and $\mathfrak{A}$ are mapped into the spaces of quantum states and observables $Q(\mathbb{B}_n)$ and $B(\mathbb{B}_n)$, respectively; see Fig. 2. The following maps on states (left-hand column) and observables (right-hand column) transform the classical system into a quantum-mechanical one on $\mathbb{B}_n$:

(1) We construct a map $\hat{\mathcal{F}}_n : X \to Q(\mathbb{B}_n)$ from classical states (points) in X to quantum states on $\mathbb{B}_n$. By analogy with the RKHS-valued feature maps in machine learning [68], $\hat{\mathcal{F}}_n$ will be referred to as a quantum feature map. To arrive at $\hat{\mathcal{F}}_n$, the classical statistical space $\mathcal{P}(X)$ is first embedded into the quantum mechanical state space $Q(\mathcal{H})$ associated with $\mathcal{H}$ through a map $P : \mathcal{P}(X) \to Q(\mathcal{H})$ [see (22) below]. The composite map $\mathcal{F} := P \circ \delta$ thus describes a one-to-one quantum feature map from X into $Q(\mathcal{H})$. Next, the infinite-dimensional space $Q(\mathcal{H})$ is projected onto a finite-dimensional quantum state space $Q(\mathcal{H}_n)$ associated with a 2^n-dimensional subspace $\mathcal{H}_n \subset \mathcal{H}$ by means of a map $\mathbf{\Pi}'_n : Q(\mathcal{H}) \to Q(\mathcal{H}_n)$. We refer to this level of description as matrix mechanical since all quantum states and observables are finite-rank operators, represented by $2^n \times 2^n$ matrices. To arrive at the quantum computational state space, we finally apply a unitary $\mathcal{W}_n : Q(\mathcal{H}_n) \to Q(\mathbb{B}_n)$, so that the full quantum feature map from X to $Q(\mathbb{B}_n)$ takes the form $\hat{\mathcal{F}}_n = \mathcal{W}_n \circ \mathbf{\Pi}'_n \circ P \circ \delta$.

(2) We construct a linear map $\hat{T}_n : \mathfrak{A} \to B(\mathbb{B}_n)$ from classical observables in $\mathfrak{A}$ to quantum mechanical observables in $B(\mathbb{B}_n)$. This map takes the form $\hat{T}_n = \mathcal{W}_n \circ \mathbf{\Pi}_n \circ T$, where

$\mathbf{\Pi}_n: B(\mathcal{H}) \to B(\mathcal{H}_n)$ is a projection, so that $\hat{T}_n$ yields a quantum computational representation of classical observables passing through intermediate quantum mechanical and matrix mechanical representations. Here, $T: \mathfrak{A} \to B(\mathcal{H})$ is one-to-one on real-valued functions in $\mathfrak{A}$, and Tf is self-adjoint whenever f is real.

Next, we describe the maps governing the temporal evolution of states and observables, represented by horizontal arrows in Fig. 2:

(1) At the quantum mechanical level, states in $Q(\mathcal{H})$ evolve under the operator Ψ^t (horizontal arrow in the left-hand column) induced by a unitary Koopman operator $U^t = e^{tV}$ on $\mathcal{H}$. This evolution is generated by a skew-adjoint generator $V: D(V) \to \mathcal{H}$, defined on a dense subspace $D(V) \subset \mathcal{H}$ and possessing a countable spectrum of eigenfrequencies.

(2) The generator V is mapped to a self-adjoint Hamiltonian $H_n: \mathbb{B}_n \to \mathbb{B}_n$ given by $H_n = \mathcal{W}_n \mathbf{\Pi}_n V/i$. This Hamiltonian is decomposable as a sum $H_n = \sum_j G_j$ of mutually commuting operators $G_j \in B(\mathbb{B}_n)$, each of which is of pure tensor product form, $G_j = \bigotimes_{i=1}^n G_{ij}$. The latter property enables quantum parallelism in the unitary evolution $\hat{\Psi}_n^t: Q(\mathbb{B}_n) \to Q(\mathbb{B}_n)$ at the quantum computational level generated by H_n (see horizontal arrow at the bottom of the left-hand column of Fig. 2). One of our main results is that $\hat{\Psi}^t$ can be implemented via a quantum circuit of size $O(n)$ and no interchannel communication (see Figs. 6 and 8).

(3) The horizontal arrow at the quantum mechanical level represents the action of the Heisenberg evolution operator $\mathcal{U}^t: B(\mathcal{H}) \to B(\mathcal{H})$. Under the assumption that the RKHS $\mathcal{H}$ is invariant under the Koopman operator, $\mathcal{U}^t$ acts on $B(\mathcal{H})$ by conjugation with U^t, i.e., $\mathcal{U}^t A = U^t A U^{t*}$.

(4) The corresponding Heisenberg evolution operator at the quantum computational level, $\hat{\mathcal{U}}_n^t: B(\mathbb{B}_n) \to B(\mathbb{B}_n)$, acting on quantum mechanical observables on the Hilbert space $\mathbb{B}_n$, is represented by the horizontal arrow at the bottom of the right-hand column. This operator is obtained by projection of $\mathcal{U}^t$, *viz.*, $\hat{\mathcal{U}}_n^t = \mathcal{W}_n \mathbf{\Pi}_n \mathcal{U}^t$.

Given a classical initial condition $x \in X$, the quantum computational system constructed by QECD makes probabilistic predictions $\hat{f}_n^{(t)}(x)$ of $f(\Phi^t(x))$ through quantum mechanical measurement of the projection-valued measure (PVM) [4,69] associated with the quantum register on the quantum state $\hat{\rho}_{x,n}^{(t)} := \hat{\Psi}_n^t(\hat{\rho}_{x,n})$, where $\hat{\rho}_{x,n} = \hat{\mathcal{F}}_n(x)$. The state $\hat{\rho}_{x,n}$ is prepared by means of a circuit of size $O(n)$, which is applied to the standard initial state vector of the quantum computer. Furthermore, the measurement step is effected by performing a rotation $\hat{\rho}_{x,n}^{(t)} \mapsto \tilde{\rho}_{x,n}^{(t)}$ by a QFT, which is implementable via a circuit of size $O(n^2)$. An ensemble of such measurements then approximates the quantum mechanical expectation value

$$\langle \hat{T}_n f \rangle_{\hat{\rho}_{x,n}^{(t)}} := f_n^{(t)}(x). \tag{4}$$

The function $x \mapsto f_n^{(t)}(x)$ converges in turn uniformly to the true classical evolution, i.e., $U^t f(x)$, in the large-qubit limit, $n \to \infty$. We will return to these points in a more detailed discussion in Secs. V–VIII.

In summary, the key distinguishing aspects of QECD are as follows:

(1) *Dynamical consistency.* The predictions made by the quantum quantum computational system via (4) converge to the true classical evolution as the number of qubits n increases. In particular, since $\dim \mathbb{B}_n = 2^n$, the convergence is exponentially fast in n.

(2) *Quantum efficiency.* The full circuit implementation of the scheme, including state preparation, dynamical evolution, and measurement, requires a circuit of size $O(n^2)$ and depth $O(n)$. Since, as just mentioned, the dimension of $\mathbb{B}_n$ increases exponentially with n, the quantum computational system constructed by QECD can simulate a 2^n-dimensional subspace of classical observables at an $O(n^2)$ computational cost. This results in an exponential quantum advantage over classical numerical techniques for approximating Koopman operators, where computational complexity scales linearly with the dimension of the subspace, i.e., is $O(2^n)$. We will make a comparison with randomized classical algorithms for quantum circuit simulation [62–64] in Sec. IX.

(3) *State preparation.* The quantum computational state $\hat{\rho}_{x,n}$ corresponding to classical state x is prepared by passing the standard initial state vector of the quantum computer through a circuit of size $O(n)$ and depth $O(1)$. This overcomes the expensive (potentially exponential) state preparation problem affecting many quantum computational algorithms.

(4) *Measurement process.* The process of querying the system to obtain predictions is a standard projective measurement of the quantum register. Importantly, no quantum state tomography or auxiliary classical computation is needed to retrieve the relevant information.

In the ensuing sections, we lay out the properties of the classical system under study (Sec. III), and describe the conversion to the quantum computational system using QECD (Secs. IV–VIII).

III. CLASSICAL DYNAMICS AND OBSERVABLES

A. Dynamical system

We focus on the class of continuous, measure-preserving, ergodic flows with a pure point spectrum generated by finitely many eigenfrequencies and continuous corresponding eigenfunctions. Every such system is topologically conjugate (for our purposes, equivalent) to an ergodic rotation on a d-dimensional torus, so we will set $X = \mathbb{T}^d$ without loss of generality. Using the notation $x = (\theta^1, \ldots, \theta^d)$ to represent a point $x \in \mathbb{T}^d$, where $\theta^j \in [0, 2\pi)$ are canonical angle coordinates, the dynamics is described by the flow map

$$\Phi^t(x) = (\theta^1 + \alpha_1 t, \ldots, \theta^d + \alpha_d t) \mod 2\pi, \tag{5}$$

where $\alpha_1, \ldots, \alpha_d$ are positive, rationally independent (incommensurate) frequency parameters. This dynamical system is also known as a linear flow on the d-torus, but note that $\mathbb{T}^d$ is not a linear space. In dimension $d > 1$, the orbits $\Phi^t(x)$ of the dynamics do not close by incommensurability of the α_j, each forming a dense subset of the torus (i.e., a given orbit passes by any point in $\mathbb{T}^d$ at an arbitrarily small distance). The case $d = 2$ is shown for two choices of α_j in Fig. 3, illustrating the difference between ergodic and nonergodic dynamics. In dimension $d = 1$, the flow map corresponds to a harmonic oscillator on the circle, $\mathbb{T}^1 = S^1$, where each orbit is periodic and samples the whole space.

It is important to note that if the dynamical system is not presented in the form of a torus rotation, then standard

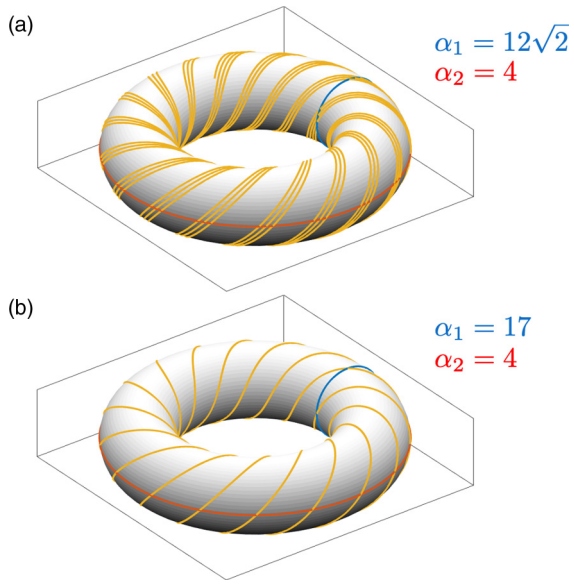

FIG. 3. Ergodic (a) and nonergodic (b) linear flows on the two-dimensional torus $\mathbb{T}^2$. In (a) the ratio of the frequency parameters α_1/α_2 is irrational, and the trajectory starts to fill the torus surface. In (b) the ratio of the frequencies is rational, and the trajectory is closed. The corresponding frequency parameters α_1 and α_2 are given to the right of each figure.

constructions from ergodic theory may be used to transform it into the form in (5). These constructions are based entirely on spectral objects (i.e., eigenfunctions and eigenfrequencies) associated with the Koopman operator of the system. See Appendix D 4 further details. The same constructions allow one to treat the case where X is a periodic or quasiperiodic attractor of a dynamical flow $\Phi^t: \mathcal{X} \to \mathcal{X}$ on a higher-dimensional space $\mathcal{X} \supseteq X$. By virtue of these facts, the quantum mechanical framework described in this paper can readily handle simulations of observables of general measure-preserving, ergodic flows with pure point spectrum. Relevant examples include ODE models on $\mathcal{X} = \mathbb{R}^m$ with quasiperiodic attractors [70], as well as PDE models where $\mathcal{X}$ is an infinite-dimensional function space. The latter class includes many pattern-forming physical systems such as thermal convection flows [71], plasmas [72], and reaction-diffusion systems [73] in moderate-forcing regimes.

At any dimension d, the flow in (5) is measure-preserving and ergodic for a probability measure μ given by the normalized Haar measure. The dynamics of classical observables $f: X \to \mathbb{C}$ is governed by the Koopman operator U^t, which is a linear operator, acting by composition with the dynamical flow in accordance with (3) [44,45,74]. The Koopman operator acts as an isometry on the Banach space of continuous functions on X, i.e., $\|U^t f\|_{C(X)} = \|f\|_{C(X)}$, where $\|f\|_{C(X)} = \max_{x\in X} |f(x)|$ is the uniform norm. In addition, U^t lifts to a unitary operator on the Hilbert space $L^2(\mu)$ associated with the invariant measure. That is, using $\langle f, g\rangle_{L^2(\mu)} = \int_X f^* g \, d\mu$ to denote the $L^2(\mu)$ inner product, we have $\langle U^t f, U^t g\rangle_{L^2(\mu)} = \langle f, g\rangle_{L^2(\mu)}$ for all $f, g \in L^2(\mu)$, which implies, in conjunction with the invertibility of Φ^t, that

$$U^{t*} = (U^t)^{-1}.$$

Here, U^{t*} denotes the operator adjoint, which is also frequently denoted as $(U^t)^\dagger$. The collection $\{U^t: L^2(\mu) \to L^2(\mu)\}_{t\in\mathbb{R}}$ then forms a strongly continuous unitary group under composition of operators [75]. See again Fig. 1.

By Stone's theorem on one-parameter unitary evolution groups [76], the Koopman group on $L^2(\mu)$ has a skew-adjoint infinitesimal generator, i.e., an operator $V: D(V) \to L^2(\mu)$ defined on a dense subspace $D(V) \subset L^2(\mu)$ satisfying

$$V^* = -V \quad \text{and} \quad Vf = \lim_{t\to 0} \frac{U^t f - f}{t}, \tag{6}$$

for all $f \in D(V)$. The generator gives the Koopman operator at any time t by exponentiation,

$$U^t = e^{tV}. \tag{7}$$

Modulo multiplication by $1/i$ to render it self-adjoint, it plays an analogous role to a quantum mechanical Hamiltonian generating the unitary Heisenberg evolution operators.

As already noted, the torus rotation in (5) is a canonical representative of a class of continuous-time continuous dynamical systems on topological spaces with quasiperiodic dynamics generated by finitely many basic frequencies. This means that every such system can be transformed into an ergodic torus rotation of a suitable dimension by a homeomorphism (continuous, invertible map with continuous inverse). By specializing to this class of systems (as opposed to a more general measure-preserving, ergodic flow), we gain two important properties:

(1) The dynamics has no mixing (chaotic) component. This implies that the spectrum of the Koopman operator for this system acting on $L^2(\mu)$, or a suitable RKHS as in what follows, is of "pure point" type, obviating complications arising from the presence of continuous spectrum as would be the case under mixing dynamics.

(2) The state space X is a smooth, closed manifold with the structure of a connected, Abelian Lie group. The Abelian group structure, in particular, renders this system amenable to analysis with Fourier analytic tools.

Below, we use a d-dimensional vector $j = (j_1, \ldots, j_d) \in \mathbb{Z}^d$ to represent a generic multi-index, and

$$\phi_j(x) = \prod_{m=1}^{d} \varphi_{j_m}(\theta^m) \quad \text{with} \quad \varphi_l(\theta) = e^{il\theta}, \tag{8}$$

to represent the Fourier functions on $\mathbb{T}^d$. In Sec. XII, we will discuss possible avenues for extending the framework presented here to other classes of dynamical systems, such as mixing dynamical systems with continuous spectra of the Koopman operators.

B. Algebra of observables

According to the scheme described in Sec. II B, we perform quantum conversion of an (Abelian) algebra $\mathfrak{A}$ of classical observables, i.e., a space of complex-valued functions on X which is closed under the pointwise product of functions. We construct $\mathfrak{A}$ such that it is a subalgebra of $C(X)$ with additional (here, C^∞) regularity and RKHS structure. This structure is induced by a smooth, positive-definite kernel

function $\tilde{k}: X \times X \to \mathbb{R}$, which has the following properties for every point $x \in X$ and function $f \in \mathfrak{A}$:

(1) The kernel section $\tilde{k}_x := \tilde{k}(x, \cdot)$ lies in $\mathfrak{A}$.

(2) Pointwise evaluation, $x \mapsto f(x)$, is continuous, and satisfies

$$f(x) = \langle \tilde{k}_x, f \rangle_{\mathfrak{A}}, \tag{9}$$

where $\langle \cdot, \cdot \rangle_{\mathfrak{A}}$ is the inner product of $\mathfrak{A}$.

Equation (9) is known as the reproducing property, and underlies the many useful properties of RKHSs for tasks such as function approximation and learning. Note, in particular, that L^2 spaces, which are more commonly employed in Koopman operator theory and numerical techniques (see Sec. III A), do not have a property analogous to (9). In fact, pointwise evaluation is not even defined for the $L^2(\mu)$ Hilbert space on $\mathbb{T}^d$. See Refs. [46,47,77] for detailed expositions on RKHS theory.

Our construction of $\mathfrak{A}$ follows Ref. [50]. We begin by setting parameters $p \in (0, 1)$ and $\tau > 0$, and defining the map $|\cdot|_p: \mathbb{Z}^d \to \mathbb{R}_+$,

$$|j|_p := |j_1|^p + \cdots + |j_d|^p,$$

and the functions $\psi_j \in C(X)$,

$$\psi_j := e^{-\tau |j|_p / 2} \phi_j \quad \text{with} \quad j \in \mathbb{Z}^d.$$

We then define a kernel $\tilde{k}: X \times X \to \mathbb{R}_+$ via the series

$$\tilde{k}(x, x') = \sum_{j \in \mathbb{Z}^d} \psi_j^*(x) \psi_j(x'), \tag{10}$$

where the sum over j converges uniformly on $X \times X$ to a smooth function. Intuitively, τ can be thought of as a locality parameter for the kernel, meaning that as τ decreases $\tilde{k}(x, x')$ becomes increasingly concentrated near $x = x'$, approaching a δ function as $\tau \to 0$.

An important property of the kernel that holds for any $\tau > 0$ is that it is translation-invariant on the Abelian group $X = \mathbb{T}^d$. That is, using additive notation to represent the binary group operation on X, we have

$$\tilde{k}(x + y, x' + y) = \tilde{k}(x, x'), \quad \forall x, x', y \in X. \tag{11}$$

In particular, setting $y = \Phi^t(e)$, where e is the identity element of X, and noticing that the dynamical flow from (5) satisfies $\Phi^t(x) = x + \Phi^t(e)$, we deduce the dynamical invariance property

$$\tilde{k}(\Phi^t(x), \Phi^t(x')) = \tilde{k}(x, x'), \quad \forall x, x' \in X, \quad \forall t \in \mathbb{R}.$$

In Ref. [50] it was shown that for every $p > 0$ and $\tau > 0$, the kernel $\tilde{k}$ in (10) is a strictly positive-definite kernel on X, so it induces an RKHS, $\mathfrak{A}$, which is a dense subspace of $C(X)$. One can verify that the collection $\{\psi_j : j \in \mathbb{Z}^d\}$ forms an orthonormal basis of $\mathfrak{A}$, consisting of scaled Fourier functions, so every observable $f \in \mathfrak{A}$ admits the expansion

$$f = \sum_{j \in \mathbb{Z}^d} \tilde{f}_j \psi_j = \sum_{j \in \mathbb{Z}^d} \tilde{f}_j e^{-\tau |j|_p / 2} \phi_j,$$

where the sum over j converges in $\mathfrak{A}$ norm. The above manifests the fact that $\mathfrak{A}$ contains continuous functions with Fourier coefficients decaying faster than any polynomial, implying in turn that every element of $\mathfrak{A}$ is a smooth function in $C^\infty(X)$.

TABLE I. Properties of representative spaces of classical observables on the compact Abelian group $X = \mathbb{T}^d$. The space $\mathfrak{A}$ is an RKHA, which, in addition to being an RKHS, has a Banach *-algebra structure.

	$L^2(\mu)$	$L^\infty(\mu)$	$C(X)$	$C^\infty(X)$	$\mathfrak{A}$
Completeness	✓	✓	✓	×	✓
Hilbert space structure	✓	×	×	×	✓
Pointwise evaluation	×	×	✓	✓	✓
*-algebra structure	×	✓	✓	✓	✓
C^∞ regularity	×	×	×	✓	✓

It can also be shown that the RKHS induced by $\tilde{k}$ acquires an important special property which is not shared by generic RKHSs—namely, it becomes an Abelian, unital, Banach *-algebra under pointwise multiplication of functions. We list the defining properties for completeness in Appendix A 1. In Ref. [50], the space $\mathfrak{A}$ was referred to as a reproducing kernel Hilbert algebra (RKHA) as it enjoys the properties of both RKHSs and Banach algebras. In particular, a distinguishing aspect of $\mathfrak{A}$ is that it simultaneously has Hilbert space structure [as $L^2(\mu)$] and Banach *-algebra structure [as $C(X)$], while also allowing pointwise evaluation by continuous functionals [i.e., the reproducing property in (9)]. The RKHAs associated with the family of kernels in (10) are examples of harmonic Hilbert spaces on locally compact Abelian groups [48], and are also closely related (by Fourier transforms) to weighted convolution algebras [49] on the dual group $\mathbb{Z}^d$ of $X = \mathbb{T}^d$.

Table I summarizes the properties of $\mathfrak{A}$ and other function spaces on X employed in this work. In what follows, we shall let $\mathfrak{A}_{\text{sa}}$ denote the set of self-adjoint elements of $\mathfrak{A}$, i.e., the elements $f \in \mathfrak{A}$ satisfying $f^* = f$. Since the * operation of $\mathfrak{A}$ corresponds to complex conjugation of functions, it follows that $\mathfrak{A}_{\text{sa}}$ contains the real-valued functions in $\mathfrak{A}$. Note that if $f = \sum_{j \in \mathbb{Z}^d} \tilde{f}_j \psi_j$ is an element of $\mathfrak{A}_{\text{sa}}$, then its expansion coefficients in the ψ_j basis satisfy $\tilde{f}_j^* = \tilde{f}_{-j}$.

Next, we state a product formula for the orthonormal basis functions ψ_j, which follows directly from their definition, *viz.,*

$$\psi_j \psi_l = c_{jl} \psi_{j+l} \quad \text{with}$$
$$c_{jl} = \exp\left(-\tau \frac{|j|_p + |l|_p - |j + l|_p}{2} \right). \tag{12}$$

In the above, we interpret the coefficients c_{jl} as "structure constants" of the RKHA $\mathfrak{A}$. Figure 4 displays representative matrices formed by the c_{jl} in dimension $d = 1$ and 2 for $p = 1/4$ and $\tau = 1/4$.

In the special case $d = 1$, we will let $\mathfrak{A}^{(1)}$ be the RKHA on the circle $S^1 \equiv \mathbb{T}^1$ constructed as above. We denote the reproducing kernel of $\mathfrak{A}^{(1)}$ by $\tilde{k}^{(1)}$, and let $\psi_j^{(1)}$, $j \in \mathbb{Z}$, be the corresponding orthonormal basis functions with $\psi_j^{(1)}(\theta) = e^{-|j|^p \tau / 2} \varphi_j(\theta)$. It then follows that $\mathfrak{A}$ admits the tensor product

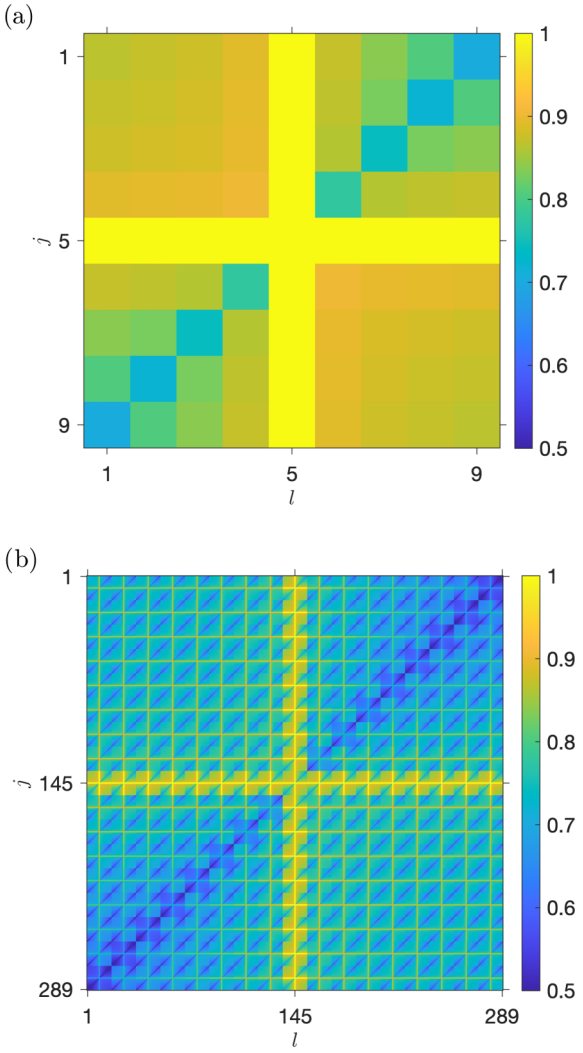

FIG. 4. Structure constant matrices c_{jl} for reproducing kernel Hilbert algebras on (a) the circle with $d = 1$ and (b) the 2-torus with $d = 2$. In both cases we use the parameter values $p = 1/4$ and $\tau = 1/4$ as given in (12). In (a), we consider indices in the range $-2^{n-1} \leqslant j, l \leqslant 2^{n-1}$ with $n = 3$. In (b), the multi-indices $j = (j_1, j_2)$ and $l = (l_1, l_2)$ satisfy $-2^{n/2-1} \leqslant j_i, l_i \leqslant 2^{n/2-1}$ with $n = 8$. In both (a) and (b), we map j and l into standard matrix indices $1, 2, \ldots, (2^{n/d} + 1)^d$ [which results in $(2^4 + 1)^2 = 289$ for (b)] by lexicographical ordering. The matrix in (b) is thus equal to the Kronecker product of the matrix in (a) with itself.

factorization

$$\mathfrak{A} = \bigotimes_{i=1}^{d} \mathfrak{A}^{(1)}, \tag{13}$$

and the reproducing kernel and orthonormal basis functions of $\mathfrak{A}$ similarly factorize as

$$\tilde{k}(x, x') = \prod_{i=1}^{d} \tilde{k}^{(1)}(\theta^i, \theta^{i\prime}), \qquad \psi_j(x) = \prod_{i=1}^{d} \psi_{j_i}^{(1)}(\theta^i),$$

where $j = (j_1, \ldots, j_d)$, and θ^i, $\theta^{i\prime}$ are canonical angle coordinates of the points $x = (\theta^1, \ldots, \theta^d)$, $x' = (\theta^{1\prime}, \ldots, \theta^{d\prime})$, respectively [see also (8)].

C. Evolution of RKHA observables

From an operator-theoretic perspective, simulating the dynamical evolution of a continuous classical observable $f \in C(X)$ can be understood as approximating the Koopman operator U^t on $C(X)$; for, if U^t were known one could use it to compute $U^t f(x) = f(\Phi^t(x))$ for every observable $f \in C(X)$, time $t \in \mathbb{R}$, and initial condition $x \in X$ (cf. Sec. II). Yet, despite its theoretical appeal, consistently approximating the Koopman operator on $C(X)$ is challenging in practice, as this space lacks the Hilbert space structure underpinning commonly employed operations used in numerical techniques, such as orthogonal projections (see Table I). For a measure-preserving, ergodic dynamical system such as the torus rotation in (5), a natural alternative is to consider the unitary Koopman operator on the $L^2(\mu)$ Hilbert space associated with the invariant measure μ. While this choice addresses the absence of orthogonal projections on $C(X)$, $L^2(\mu)$ lacks the notion of pointwise evaluation of functions, so one must correspondingly abandon the notion of pointwise forecasting in this space.

In light of the above considerations, RKHSs emerge as attractive candidates of spaces of classical observables in which to perform simulation, as they allow pointwise evaluation through the reproducing property in (9) while having a Hilbert space structure. Unfortunately, an obstruction to using RKHSs in dynamical systems forecasting is that a general RKHS $\mathcal{H}$ on X need not be preserved under the dynamics, even if the reproducing kernel k is continuous. That is, in general, if $f: X \to \mathbb{C}$ lies in an RKHS, the composition $f \circ \Phi^t$ need not lie in the same space, and thus the Koopman operator is not well defined as an operator mapping the RKHS into itself [27]. Intuitively, this is because membership of a function f in an RKHS generally imposes stringent requirements in its regularity, as we discussed for example in Sec. IV A with the rapid decay of Fourier coefficients, which need not be preserved by the dynamical flow.

An exception to this obstruction occurs when the reproducing kernel is translation-invariant, which holds true for the class of kernels introduced in Sec. III B [see (11)]. In fact, it can be shown [55] that the RKHA $\mathfrak{A}$ associated with the kernel $\tilde{k}$ in (10), is invariant under the Koopman operator U^t for all $t \in \mathbb{R}$, and $U^t: \mathfrak{A} \to \mathfrak{A}$ is unitary and strongly continuous. Analogously to the $L^2(\mu)$ case, the evolution group $\{U^t: \mathfrak{A} \to \mathfrak{A}\}_{t\in\mathbb{R}}$ is uniquely characterized through its skew-adjoint generator $V: D(V) \to \mathfrak{A}$, defined on a dense subspace $D(V) \subset \mathfrak{A}$, and acting on observables as displayed in (6).

For the torus rotation in (5), V is diagonalizable in the $\{\psi_j\}$ basis of $\mathfrak{A}$. That is, for $j = (j_1, \ldots, j_d) \in \mathbb{Z}^d$, we have

$$V\psi_j = i\omega_j\psi_j,$$

where ω_j is a real eigenfrequency given by

$$\omega_j = j_1\alpha_1 + \cdots + j_d\alpha_d. \tag{14}$$

Moreover, V admits a decomposition into mutually commuting, skew-adjoint generators $V_1, \ldots, V_d$ satisfying

$$V_l\psi_j = i j_l \alpha_l \psi_j \quad \text{with} \quad l = 1, \ldots, d. \tag{15}$$

In particular, since $\{\psi_j\}$ is an orthonormal basis, (15) completely characterizes V_l, and we have

$$V = V_1 + \cdots + V_d, \quad [V_j, V_l] = 0, \quad [V_j, V] = 0. \tag{16}$$

It should be noted that the Koopman generator on $L^2(\mu)$ admits a similar decomposition to (16); see e.g., Ref. [25] for further details. Analogously to the $L^2(\mu)$ case, the Koopman operator on $\mathfrak{A}$ can be recovered at any $t \in \mathbb{R}$ from the generator by exponentiation as given in (7).

IV. EMBEDDING INTO AN INFINITE-DIMENSIONAL QUANTUM SYSTEM

The initial stages of the QECD procedure outlined in Sec. II involve embedding classical states and observables into states and observables of quantum system associated with an infinite-dimensional RKHS $\mathcal{H}$, arriving at the quantum mechanical level depicted in Fig. 2. In this section, we describe the construction of this quantum system and associated embeddings of classical states and observables. First, in Sec. IV A we build $\mathcal{H}$ as a subspace of the RKHA $\mathfrak{A}$ from Sec. III B. Then, in Secs. IV B and IV C we establish representation maps $Q: X \to Q(\mathcal{H})$ and $T: \mathfrak{A} \to B(\mathcal{H})$ from classical states and observables into quantum mechanical states and observables, respectively, on $\mathcal{H}$. Note that the quantum mechanical embedding of states Q passes through an intermediate classical statistical level associated with probability measures on the classical state space (second row in the left-hand column of Fig. 2). In Secs. IV D and IV E, we establish the classical-quantum consistency and associated dynamical properties of our embeddings.

A. Reproducing kernel Hilbert space

We choose $\mathcal{H}$ as an infinite-dimensional subspace of the RKHA $\mathfrak{A}$ containing zero-mean functions. For that, we introduce the (infinite) index set

$$J = \{(j_1, \ldots, j_d) \in \mathbb{Z}^d : j_i \neq 0\}, \tag{17}$$

and define $\mathcal{H}$ as the corresponding infinite-dimensional closed subspace

$$\mathcal{H} = \overline{\text{span}\{\psi_j : j \in J\}}.$$

The space $\mathcal{H}$ is then an RKHS with the reproducing kernel

$$k(x, x') = \sum_{j \in J} \psi_j^*(x)\psi_j(x'). \tag{18}$$

In particular, for every $f \in \mathcal{H}$, which is necessarily an element of $\mathfrak{A}$, the reproducing property in (9) reads

$$f(x) = \langle k_x, f\rangle_{\mathcal{H}} = \langle \tilde{k}_x, f\rangle_{\mathfrak{A}},$$

where $k_x := k(x, \cdot)$ is the section of the kernel k at $x \in X$, and $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ denotes the inner product of $\mathcal{H}$.

By excluding zero indices from the index set J, every element f of $\mathcal{H}$ has zero mean, $\int_X f \, d\mu = 0$, as noted above. The reason for adopting this particular definition for $\mathcal{H}$, instead of, e.g., working with the entire space $\mathfrak{A}$, is that later on it will facilitate construction of 2^n-dimensional subspaces $\mathcal{H}_n \subset \mathcal{H}$ suitable for quantum computation (see Sec. V). In what follows, $\Pi: \mathfrak{A} \to \mathfrak{A}$ will denote the orthogonal projection with $\text{ran}\Pi = \mathcal{H}$. Moreover, we set

$$\kappa = k(x, x) = \sum_{j \in J} e^{-\tau|j|_p}, \quad \tilde{\kappa} = \tilde{k}(x, x) = \sum_{j \in \mathbb{Z}^d} e^{-\tau|j|_p},$$

where these definitions are independent of the point $x \in X$ by (11). We also note that, by construction, $\mathcal{H}$ is a Koopman-invariant subspace of $\mathfrak{A}$, so we may define unitary Koopman operators $U^t: \mathcal{H} \to \mathcal{H}$ by restriction of $U^t : \mathfrak{A} \to \mathfrak{A}$ from Sec. III C.

B. Representation of states with a quantum feature map

For our purposes, a key property that the RKHS structure of $\mathcal{H}$ endows is the feature map, which is the continuous map $F: X \to \mathcal{H}$ mapping classical state $x \in X$ to the RKHS function

$$F(x) = k_x. \tag{19}$$

It can be shown that for the choice of kernel in (18), F is an injective map, and the functions $\{F(x) \in \mathcal{H} : x \in X\}$ are linearly independent. It is then natural to think of the normalized feature vectors

$$\xi_x := \frac{k_x}{\|k_x\|_{\mathcal{H}}} = \frac{k_x}{\sqrt{\kappa}} \tag{20}$$

as "wave functions" corresponding to classical states $x \in X$.

We can generalize this idea by associating every such wave function ξ_x with the pure quantum state $\rho_x = \langle \xi_x, \cdot \rangle_{\mathcal{H}} \xi_x$. The mapping $\mathcal{F}: X \to Q(\mathcal{H})$ with

$$\mathcal{F}(x) = \rho_x \tag{21}$$

then describes an embedding of the classical state space X into quantum mechanical states in $Q(\mathcal{H})$, which we refer to as a quantum feature map. Note that there is no loss of information in representing $x \in X$ by $\rho_x \in Q(\mathcal{H})$. Moreover, $\mathcal{F}$ can be understood as a composition $\mathcal{F} = P \circ \delta$, where $\delta: X \to \mathcal{P}(X)$ maps classical state $x \in X$ to the Dirac probability measure $\delta_x \in \mathcal{P}(X)$, and $P: \mathcal{P}(X) \to Q(\mathcal{H})$ is a map from classical probability measures on X to quantum states on $\mathcal{H}$, such that

$$P(p) = \int_X \rho_x \, dp(x). \tag{22}$$

The map P describes an embedding of the state space X into the space of probability measures $\mathcal{P}(X)$, i.e., the classical statistical level in the left-hand column of Fig. 2. See Ref. [50] for further details on the properties of this map.

By virtue of it being an RKHS, we can also define classical and quantum feature maps for the RKHA $\mathfrak{A}$. Specifically, we set $\tilde{F}: X \to \mathfrak{A}$ and $\tilde{\mathcal{F}}: X \to Q(\mathfrak{A})$, where

$$\tilde{F}(x) = \langle \tilde{k}_x, \cdot \rangle_{\mathfrak{A}}, \quad \tilde{\mathcal{F}}(x) = \langle \tilde{\xi}_x, \cdot \rangle_{\mathfrak{A}} \tilde{\xi}_x, \tag{23}$$

and $\tilde{\xi}_x = \tilde{k}_x / \|\tilde{k}_x\|_{\mathfrak{A}}$. The feature maps $\tilde{F}$ and $\tilde{\mathcal{F}}$ have analogous properties to F and $\mathcal{F}$, respectively, which we do not discuss here in the interest of brevity.

C. Representation of observables

The quantum mechanical representation of classical observables in $\mathfrak{A}$ is considerably facilitated by the Banach algebra structure of that space. In Sec. IV C 1, we leverage

that structure to build representation maps from functions in $\mathfrak{A}$ to bounded linear operators in $B(\mathfrak{A})$. Then, in Sec. IV C 2, we consider associated representations mapping into bounded linear operators on the RKHS $\mathcal{H}$ (which is a strict subspace of $\mathfrak{A}$), arriving at the map $T:\mathfrak{A}\to B(\mathcal{H})$ depicted in the right-hand column of Fig. 2. Additional details on the construction are provided in Appendix A.

1. Representation on the RKHA $\mathfrak{A}$

We begin by noting that the joint continuity of the multiplication operation of Banach algebras [see (A1)] implies that for every $f\in\mathfrak{A}$ the multiplication operator $A_f: g\mapsto fg$ is well defined as a bounded operator in $B(\mathfrak{A})$. This leads to the regular representation $\pi:\mathfrak{A}\to B(\mathfrak{A})$, which is the algebra homomorphism of $\mathfrak{A}$ into $B(\mathfrak{A})$, mapping classical observables in $\mathfrak{A}$ to their corresponding multiplication operator,

$$\pi f := A_f. \tag{24}$$

This mapping is a homomorphism since

$$\pi(fg) = A_{fg} = A_f A_g, \quad \forall f,g\in\mathfrak{A},$$

and it is injective (i.e., faithful as a representation) since $[\pi(f-f')]1_X = f-f'\neq 0$ whenever $f\neq f'$. However, π is not a *-representation; i.e., $\pi(f^*)$ is not necessarily equal to A_f^*. In particular, A_f need not be a self-adjoint operator in $B(\mathfrak{A})$ if f is a self-adjoint element in $\mathfrak{A}_{\rm sa}$. To construct a map from $\mathfrak{A}$ into the self-adjoint operators in $B(\mathfrak{A})$, we define $\tilde{T}:\mathfrak{A}\to B(\mathfrak{A})$ with

$$\tilde{T}f = \frac{\pi f + (\pi f)^*}{2}. \tag{25}$$

By construction, $\tilde{T}f$ is self-adjoint for all $f\in\mathfrak{A}$, and it can also be shown (see Appendix A 2) that $\tilde{T}$ is injective on $\mathfrak{A}_{\rm sa}$. That is, $\tilde{T}$ provides a one-to-one mapping between real-valued functions in $\mathfrak{A}$ and self-adjoint operators in $B(\mathfrak{A})$.

It follows from the product formula in (12) that if $f\in\mathfrak{A}$ has the expansion $f=\sum_{j\in\mathbb{Z}^d}\tilde{f}_j\psi_j$, where $\tilde{f}_j=\langle\psi_j,f\rangle_{\mathfrak{A}}$, then the corresponding multiplication operator $A_f=\pi f$ has the matrix elements

$$(A_f)_{ij} := \langle\psi_i, A_f\psi_j\rangle_{\mathfrak{A}} = \langle\psi_i, f\psi_j\rangle_{\mathfrak{A}},$$

and thus

$$(A_f)_{ij} = c_{j,i-j}\tilde{f}_{i-j}. \tag{26}$$

Correspondingly, the matrix elements of the self-adjoint operator $S_f := \tilde{T}f$ are given by

$$(S_f)_{ij} := \langle\psi_i, S_f\psi_j\rangle_{\mathfrak{A}} = \frac{(A_f)_{ij}+(A_f)^*_{ji}}{2}.$$

If, in addition, f lies in $\mathfrak{A}_{\rm sa}$, then we have $\tilde{f}^*_{j-i}=\tilde{f}_{i-j}$ and the formula above reduces to

$$(S_f)_{ij} = \frac{c_{j,i-j}+c_{i,j-i}}{2}\tilde{f}_{i-j}. \tag{27}$$

Here, of particular interest are the multiplication and self-adjoint operators representing the basis elements of $\mathfrak{A}$, i.e., A_{ψ_l} and S_{ψ_l}, respectively, for $l\in\mathbb{Z}^d$. Since $\tilde{f}_{i-j}=\delta_{l,i-j}$ for $f=\psi_l$, it follows from (26) that after a suitable lexicographical ordering of multi-indices (as in Fig. 4), $(A_f)_{ij}$ forms a banded matrix with nonzero elements only in the diagonal corresponding to multi-index k. Figure 5(a) illustrates the nonzero matrix elements of A_{ψ_1} in the one-dimensional case, $d=1$. Similarly, the self-adjoint operator S_{ψ_1} is a bidiagonal matrix with nonzero entries in the diagonals corresponding to ±1, as shown in Fig. 5(b).

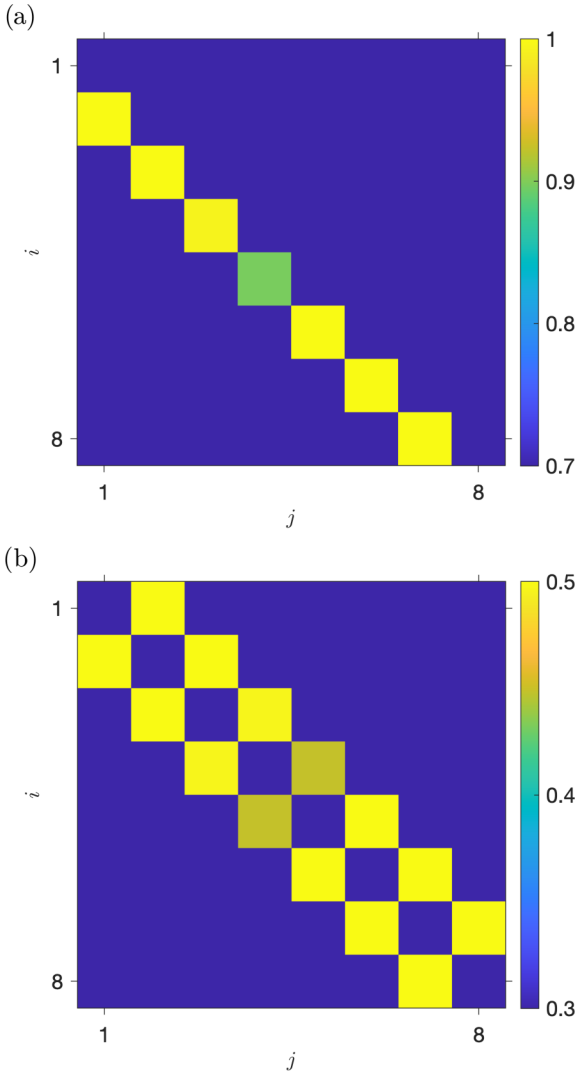

FIG. 5. Matrix elements $(A_{\psi_1})_{ij}$ (a) and $(S_{\psi_1})_{ij}$ (b) of the multiplication operator A_{ψ_1} [Eq. (26)] and the self-adjoint operator S_{ψ_1} [Eq. (27)] representing the basis function ψ_1 for dimension $d=1$. As in Fig. 4(a), we consider the reproducing kernel Hilbert algebra $\mathfrak{A}$ on the circle from with $p=1/4$ and $\tau=1/4$, and map the indices i and j into standard matrix indices $1,2,\ldots,2^n$ with $n=3$. The matrix in (a) has nonzero elements only in the first lower diagonal, $i-j=1$. The matrix in (b) is a symmetric bidiagonal matrix with elements in the first upper and lower diagonals, $i-j=\pm1$.

We deduce from these observations that if f is a bandlimited observable (i.e., expressible as a finite linear combination of Fourier functions ϕ_j), A_f is represented by a banded matrix, whose lth diagonal comprises the structure constants c_{lj} multiplied by $\tilde{f}_l$. The matrix representing S_f is also banded whenever f is bandlimited. If, in addition, f is real, the lth diagonal of $(S_f)_{ij}$ is given by the multiple of $(c_{lj}+c_{li})/2$ with $\tilde{f}_l$.

2. Representation on the RKHS $\mathcal{H}$

We now take up the task of defining analogs of the maps $\pi:\mathfrak{A}\to B(\mathfrak{A})$ and $T:\mathfrak{A}\to B(\mathfrak{A})$ from Sec. IV C 2, mapping elements of $\mathfrak{A}$ to bounded operators on the RKHS $\mathcal{H}\subset\mathfrak{A}$ (i.e., the Hilbert space underlying the infinite-dimensional system at the quantum mechanical level). To that end, let $\mathbf{\Pi}$ be the projection map from $B(\mathfrak{A})$ to $B(\mathcal{H})$, defined as

$$\mathbf{\Pi}A := \Pi A\Pi, \tag{28}$$

where Π is the orthogonal projection from $\mathfrak{A}$ to $\mathcal{H}$ introduced in Sec. IV A. One can explicitly verify that the map $\mathbf{\Pi}\circ\pi:\mathfrak{A}\to B(\mathcal{H})$ is injective, so there is no loss of information in representing $f\in\mathfrak{A}$ by $\mathbf{\Pi}(\pi f)\in B(\mathcal{H})$ as opposed to $\pi f\in B(\mathfrak{A})$. For our purposes, however, in addition to injectivity we require that our representation maps provide value-level consistency between classical and quantum measurements (in a sense made precise in Sec. IV D below). For that, it becomes necessary to modify the map $\mathbf{\Pi}\circ\pi$, as well as its self-adjoint counterpart $\mathbf{\Pi}\circ\tilde{T}$, to take into account the contractive effect of the projection $\mathbf{\Pi}$.

In Appendix A 3, we construct a self-adjoint, invertible operator $L:\mathfrak{A}\to\mathfrak{A}$, which is diagonal in the $\{\psi_j\}$ basis, and whose role is to counterbalance that contraction. Specifically, we define $\varpi:\mathfrak{A}\to B(\mathcal{H})$ and $T:\mathfrak{A}\to B(\mathcal{H})$ with

$$\varpi = \mathbf{\Pi}\circ\pi\circ L^{-1},\quad T = \mathbf{\Pi}\circ\tilde{T}\circ L^{-1}. \tag{29}$$

Here, L^{-1} inflates the expansion coefficients of functions in the $\{\psi_j\}$ basis of $\mathfrak{A}$, absorbing the contractive action of $\mathbf{\Pi}$. Analogously to π and $\tilde{T}$, respectively, ϖ is one-to-one, and T is one-to-one on the real functions in $\mathfrak{A}_{\mathrm{sa}}$. Moreover, every operator in the range of T is self-adjoint. The map T provides the representation of classical observables in $\mathfrak{A}$ by self-adjoint operators in $B(\mathcal{H})$ at the quantum mechanical level, depicted by vertical arrows in the right-hand column of Fig. 2.

D. Classical-quantum consistency

We now come to a key property of the regular representation π and the associated map $\tilde{T}$, which is a consequence of the reproducing property and Banach algebra structure of $\mathfrak{A}$. Namely, π and $\tilde{T}$ provide a consistent correspondence between evaluation of classical observables and quantum mechanical expectation values. To see this, for any quantum state $\varrho\in Q(\mathfrak{A})$ and quantum mechanical observable $A\in B(\mathfrak{A})$, let

$$\langle A\rangle_\varrho := \mathrm{tr}(\varrho A) \tag{30}$$

be the standard quantum mechanical expectation functional. Then, it follows from the reproducing property in (9), the definition of the quantum feature map $\tilde{\mathcal{F}}$ in (23), and the definition of the regular representation in (24) that for any observable $f\in\mathfrak{A}$ and classical state $x\in X$,

$$f(x) = \langle\pi f\rangle_{\varrho_x} = \langle\tilde{T}f\rangle_{\varrho_x}, \tag{31}$$

where $\varrho_x=\tilde{\mathcal{F}}(x)$. The last equality in (31) requires that f is a self-adjoint element in $\mathfrak{A}_{\mathrm{sa}}$; see Ref. [50] for further details. Equation (31) shows, in particular, that by passing to the quantum mechanical representation we maintain pointwise consistency with the classical measurement processes for special sets of quantum mechanical observables and states. These are the self-adjoint operators S_f and the pure states ϱ_x.

To express these relationships in terms of matrix elements, note first that the quantum state ϱ_x satisfies

$$(\varrho_x)_{ij} := \langle\psi_i,\varrho_x\psi_j\rangle_{\mathfrak{A}} = \frac{\langle\psi_i,\tilde{k}_x\rangle_{\mathfrak{A}}\langle\tilde{k}_x,\psi_j\rangle_{\mathfrak{A}}}{\tilde{\kappa}} = \frac{\psi_i^*(x)\psi_j(x)}{\tilde{\kappa}}. \tag{32}$$

Combining this result with (26), we obtain

$$f(x) = \mathrm{tr}(\varrho_x(\pi f)) = \sum_{i,j\in\mathbb{Z}^d}(\varrho_x)_{ij}(A_f)_{ji} = \sum_{i,j\in\mathbb{Z}^d}\frac{\psi_i^*(x)\psi_j(x)c_{i,j-i}\tilde{f}_{j-i}}{\tilde{\kappa}},$$

and this relationship holds irrespective of whether f is self-adjoint or not. If f is a self-adjoint element in $\mathfrak{A}_{\mathrm{sa}}$, then we can use the matrix elements of the self-adjoint operator S_f from (27), in conjunction with the fact that ϱ_x is also self-adjoint, to arrive at the expression

$$\begin{aligned} f(x) &= \mathrm{tr}(\varrho_x(\tilde{T}f)) \\ &= \sum_{i,j\in\mathbb{Z}^d}(\varrho_x)_{ij}(S_f)_{ji} \\ &= \sum_{i,j\in\mathbb{Z}^d}\frac{\psi_i^*(x)\psi_j(x)(c_{i,j-i}+c_{j,i-j})\tilde{f}_{j-i}}{2\tilde{\kappa}}. \end{aligned}$$

Even though $\mathcal{H}$ is a strict subspace of the RKHA $\mathfrak{A}$, it is still possible to consistently recover all predictions made for classical observables, as we describe in Appendix A 3. There, we show that the modified versions $\varpi:\mathfrak{A}\to B(\mathcal{H})$ and $T:\mathfrak{A}\to B(\mathcal{H})$ of $\pi:\mathfrak{A}\to B(\mathfrak{A})$ and $\tilde{T}:\mathfrak{A}\to B(\mathfrak{A})$, respectively [defined in (29)], satisfy the analogous consistency relation to (31), i.e.,

$$f(x) = \langle\varpi f\rangle_{\rho_x} = \langle Tf\rangle_{\rho_x}, \tag{33}$$

where $\rho_x=\mathcal{F}(x)$ is the quantum state on $\mathcal{H}$ obtained from the feature map in (21). As with (31), the first equality in (33) holds for any $f\in\mathfrak{A}$ and the second holds for real-valued elements $f\in\mathfrak{A}_{\mathrm{sa}}$.

E. Dynamical evolution

In this section, we describe the dynamics of quantum states and observables associated with the RKHA $\mathfrak{A}$ and RKHS $\mathcal{H}\subset\mathfrak{A}$, and establish consistency relations between the classical and quantum evolution.

First, recall that the Koopman operators U^t act on $\mathfrak{A}$ as a unitary evolution group. As a result, there is an induced action $\mathcal{U}^t:B(\mathfrak{A})\to B(\mathfrak{A})$ on quantum mechanical observables in $B(\mathfrak{A})$, given by

$$\mathcal{U}^tA = U^tAU^{t*}. \tag{34}$$

This action has the important property of being compatible with the action of the Koopman operator on functions in $\mathfrak{A}$ under the regular representation. Specifically, for every $f\in\mathfrak{A}$ and $t\in\mathbb{R}$, we have

$$\mathcal{U}^t(\pi f) = \pi(U^tf). \tag{35}$$

The unitary evolution in (34) has a corresponding dual action $\Psi^t: Q(\mathfrak{A}) \to Q(\mathfrak{A})$ on quantum states, given by

$$\Psi^t(\varrho) = U^{t*}\varrho U^t \equiv \mathcal{U}^{-t}\varrho. \tag{36}$$

One can verify that this action is compatible with the classical dynamical flow under the feature map $\tilde{\mathcal{F}}: X \to Q(\mathfrak{A})$, *viz.*,

$$\Psi^t(\tilde{\mathcal{F}}(x)) = \tilde{\mathcal{F}}(\Phi^t(x)). \tag{37}$$

Using (31), (34), and (36), we arrive at the consistency relationships

$$U^t f(x) = \langle \mathcal{U}^t(\pi f)\rangle_{\varrho_x} = \langle \pi f\rangle_{\Psi^t(\varrho_x)}, \tag{38}$$

with $\varrho_x = \tilde{\mathcal{F}}(x)$. This holds for every classical observable $f \in \mathfrak{A}$, initial condition $x \in X$, and evolution time $t \in \mathbb{R}$. If, in addition, f is a self-adjoint element in $\mathfrak{A}_{\rm sa}$, we may compute the evolution $U^t f$ using the self-adjoint operator $\tilde{T} f$, which is accessible via physical measurements. That is, for $f \in \mathfrak{A}_{\rm sa}$ we have

$$U^t f(x) = \langle \mathcal{U}^t(\tilde{T} f)\rangle_{\varrho_x} = \langle \tilde{T}_f\rangle_{\Psi^t(\varrho_x)}. \tag{39}$$

In summary, we have constructed a dynamically consistent embedding of the torus rotation from (5) into a quantum mechanical system on the RKHA $\mathfrak{A}$. For completeness, we note that the matrix elements of the state $\Psi^t(\rho_x)$ are given by

$$\begin{aligned}\langle \psi_i, \Psi^t(\varrho_x)\psi_j\rangle_{\mathfrak{A}} &= \langle U^t\psi_i, U^t\varrho_x\psi_j\rangle_{\mathfrak{A}} \\ &= e^{i(\omega_j-\omega_i)t}(\varrho_x)_{ij}.\end{aligned}$$

Using this formula together with the expressions for the matrix elements of $\tilde{T} f$ in (27), respectively, we arrive at the expression

$$U^t f = \sum_{i,j\in\mathbb{Z}^d} e^{i(\omega_j-\omega_i)t}\frac{\psi_i^*(x)\psi_j(x)(c_{i,j-i}+c_{j,i-j})\tilde{f}_{j-i}}{2\tilde{\kappa}},$$

which holds for all self-adjoint elements $f = \sum_{j\in\mathbb{Z}^d}\tilde{f}_j\psi_j \in \mathfrak{A}_{\rm sa}$.

Our discussion was thus far based on the RKHA $\mathfrak{A}$, as opposed to the RKHS $\mathcal{H}$. In Appendix A 4, we establish that the dynamics of classical states and observables can be represented consistently through their representatives on $\mathcal{H}$ using the maps ϖ and T in (29). Specifically, we show that for any $f \in \mathfrak{A}$,

$$U^t f(x) = \langle \mathcal{U}^t(\varpi f)\rangle_{\rho_x} = \langle \varpi f\rangle_{\Psi^t(\rho_x)},$$

while for any real-valued $f \in \mathfrak{A}_{\rm sa}$,

$$U^t f(x) = \langle \mathcal{U}^t(T f)\rangle_{\rho_x} = \langle T f\rangle_{\Psi^t(\rho_x)},$$

where $\rho_x = \mathcal{F}(x)$. In the above, $\mathcal{U}^t: B(\mathcal{H}) \to B(\mathcal{H})$ and $\Psi^t: Q(\mathcal{H}) \to Q(\mathcal{H})$ are evolution operators on quantum observables and states on $\mathcal{H}$, respectively, defined analogously to their counterparts on $\mathfrak{A}$ using the Koopman operator $U^t: \mathcal{H} \to \mathcal{H}$ (see Sec. IV A).

V. PROJECTION TO FINITE DIMENSIONS

While being dynamically consistent with the underlying classical evolution, the quantum system constructed in Sec. IV is infinite-dimensional, and thus not directly accessible to simulation by a quantum computer. We now describe an approach for projecting the infinite-dimensional quantum system to a finite-dimensional system. In Fig. 2 we refer to this level of representation as matrix mechanical, since all linear operators involved have finite rank and are representable by matrices. Our objectives are to construct this projection such that (1) it is refinable, i.e., the original quantum system is recovered in a limit of infinite dimension (number of qubits) and (2) it facilitates the eventual passage to the quantum computational level (to be described in Sec. VI).

We begin by fixing a positive integer parameter n (the number of qubits), chosen such that it is a multiple of the dimension d of the classical state space X, and defining the index sets

$$\begin{aligned}J_{n,d} &= \{-2^{n/d-1}, \ldots, -1, 1, \ldots, 2^{n/d-1}\}, \\ J_n &= \{(j_1, \ldots, j_d) \in \mathbb{Z}^d : j_i \in J_{n,d}\}.\end{aligned} \tag{40}$$

Note that J_n is a subset of J from (17) with $N \equiv 2^n$ elements. Next, consider the N-dimensional subspace of $\mathcal{H}$ given by

$$\mathcal{H}_n = \text{span}\{\psi_j : j \in J_n\},$$

and let $\Pi_n: \mathcal{H} \to \mathcal{H}$ be the orthogonal projection mapping into $\mathcal{H}_n$. When appropriate, we will interpret Π_n as a map into its range, i.e., $\Pi_n: \mathcal{H} \to \mathcal{H}_n$, without change of notation. The subspace $\mathcal{H}_n$ has the structure of an RKHS of dimension 2^n, associated with the spectrally truncated reproducing kernel

$$k_n(x, x') = \sum_{j\in J_n}\psi_j^*(x)\psi_j(x'). \tag{41}$$

Moreover, $\mathcal{H}_d, \mathcal{H}_{2d}, \mathcal{H}_{3d}, \ldots$ is a nested family of subspaces, increasing towards $\mathcal{H}$.

By virtue of being spanned by eigenfunctions of the generator V, $\mathcal{H}_n$ is invariant under the Koopman operator, i.e., $U^t\mathcal{H}_n = \mathcal{H}_n$ for all $t \in \mathbb{R}$. Moreover, the projection Π_n commutes with both V and U^t,

$$[V, \Pi_n] = 0, \quad [U^t, \Pi_n] = 0.$$

These invariance properties allow us to define a projected generator

$$V_n \equiv \Pi_n V \Pi_n, \tag{42}$$

and an associated Koopman operator

$$U_n^t := e^{tV_n} \equiv \Pi_n U^t \Pi_n,$$

such that the following diagram commutes for all $t \in \mathbb{R}$:

$$\begin{array}{ccc}\mathcal{H} & \xrightarrow{U^t} & \mathcal{H} \\ {\scriptstyle\Pi_n}\downarrow & & \downarrow{\scriptstyle\Pi_n} \\ \mathcal{H}_n & \xrightarrow{U_n^t} & \mathcal{H}_n\end{array}. \tag{43}$$

Similarly, we define a finite-rank Heisenberg operator

$$\mathcal{U}_n^t := \mathcal{U}^t\mathbf{\Pi}_n,$$

where $\mathbf{\Pi}_n: B(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is the projection on $B(\mathcal{H})$ defined as $\mathbf{\Pi}_n A = \Pi_n A \Pi_n$. This leads to an analogous commutative

diagram to that in (43), *viz.*,

$$\begin{array}{ccc} B(\mathcal{H}) & \xrightarrow{\mathcal{U}^t} & B(\mathcal{H}) \\ {\scriptstyle \mathbf{\Pi}_n}\downarrow & & \downarrow{\scriptstyle \mathbf{\Pi}_n} \\ B(\mathcal{H}_n) & \xrightarrow{\mathcal{U}_n^t} & B(\mathcal{H}_n) \end{array} . \tag{44}$$

Next, we introduce a spectrally truncated feature map $F_n: X \to \mathcal{H}_n$, defined analogously to (19) as

$$F_n(x) = k_{x,n} := k_n(x, \cdot),$$

as well as a corresponding quantum feature map $\mathcal{F}_n: X \to Q(\mathcal{H}_n)$, such that $\mathcal{F}_n(x) = \rho_{x,n}$ is given by

$$\rho_{x,n} = \langle \xi_{x,n}, \cdot \rangle_{\mathcal{H}_n} \xi_{x,n} \quad \text{with} \quad \xi_{x,n} = \frac{k_{x,n}}{\sqrt{\kappa_n}},$$
$$\kappa_n = k_n(x,x) = \sum_{j \in J_n} e^{-\tau |j|_p}. \tag{45}$$

In the sequel, we will use the states $\rho_{x,n}$ as approximations of the states $\rho_x = \mathcal{F}(x)$. These approximations have the following properties:

(1) The dynamical evolution of $\rho_{x,n}$ is governed by a finite-rank operator $\Psi_n^t: Q(\mathcal{H}_n) \to Q(\mathcal{H}_n)$, where

$$\Psi_n^t(\rho_{x,n}) = U_n^{t*} \rho_{x,n} U_n^t.$$

(2) As $n \to \infty$ (i.e., in the infinite qubit limit), $\Psi_n^t(\rho_{x,n})$ converges to $\Psi^t(\rho_x)$, in the sense that for any quantum mechanical observable $A \in B(\mathcal{H})$,

$$\langle A_n \rangle_{\rho_{x,n}} \xrightarrow{n \to \infty} \langle A \rangle_{\rho_x}, \tag{46}$$

where $A_n = \mathbf{\Pi}_n A$, and the convergence is uniform with respect to $x \in X$.

See Appendix A 5 for further details.

In light of the above, we employ the following approximations to the quantum mechanical representation of the evolution of classical observables from Sec. IV E (see also Appendix A 4),

$$\begin{aligned} \check{f}_n^{(t)}(x) &:= \langle \mathbf{\Pi}_n(\varpi f) \rangle_{\Psi_n^t(\rho_{x,n})}, \\ f_n^{(t)}(x) &:= \langle \mathbf{\Pi}_n(T f) \rangle_{\Psi_n^t(\rho_{x,n})}. \end{aligned} \tag{47}$$

By (46), for every function $f \in \mathfrak{A}$, $\check{f}_n^{(t)}(x)$ converges as $n \to \infty$ to $U^t f(x)$, uniformly with respect to the initial condition $x \in X$ and evolution time $t \in \mathbb{R}$, whereas $f_n^{(t)}(x)$ converges to $U^t f(x)$ if f is self-adjoint (real-valued).

VI. REPRESENTATION ON A QUANTUM COMPUTER

We are now ready to perform the final step in the QECD pipeline, namely passage from the matrix mechanical level to the quantum computational level associated with the n-qubit Hilbert space $\mathbb{B}_n = \mathbb{B}^{\otimes n}$ (see bottom row in Fig. 2). We will do so by applying a unitary map, so that the systems in the matrix mechanical and quantum computational levels are isomorphic as quantum systems. However, the key aspects that the quantum computational system provides are that (1) it can be efficiently implemented as a quantum circuit with a quadratic number of gates in n and (2) information about the evolution of classical observables can be extracted by measurement of the standard projection-valued measure associated with the computational basis. We describe the construction of the unitary map from the matrix mechanical to quantum computational levels and the properties of the resulting quantum system in Secs. VI A and VI B, respectively.

A. Quantum computational system on the tensor product Hilbert space

Being expressible in terms of finite-rank quantum states, observables, and evolution operators, the approximation framework described in Sec. V can be encoded in a quantum computing system operating on a finite-dimensional Hilbert space. In particular, letting $\mathbb{B} \simeq \mathbb{C}^2$ denote the two-dimensional Hilbert space associated with a single qubit, it follows immediately from the fact that $\mathcal{H}_n$ is a 2^n-dimensional Hilbert space that there exists a unitary map $W_n: \mathcal{H}_n \to \mathbb{B}_n$, where

$$\mathbb{B}_n := \mathbb{B}^{\otimes n} \simeq \underbrace{\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{n} \tag{48}$$

is the tensor product Hilbert space associated with n qubits. Under such a unitary, the projected generator V_n from (42) maps to a skew-adjoint operator $\hat{V}_n := W_n V_n W_n^*$, inducing a self-adjoint Hamiltonian

$$H_n := \frac{1}{i} \hat{V}_n, \tag{49}$$

and a corresponding unitary evolution operator $\hat{U}_n^t := e^{iH_n t}$ on $\mathbb{B}_n$. This leads to the commutative diagram

$$\begin{array}{ccc} \mathcal{H}_n & \xrightarrow{U_n^t} & \mathcal{H}_n \\ {\scriptstyle W_n}\downarrow & & \downarrow{\scriptstyle W_n} \\ \mathbb{B}_n & \xrightarrow{\hat{U}_n^t} & \mathbb{B}_n \end{array} ,$$

expressing the fact that elements of $\mathcal{H}_n$ and $\mathbb{B}_n$ evolve consistently under U_n^t and $\hat{U}_n^t$, respectively. Note that we work here with the self-adjoint Hamiltonian H_n as opposed to the skew-adjoint generator $\hat{V}_n$ for consistency with the usual convention in quantum mechanics.

In addition, W_n induces a unitary $\mathcal{W}_n: B(\mathcal{H}_n) \to B(\mathbb{B}_n)$, with $\mathcal{W}_n A = W_n A W_n^*$, mapping quantum mechanical observables on $\mathcal{H}_n$ to quantum mechanical observables on $\mathbb{B}_n$. The restriction of $\mathcal{W}_n$ on $Q(\mathcal{H}_n) \subset B(\mathcal{H}_n)$ then induces a continuous, invertible map $\mathcal{W}_n: Q(\mathcal{H}_n) \to Q(\mathbb{B}_n)$ from quantum states on $\mathcal{H}_n$ to quantum states on $\mathbb{B}_n$ (which we continue to denote using the symbol $\mathcal{W}_n$). Moreover, we have the evolution maps

$$\begin{aligned} \hat{\Psi}_n^t &: Q(\mathbb{B}_n) \to Q(\mathbb{B}_n): \hat{\rho}_n \mapsto \hat{\Psi}_n^t(\hat{\rho}_n) = \hat{U}_n^{t*} \hat{\rho}_n \hat{U}_n^t, \\ \hat{\mathcal{U}}_n^t &: B(\mathbb{B}_n) \to B(\mathbb{B}_n): \hat{A}_n \mapsto \hat{\mathcal{U}}_n^t \hat{A}_n = \hat{U}_n^t \hat{A}_n \hat{U}_n^{t*}, \end{aligned} \tag{50}$$

such that the maps for states and observables between and within the matrix mechanical and quantum computational level in Fig. 2 constitute commutative diagrams. In particular, following the vertical arrows in the left- and right-hand columns from the classical level to the quantum computational level gives the maps $\hat{\mathcal{F}}_n: X \to Q(\mathbb{B}_n)$ and $\hat{T}_n: \mathfrak{A} \to B(\mathbb{B}_n)$,

where

$$\hat{\mathcal{F}}_n = \mathcal{W}_n \circ \mathbf{\Pi}'_n \circ P \circ \delta, \quad \hat{T}_n = \mathcal{W}_n \circ \mathbf{\Pi}_n \circ T. \tag{51}$$

The maps $\hat{\mathcal{F}}_n$ and $\hat{T}_n$ provide the quantum computational representation of classical states and observables, respectively, which are two of the main ingredients of QECD (see Sec. II). By unitary equivalence, they have analogous convergence properties in the $n \to \infty$ limit as those of their matrix mechanical counterparts $\mathcal{F}_n$ and T_n, respectively, described in Sec. V. We also note that the evolution operator $\hat{U}_n^t$ at the quantum computational level can be equivalently obtained as a projection of the Koopman operator U^t on $\mathfrak{A}$, i.e.,

$$\hat{U}_n^t = (\mathcal{W}_n \circ \mathbf{\Pi}_n \circ \mathbf{\Pi}) U^t. \tag{52}$$

B. Factorizing the Hamiltonian in tensor product form

In order for the representation of the dynamics on $\mathbb{B}_n$ to exhibit robust quantum parallelism, i.e., implementation on a quantum circuit of small depth, it is highly beneficial that the Hamiltonian H_n can be decomposed as a sum of commuting operators in pure tensor product form, i.e.,

$$H_n = \sum_{j \in J_n} G_j = \sum_{j \in J_n} G_{1j} \otimes \cdots \otimes G_{nj}, \tag{53}$$

where $[G_i, G_j] = 0$ and $G_{lj} : \mathbb{B} \to \mathbb{B}$ are mutually commuting, single-qubit Hamiltonians. With such a decomposition, the unitary operator $\hat{U}_n^t = e^{iH_n t}$ generated by H_n factorizes as

$$\hat{U}_n^t = \exp\left(i \sum_{j \in J_n} G_j t\right) = \prod_{j \in J_n} \exp\left(\bigotimes_{l=1}^{n} i G_{lj} t\right). \tag{54}$$

Thus, $\hat{U}_n^t$ can be split into a composition of up to 2^n unitaries $\exp(iG_j t)$ [depending on the number of nonzero terms G_j in the right-hand side of (53)], which can be applied in any order by commutativity of the G_j. Moreover, each unitary $\exp(iG_j t)$ has a generator of pure tensor product form, and thus can be represented as a quantum circuit with at most n quantum gates for rotations of the individual qubits.

In fact, as we will now show, using a Walsh operator representation [56], for a dynamical system with pure point spectrum the decomposition in (53) has only n nonzero terms G_j, and for each nonzero term, the tensor product factorization $G_j = \bigotimes_{l=1}^{n} G_{lj}$ has all but one factors G_{lj} equal to the identity. As a result,

$$\exp\left(\bigotimes_{l=1}^{n} i G_{lj} t\right) = \bigotimes_{l=1}^{n} \exp(i G_{lj} t),$$

and the decomposition in (54) reduces to a tensor product of n unitaries,

$$\hat{U}_n^t = \bigotimes_{l=1}^{n} \Xi_l^t \quad \text{with} \quad \Xi_l^t = \exp\left(i \sum_{j \in J_n} G_{lj} t\right). \tag{55}$$

The key point about (55) is that $\hat{U}_n^t$ can be implemented via a quantum circuit of n qubit channels with no cross-channel communication. We will return to this point in Sec. VII.

1. Walsh-Fourier transform and Walsh operators

Classical states and observables of the dynamical system have been transformed into pure state density operators and self-adjoint operators on the 2^n-dimensional Hilbert space $\mathbb{B}_n$ which is a tensor product of the single qubit quantum state spaces as given in (48). We will employ the commonly used Dirac bra-ket notation [5] to denote vectors in $\mathbb{B}_n$. We let $\{|0\rangle, |1\rangle\}$ be the standard orthonormal basis of the single-qubit Hilbert space $\mathbb{B} \simeq \mathbb{C}^2$ comprising of eigenvectors of the Pauli Z operator,

$$Z|0\rangle = |0\rangle, \quad Z|1\rangle = -|1\rangle,$$

with

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Thus, each vector $|\psi\rangle \in \mathbb{B}$ can be expanded in this basis as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad \text{with} \quad \alpha, \beta \in \mathbb{C}.$$

In order to arrive at the decomposition in (55), we employ the approach developed in Ref. [56], which is based on discrete Walsh-Fourier transforms, and the associated Walsh operators, as follows. First, for any nonnegative integer $j \in \mathbb{N}_0$, we let $\beta(j) = (\beta_1(j), \ldots, \beta_l(j)) \in \{0,1\}^l$ be its binary expansion; that is,

$$j = \sum_{i=1}^{l} \beta_i(j) 2^{i-1} = \beta_1(j)2^0 + \beta_2(j)2^1 + \cdots + \beta_l(j)2^l,$$

where $l \in \mathbb{N}$ is the smallest positive integer such that $j \leqslant 2^l - 1$. For example, we have $\beta(0) = 0$, $\beta(1) = 1$, $\beta(2) = (0,1)$, $\beta(3) = (1,1)$, and $\beta(4) = (0,0,1)$. Moreover, for every real number $u \in [0,1)$ we let $\gamma(u) = (\gamma_1(u), \gamma_2(u), \ldots) \in \{0,1\}^{\mathbb{N}}$ be its dyadic expansion, i.e.,

$$u = \sum_{i=1}^{\infty} \gamma_i(u) 2^{-i} = \frac{\gamma_1(u)}{2} + \frac{\gamma_2(u)}{4} + \frac{\gamma_3(u)}{8} + \cdots.$$

Note that the most significant digit in $\beta(j)$ is the last one, $\beta_l(j)$, whereas the most significant digit in $\gamma(u)$ is the first one, $\gamma_1(u)$.

With this notation, for every $j \in \mathbb{N}_0$ we define the Walsh function $w_j : [0,1) \to \{0,1\}$ as

$$w_j(u) = (-1)^{\sum_{i=1}^{l} \beta_i(j)\gamma_i(u)}.$$

Furthermore, for any $n \in \mathbb{N}_0$ and $j \in \{0, \ldots, 2^n - 1\}$, we define the discrete Walsh function of order n, $w_j^{(n)} : \{0, \ldots, 2^n - 1\} \to \{0,1\}$ as

$$w_j^{(n)}(m) = w_j(m/2^n), \quad \text{with} \quad m = 0, \ldots, 2^n - 1.$$

It then follows that

$$w_j^{(n)}(m) = (-1)^{\sum_{i=1}^{l} \beta_i(j)\gamma_i(m/2^n)} = (-1)^{\sum_{i=1}^{n} \beta_i^{(n)}(j)\tilde{\beta}_i^{(n)}(m)}.$$

Here, $\beta^{(n)}(j) = (\beta_1, \ldots, \beta_l, 0, \ldots, 0) \in \{0,1\}^n$ is the n-digit binary expansion of j obtained by padding $\beta(j)$ to the right with zeros, as needed. Moreover,

$$\tilde{\beta}^{(n)}(j) = \left(\tilde{\beta}_1^{(n)}(j), \ldots, \tilde{\beta}_n^{(n)}(j)\right) = (\gamma_1(j/2^m), \ldots, \gamma_n(j/2^m))$$

is the n-digit reversed binary representation of m. Thus, the exponent in the expression for $w_j^{(n)}(m)$ is given by the inner

product between the n-digit binary expansion of j with the bit-reversed binary expansion of m. For example, with $n = 2$ and $m = 0, 1, 2, 3$, we have $w_0^{(2)}(m) = \{1, 1, 1, 1\}$, $w_1^{(2)}(m) = \{1, 1, -1, -1\}$, $w_2^{(2)}(m) = \{1, -1, 1, -1\}$, and $w_3^{(2)}(m) = \{1, -1, -1, 1\}$.

Among the Walsh functions w_j, those with $j = 1, 2, 4, \ldots, 2^l$ for $l \in \mathbb{N}_0$ are called Rademacher functions, R_l, and satisfy

$$w_{2^l}(u) \equiv R_l(u) = (-1)^{\gamma_l(u)}. \tag{56}$$

That is, $R_l(u)$ depends only on the $(l+1)$-th bit in the dyadic expansion of u. Using (56), it follows that for any integer $m \in \{0, \ldots, 2^n - 1\}$, we have

$$\frac{m}{2^n} = \sum_{i=0}^{n-1} \frac{1 - R_{i+1}(m/2^n)}{2^{i+2}},$$

meaning that we can express the ith bit in the dyadic decomposition of $m/2^n$ in terms of the $(i-1)$-th Rademacher function,

$$\gamma_i(m/2^n) = \frac{1 - R_{i-1}(m/2^n)}{2}. \tag{57}$$

It is known that the set $\{w_j\}_{j\in\mathbb{N}_0}$ forms an orthonormal basis of the Hilbert space $L^2([0,1])$ with respect to Lebesgue measure. In the discrete case, we let $L_n^2([0,1])$ be the N-dimensional Hilbert space, $N \equiv 2^n$, with respect to the normalized counting measure supported on $\{0, 1/N, 2/N, \ldots, (N-1)/N\}$. Then, the set of discrete Walsh functions of order n, $\{w_j^{(n)}\}_{j=0}^{N-1}$, is an orthonormal basis of $L_n^2([0,1])$. One obtains

$$f = \sum_{j=0}^{N-1} \hat{f}_j w_j^{(n)} \in L_n^2([0,1]) \quad \text{with}$$

$$\hat{f}_j = \frac{1}{N} \sum_{m=0}^{N-1} w_j^{(n)}(m) f(m/N).$$

The map $\mathsf{F}_n : L_n^2([0,1]) \to \mathbb{C}^N : f \mapsto (\hat{f}_0, \ldots, \hat{f}_{N-1})$ is called the discrete Walsh-Fourier transform of the function $f \in L_n^2([0,1])$.

Next, consider the tensor product basis $\{|\boldsymbol{b}\rangle = |b_1\rangle \otimes \cdots \otimes |b_n\rangle\}$ of $\mathbb{B}_n$ with $|b_i\rangle \in \{|0\rangle, |1\rangle\}$, where the multi-index $\boldsymbol{b} = (b_1, \ldots, b_n) \in \{0,1\}^n$ runs over all binary strings of length n. Whenever convenient, we will employ the notation $|b\rangle \equiv |\boldsymbol{b}\rangle$, where $\boldsymbol{b} = \tilde{\beta}^{(n)}(b)$. That is, b is an integer in the range $0, \ldots, 2^n - 1$, whose reversed binary representation is equal to $\boldsymbol{b}$,

$$b = \sum_{i=1}^{n} \tilde{\beta}_i^{(n)}(b) 2^{n-i} = \sum_{i=1}^{n} b_i 2^{n-i}.$$

For example, in a system with $n = 3$ qubits $|b\rangle = |6\rangle$ corresponds to $|\boldsymbol{b}\rangle = |110\rangle$, where the least significant bit is the one to the right. Note that $\{|b\rangle\}_{b=0}^{2^n-1}$ is also the standard quantum computational basis for an n-qubit problem in the Qiskit framework [61] that we will employ in Sec. X [78].

For every $\boldsymbol{b} \in \{0,1\}^n$, we define the associated Walsh operator $Z_{\boldsymbol{b}} : \mathbb{B}_n \to \mathbb{B}_n$ as

$$Z_{\boldsymbol{b}} = Z^{b_1} \otimes Z^{b_2} \otimes \cdots \otimes Z^{b_n}.$$

By construction, the $Z_{\boldsymbol{b}}$ form a collection of mutually commuting, self-adjoint operators, which have pure tensor product form and are diagonal in the $\{|\boldsymbol{b}\rangle\}$ basis of $\mathbb{B}_n$, i.e.,

$$Z_{\boldsymbol{b}}|\boldsymbol{c}\rangle = \left(\prod_{i=1}^{n} (1 - 2c_i)^{b_i}\right)|\boldsymbol{c}\rangle,$$

where $|\boldsymbol{c}\rangle$ is again a quantum computational basis vector. For example, for $n = 2$ qubits, the Walsh operator $Z_{\boldsymbol{b}}$ with $\boldsymbol{b} = |01\rangle$, and the basis vector $|\boldsymbol{c}\rangle = |\boldsymbol{b}\rangle$, one obtains

$$Z_{\boldsymbol{b}}|\boldsymbol{c}\rangle = (I \otimes Z)(|0\rangle \otimes |1\rangle) = -|\boldsymbol{c}\rangle.$$

It follows from a counting argument that the collection $\{Z_{\boldsymbol{b}}\}_{\boldsymbol{b}\in\{0,1\}^n}$ forms a basis of the vector space of operators in $B(\mathbb{B}_n)$ which are diagonal in the $\{|\boldsymbol{b}\rangle\}$ basis. In Ref. [56], it was shown that if $A \in B(\mathbb{B}_n)$ is such a diagonal operator,

$$A|\boldsymbol{b}\rangle = a_{\boldsymbol{b}}|\boldsymbol{b}\rangle \quad \text{with} \quad a_{\boldsymbol{b}} \in \mathbb{C},$$

then it admits the expansion

$$A = \sum_{j=0}^{N-1} \hat{f}_j Z_{\beta^{(n)}(j)}, \quad \hat{f}_j \in \mathbb{C}, \tag{58}$$

where the expansion coefficients $\hat{f}_j$ are the complex Walsh-Fourier coefficients $(\hat{f}_0, \ldots, \hat{f}_{N-1}) = \mathsf{F}_n f$ of the function $f = (f_0, \ldots, f_{N-1}) \in L_n^2([0,1])$ with $f_j = a_{\beta^{(n)}(j)}$. That is, f_j is equal to the eigenvalue $a_{\boldsymbol{b}}$, where $\boldsymbol{b}$ is the n-digit binary representation of the integer j.

2. *Walsh representation of the Hamiltonian*

In order to effect the decomposition in (53) for the generator-induced Hamiltonian from (49), let $o : J_{n,d} \to \{0, \ldots, 2^{n/d} - 1\}$ be the enumeration on the index set $J_{n,d}$ from (40) based on the standard order of integers, i.e., $o(-2^{n/d-1}) = 0, 1, \ldots, 2^{n/d} - 1 = o(2^{n/d-1})$. For example, $n = 2$, $d = 1$ gives $j \in J_{2,1} = \{-2, -1, 1, 2\}$, which is mapped to $o(j) = \{0, 1, 2, 3\}$. The mapping of $J_{2,2}$ (for $n = 2$, $d = 2$) is displayed in first and third columns of Table II. We define $W_n : \mathcal{H}_n \to \mathbb{B}_n$ as the unique (unitary) linear map such that for $j = (j_1, \ldots, j_d)$,

$$W_n \psi_j = |\boldsymbol{b}\rangle \quad \text{with} \quad \boldsymbol{b} = \boldsymbol{\eta}(j) := (\eta(j_1), \ldots, \eta(j_d)),$$
$$\eta(j_i) = \tilde{\beta}^{(n/d)}(o(j_i)). \tag{59}$$

That is, W_n maps the basis element ψ_j of $\mathcal{H}_n$ with multi-index $j = (j_1, \ldots, j_d) \in J_n$ to the tensor product basis element $|\boldsymbol{b}\rangle$, with $\boldsymbol{b}$ given by an invertible binary string encoding of j. Here, $\boldsymbol{b}$ is obtained as a concatenation $(\eta(j_1), \ldots, \eta(j_d))$ of d binary strings of length n/d, corresponding to the dyadic decompositions of $o(j_1), \ldots, o(j_d)$, respectively. See again Table II, where we list the mapping for a two-dimensional torus with $2 = n/d$ qubits for each torus dimension.

Since $\hat{V}_n \psi_j = i\omega_j \psi_j$ with ω_j given by (14), we have

$$H_n|\boldsymbol{b}\rangle = \omega_{\boldsymbol{\eta}^{-1}(\boldsymbol{b})}|\boldsymbol{b}\rangle, \tag{60}$$

Thus, in order to decompose H_n into Walsh operators as in (58), we need to compute the discrete Walsh transform of the function $h \in L_n^2([0,1])$ with

$$h(m/N) = \omega_j, \quad j = \boldsymbol{\eta}^{-1}(\tilde{\beta}^{(n)}(m)). \tag{61}$$

TABLE II. Binary encodings $\boldsymbol{\eta}(j) = (\eta(j_1), \eta(j_2))$ and enumeration $b = (\tilde{\beta}^{(n)})^{-1}(\boldsymbol{\eta}(j)) = 0, \ldots, 2^n - 1$ of the eigenfrequencies ω_j with multi-index $j = (j_1, j_2)$ of a quasiperiodic system on a two-dimensional torus ($d = 2$) with basis frequencies α_1 and α_2. The total number of qubits is $n = 4$ with two qubits for each torus dimension.

(j_1, j_2)	$(\eta(j_1), \eta(j_2))$	b	ω_j
$(-2, -2)$	$((0,0),(0,0))$	0	$-2\alpha_1 - 2\alpha_2$
$(-2, -1)$	$((0,0),(0,1))$	1	$-2\alpha_1 - 1\alpha_2$
$(-2, +1)$	$((0,0),(1,0))$	2	$-2\alpha_1 + 1\alpha_2$
$(-2, +2)$	$((0,0),(1,1))$	3	$-2\alpha_1 + 2\alpha_2$
$(-1, -2)$	$((0,1),(0,0))$	4	$-1\alpha_1 - 2\alpha_2$
$(-1, -1)$	$((0,1),(0,1))$	5	$-1\alpha_1 - 1\alpha_2$
$(-1, +1)$	$((0,1),(1,0))$	6	$-1\alpha_1 + 1\alpha_2$
$(-1, +2)$	$((0,1),(1,1))$	7	$-1\alpha_1 + 2\alpha_2$
$(+1, -2)$	$((1,0),(0,0))$	8	$+1\alpha_1 - 2\alpha_2$
$(+1, -1)$	$((1,0),(0,1))$	9	$+1\alpha_1 - 1\alpha_2$
$(+1, +1)$	$((1,0),(1,0))$	10	$+1\alpha_1 + 1\alpha_2$
$(+1, +2)$	$((1,0),(1,1))$	11	$+1\alpha_1 + 2\alpha_2$
$(+2, -2)$	$((1,1),(0,0))$	12	$+2\alpha_1 - 2\alpha_2$
$(+2, -1)$	$((1,1),(0,1))$	13	$+2\alpha_1 - 1\alpha_2$
$(+2, +1)$	$((1,1),(1,0))$	14	$+2\alpha_1 + 1\alpha_2$
$(+2, +2)$	$((1,1),(1,1))$	15	$+2\alpha_1 + 2\alpha_2$

This calculation is detailed in Appendix B. The eigenvalues ω_j for the example of a two-dimensional torus with $n = 2$ qubits are listed in the fourth column of Table II.

By virtue of the decomposition in (B1), the only nonzero coefficients in the Walsh-Fourier transform $\hat{h} = (\hat{h}_0, \ldots, \hat{h}_{N-1}) = \mathsf{F}_n h$ are the coefficients $\hat{h}_j$ with $j = 2^{l+(i-1)d}$ and $1 \leqslant l \leqslant n/d$, $1 \leqslant i \leqslant d$. Correspondingly, the only nonzero terms $\hat{h}_j Z_{\beta^{(n)}(j)}$ in the Walsh operator expansion from (58) for the Hamiltonian in (60) are those for which the binary string $\boldsymbol{\eta}(j)$ has exactly one bit equal to 1 and the remaining $n - 1$ bits equal to 0. In particular, we have

$$\begin{aligned} H_n &= \sum_{i=1}^{d} \sum_{l=0}^{\frac{n}{d}-1} \hat{h}_{2^{l+(i-1)n/d}} Z_{\beta(2^{l+(i-1)n/d})} \\ &= \hat{h}_1 Z \otimes I \otimes I \otimes \cdots \otimes I \\ &\quad + \hat{h}_2 I \otimes Z \otimes I \otimes I \otimes \cdots \otimes I \\ &\quad + \hat{h}_4 I \otimes I \otimes Z \otimes I \otimes \cdots \otimes I + \cdots \\ &\quad + \hat{h}_{2^{n-1}} I \otimes \cdots \otimes I \otimes Z, \end{aligned} \tag{62}$$

where the coefficients $\hat{h}_1, \hat{h}_2, \hat{h}_4, \ldots, \hat{h}_{2^{n-1}}$ have closed-form expressions; see (B7).

Equation (62) verifies the assertion made earlier that the decomposition of H_n in (53) can be arranged to have n nonzero terms, each of which factorizes as a tensor product of n operators, with all but one factors equal to the identity. Since

$$e^{itI\otimes\cdots\otimes I\otimes Z\otimes I\cdots\otimes I} = I \otimes \cdots \otimes I \otimes e^{itZ} \otimes I \otimes \cdots \otimes I,$$

we conclude that

$$\hat{U}_n^t = e^{iH_n t} = \bigotimes_{l=0}^{n-1} \exp(it\hat{h}_{2^l} Z), \tag{63}$$

which is consistent with the decomposition in (55).

3. Cost of circuit construction

Before closing this section, we briefly assess the computational cost of constructing a quantum circuit that implements the unitary Koopman evolution $\hat{U}_n^t$. Note that this is a one-off cost that does not affect the running time of our algorithm.

According to (55), $\hat{U}_n^t$ is decomposed as a tensor product of n single-qubit unitaries $\exp(it\hat{h}_{2^l} Z)$ with $l = 0, \ldots, n-1$. Each such unitary can be implemented by an R_z gate that rotates by the angle $\vartheta_l(t) = 2\alpha t \hat{h}_{2^l}$. That is, we have

$$\exp(-it\hat{h}_l Z) = \mathsf{R}_z(\vartheta_l(t)), \tag{64}$$

where

$$\mathsf{R}_z(\vartheta) = e^{-i\vartheta Z/2} = \begin{pmatrix} e^{-i\vartheta/2} & 0 \\ 0 & e^{i\vartheta/2} \end{pmatrix} \tag{65}$$

is the matrix representation of the R_z rotation gate by angle ϑ in the $\{|0\rangle, |1\rangle\}$ basis of $\mathbb{B}$. Using the closed-form expressions for $\hat{h}_{2^l}$ in (B7), we can evaluate each rotation angle $\vartheta_l(t)$ at $O(1)$ cost, so the computational cost of building our n-qubit Koopman evolution circuit for time t is $O(n)$.

VII. PROJECTIVE MEASUREMENT OF OBSERVABLES

In the classical setting, the process of obtaining the results of a computation is a straightforward readout of the state of the computer. In contrast, in quantum computing, extracting information from the system is a nontrivial process, as it must invariably confront with the intricacies of quantum measurement. In this section, we describe how the QECD performs probabilistic predictions of the evolution of classical observables through projective measurement of quantum computational observables. First, in Sec. VII A, we consider an idealized measurement scenario, where one has access to the spectral measure of the observable of interest. Then, in Sec. VII B we develop an approximate measurement procedure based on the QFT, which yields asymptotically consistent results with the idealized measurement, while maintaining an exponential quantum advantage. Additional technical results are provided in Appendix C.

A. Idealized quantum measurement

Our goal is to approximate the classical evolution $U^t f(x)$ through projective measurement of the quantum mechanical observable $\hat{S}_n := \hat{T}_n f$ on the quantum state

$$\hat{\rho}_{x,n}^{(t)} := \hat{\Psi}_n^t(\hat{\rho}_{x,n}), \quad \hat{\rho}_{x,n} = \hat{\mathcal{F}}_n(x), \tag{66}$$

where the representation maps $\hat{T}_n$ and $\hat{\mathcal{F}}_n$ are defined in (51), and the evolution map $\hat{\Psi}_n^t$ is defined in (50) (see also Fig. 2). Since $\hat{S}_n$ is a finite-rank, self-adjoint operator, it has a spectral resolution

$$\hat{S}_n = \sum_{s \in \sigma(\hat{S}_n)} s P_s, \tag{67}$$

where $\sigma(\hat{S}_n)$ is the spectrum of $\hat{S}_n$, i.e., the set of its eigenvalues, and $P_s \in B(\mathbb{B}_n)$ are the orthogonal projections onto the corresponding eigenspaces. For example, if $s \in \sigma(\hat{S}_n)$ is an eigenvalue of multiplicity 1 with a corresponding normalized eigenvector $|s\rangle$, then P_s is the rank-1 projection given by

$P_s = |s\rangle\langle s|$. The collection $\{P_s\}$ defines a projection-valued measure (PVM) on $\sigma(\hat{S}_n)$, i.e., a map $\mathcal{S}_n\colon \Sigma(\hat{S}_n) \to B(\mathbb{B}_n)$ given by

$$\mathcal{S}_n(\Upsilon) = \sum_{s\in\Upsilon} P_s, \tag{68}$$

where $\Sigma(\hat{S}_n)$ is the collection (σ-algebra) of all subsets of $\sigma(\hat{S}_n)$, and Υ a set in $\Sigma(\hat{S}_n)$. A projective measurement of $\hat{S}_n$ on the quantum state $\hat{\rho}_{x,n}^{(t)}$ then corresponds to a randomly drawn eigenvalue $\hat{s}$ from the spectrum $\sigma(\hat{S}_n)$ with probability

$$\mathbb{P}_{\hat{\rho}_{x,n}^{(t)}}(\hat{s}) = \mathrm{tr}\big(\hat{\rho}_{x,n}^{(t)} P_s\big).$$

The random draws $\hat{s}$ have expectation

$$\sum_{s\in\sigma(\hat{S}_n)} s\mathbb{P}_{\hat{\rho}_{x,n}^{(t)}}(s) = \sum_{s\in\sigma(\hat{S}_n)} \mathrm{tr}\big(\hat{\rho}_{x,n}^{(t)} P_s\big) =: f_n^{(t)}(x),$$

which is equivalent to (47) by unitarity of the transformations from the matrix mechanical to quantum computational level.

One can compute a Monte Carlo (ensemble) estimate of $f_n^{(t)}(x)$ by performing a collection $\{\hat{s}_1, \ldots, \hat{s}_K\}$ of measurements of $\hat{S}_n$ on K independently and identically prepared quantum systems. The number K is oftentimes referred to as the number of shots. The ensemble mean,

$$\hat{f}_n^{(t)}(x) := \frac{1}{K}\sum_{k=1}^{K} \hat{s}_k \tag{69}$$

converges as $K \to \infty$ to the expectation $f_n^{(t)}(x)$. The latter, converges in turn to the true value $U^t f(x)$ in the infinite-qubit limit, $n \to \infty$; that is, we have

$$\lim_{n\to\infty}\lim_{K\to\infty} \hat{f}_n^{(t)}(x) = U^t f(x). \tag{70}$$

B. Approximate quantum measurement using quantum Fourier transforms

Despite its theoretical consistency, the quantum measurement process described in Sec. VII A is not well suited for practical quantum computation. The reason is that, in general, a quantum computing platform does not support the measurement of arbitrary PVMs such as $\mathcal{S}_n$ in (68), and instead allows only measurement of the PVM associated with the quantum register. For an n-qubit system, the latter is defined as the PVM $\mathcal{E}_n\colon \Sigma(\{0,1\}^n) \to B(\mathbb{B}_n)$,

$$\mathcal{E}_n(\Upsilon) = \sum_{\boldsymbol{b}\in\Upsilon} E_{\boldsymbol{b}} \quad \text{with} \quad E_{\boldsymbol{b}} = |\boldsymbol{b}\rangle\langle\boldsymbol{b}|,$$

where $E_{\boldsymbol{b}}$ is the orthogonal projection along the computational basis vector $|\boldsymbol{b}\rangle$.

In order to transform a measurement of $\mathcal{S}_n$ to an equivalent measurement of $\mathcal{E}_n$, one must apply a unitary transformation $\hat{\rho}_{x,n}^{(t)} \mapsto \Lambda_n \hat{\rho}_{x,n}^{(t)} \Lambda_n^*$ to the quantum state $\hat{\rho}_{x,n}^{(t)}$, where $\Lambda_n\colon \mathbb{B}_n \to \mathbb{B}_n$ is a unitary map that diagonalizes $\hat{S}_n$, i.e., $\Lambda_n^* \hat{S}_n \Lambda_n$ is a diagonal operator in the $\{|\boldsymbol{b}\rangle\}$ basis of $\mathbb{B}_n$. Two issues arise with this approach. First, Λ_n is generally not known in closed form, and must be determined by solving an (exponentially large) eigenvalue problem for $\hat{S}_n$. Second, even if Λ_n were known explicitly, it would likely be difficult to implement efficiently in a quantum circuit as it would generally be represented by a fully occupied matrix.

To overcome these challenges, instead of working with Λ_n directly, we will employ a different unitary map on $\mathbb{B}_n$ associated with the QFT. As is well known, the QFT on the n-qubit space $\mathbb{B}_n$ has a circuit implementation of size $O(n^2)$ and depth $O(n)$ [5,59,60]. Thus, including it in the QECD pipeline does not result in loss of an exponential advantage in n over deterministic classical computation. Crucially for our purposes, moreover, the class of operators $\hat{S}_n$ induced from multiplication operators πf by classical observables on X turns out to be approximately diagonalized by the QFT, with an error that vanishes in a suitable asymptotic limit.

In more detail, for any $n \in \mathbb{N}$, let $\mathfrak{F}_n\colon \mathbb{B}_n \to \mathbb{B}_n$ be the Fourier operator on $\mathbb{B}_n$, defined as

$$\mathfrak{F}_n|m\rangle = \frac{1}{\sqrt{2^n}}\sum_{p=0}^{2^n-1} e^{-2\pi i p m/2^n}|p\rangle, \tag{71}$$

where $|m\rangle$ and $|p\rangle$ are again two basis vectors of $\mathbb{B}_n$, parameterized by integers m and p, respectively, by conversion of the corresponding binary sequences. Moreover, for n divisible by the state space dimension d, let $\mathfrak{F}_{n,d}\colon \mathbb{B}_n \to \mathbb{B}_n$ be the tensor product operator defined as

$$\mathfrak{F}_{n,d} = \underbrace{\mathfrak{F}_{n/d} \otimes \cdots \otimes \mathfrak{F}_{n/d}}_{d}, \tag{72}$$

and $\mathfrak{F}_{n,d}\colon B(\mathbb{B}_n) \to B(\mathbb{B}_n)$ the induced operator on quantum computational observables, given by

$$\mathfrak{F}_{n,d} A = \mathfrak{F}_{n,d} A \mathfrak{F}_{n,d}^*. \tag{73}$$

In Appendix C, we show that $\tilde{S}_n := \mathfrak{F}_{n,d}\hat{S}_n\mathfrak{F}_{n,d}^*$ is an approximately diagonal operator in the computational basis $\{|\boldsymbol{b}\rangle\}$. In particular, decomposing $\boldsymbol{b} = (\boldsymbol{b}^{(1)}, \ldots, \boldsymbol{b}^{(d)})$, where $\boldsymbol{b}^{(i)} = (b_1^{(i)}, \ldots, b_{n/d}^{(i)})$ are binary strings of length n/d, and defining the points

$$x_{\boldsymbol{b}} = (\theta_{\boldsymbol{b}^{(1)}}, \ldots, \theta_{\boldsymbol{b}^{(d)}}) \in \mathbb{T}^d \tag{74}$$

with the canonical angle coordinates

$$\theta_{\boldsymbol{b}^{(i)}} = \frac{2\pi(\tilde{\beta}^{(n/d)})^{-1}(\boldsymbol{b}^{(i)})}{2^{n/d}},$$

we have

$$\tilde{S}_n|\boldsymbol{b}\rangle = \tilde{s}_{\boldsymbol{b}}|\boldsymbol{b}\rangle + |r_{n\boldsymbol{b}}\rangle, \quad \tilde{s}_{\boldsymbol{b}} = f(x_{\boldsymbol{b}}). \tag{75}$$

Here, $|r_{n\boldsymbol{b}}\rangle$ is a residual that vanishes as $n \to \infty$. Effectively, the points $x_{\boldsymbol{b}}$ define a uniform grid on the d-torus $\mathbb{T}^d$, indexed by the n-digit binary strings $\boldsymbol{b}$. The quantities $\tilde{s}_{\boldsymbol{b}}$ can thus be interpreted as approximate eigenvalues of $\tilde{S}_n$, which can be obtained from classical measurement of f at the points $x_{\boldsymbol{b}}$, avoiding the need to solve an exponentially large eigenvalue problem for $\hat{S}_n$.

By virtue of these facts, and since

$$\mathrm{tr}\big(\hat{\rho}_{x,n}^{(t)}\hat{S}_n\big) = \mathrm{tr}\big(\tilde{\rho}_{x,n}^{(t)}\tilde{S}_n\big),$$

with

$$\tilde{\rho}_{x,n}^{(t)} = \mathfrak{F}_{n,d}\hat{\rho}_{x,n}^{(t)}, \tag{76}$$

we can approximate a measurement of $\hat{S}_n$ on the state $\hat{\rho}^{(t)}_{x,n}$ by a measurement of the PVM $\mathcal{E}_n$ on the state $\tilde{\rho}^{(t)}_{x,n}$. The latter measurement returns a random string $\boldsymbol{b} \in \{0,1\}^n$ with probability

$$\mathbb{P}_{\tilde{\rho}^{(t)}_{x,n}}(\boldsymbol{b}) = \mathrm{tr}\big(\tilde{\rho}^{(t)}_{x,n} E_{\boldsymbol{b}}\big) = \langle \boldsymbol{b}|\tilde{\rho}^{(t)}_{x,n}|\boldsymbol{b}\rangle,$$

inducing a sample $\tilde{s}_{\boldsymbol{b}} = f(x_{\boldsymbol{b}})$. Analogously to (69), we estimate $U^t f(x)$ by forming an ensemble of K independent measurements $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_K$ of $\mathcal{E}_n$, and computing the ensemble mean by

$$\hat{f}^{(t)}_n(x) := \frac{1}{K}\sum_{k=1}^{K} \tilde{s}_{\boldsymbol{b}_k}. \tag{77}$$

Further details on this approximation, such as the proof of asymptotic consistency, can be found in Appendix C. Here, we note that due to errors associated with the QFT-based measurement process, the convergence of $\hat{f}^{(t)}_n$ to $U^t f$ is not unconditional, but requires taking a sequence of decreasing RKHA parameters τ [unlike the limit in (70) which holds for any $\tau > 0$]. It should also be noted that it is possible to simulate multiple classical observables using the same circuit and ensemble of quantum measurements $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_K\}$. That is, to simulate the evolution of a different observable $g: X \to \mathbb{C}$, we use the $\boldsymbol{b}_k$ to generate samples $\tilde{\tilde{s}}_{\boldsymbol{b}_k} = g(x_{\boldsymbol{b}_k})$, and estimate $U^t g(x)$ by $\hat{g}^{(t)}_n(x) := \sum_{k=1}^{K} \tilde{\tilde{s}}_{\boldsymbol{b}_k}/K$, analogously to (77).

VIII. STATE PREPARATION

Besides measurement of observables, the preparation, or loading, of the quantum state representing the input (initial conditions) to a quantum computer is challenging. In a typical scenario involving an n-qubit computation, the register of a quantum computer is initialized with a state vector associated with an unentangled tensor product state,

$$|\mathbf{0}\rangle \equiv |0\rangle^{\otimes n}. \tag{78}$$

The desired initial state must be prepared by applying a unitary transformation (encoder) to $|\mathbf{0}\rangle$, which may in general require a circuit of exponential depth in n when the algorithm is broken down to elementary gate operations [57,58]. This poses a potentially significant obstruction to the scalability of quantum computational algorithms.

In QECD, our task is to prepare the quantum state $\hat{\rho}_{x,n} = \hat{\mathcal{F}}_n(x)$ from (66) associated with the classical initial condition $x \in X$. This state is a pure state,

$$\hat{\rho}_{x,n} = |\hat{\xi}_{x,n}\rangle\langle\hat{\xi}_{x,n}|,$$

where the state vector $|\hat{\xi}_{x,n}\rangle = W_n \xi_{x,n}$ is obtained by application of the unitary $W_n: \mathcal{H}_n \to \mathbb{B}_n$ from (59) on the normalized RKHS feature vector $\xi_{x,n}$ from (45). Specifically, we have

$$\xi_{x,n} = \frac{k_{x,n}}{\sqrt{\kappa_n}} = \frac{1}{\sqrt{\kappa_n}}\sum_{j\in J_n} \psi^*_j(x)\psi_j,$$

and thus

$$|\hat{\xi}_{x,n}\rangle = W_n\xi_{x,n} = \sum_{j\in J_n}\frac{\psi^*_j(x)}{\sqrt{\kappa_n}} W_n\psi_j = \sum_{\boldsymbol{b}\in\{0,1\}^n}\frac{\psi^*_{\boldsymbol{\eta}^{-1}(\boldsymbol{b})}(x)}{\sqrt{\kappa_n}}|\boldsymbol{b}\rangle. \tag{79}$$

We now describe how, in the limit of small RKHA parameter τ, this state can be prepared to any degree of accuracy using a circuit of size $O(n)$ and depth $O(1)$.

Let $|\Omega\rangle \in \mathbb{B}_n$ be the state vector associated with a uniform superposition of the quantum computational basis vectors,

$$|\Omega\rangle = \frac{1}{\sqrt{N}}\sum_{\boldsymbol{b}\in\{0,1\}^n}|\boldsymbol{b}\rangle.$$

The state vector $|\Omega\rangle$ can be prepared from $|\mathbf{0}\rangle$ using a circuit of depth 1, associated with an n-fold tensor product of Hadamard gates, i.e.,

$$|\Omega\rangle = \left(\bigotimes_{i=1}^{n}\mathsf{H}\right)|\mathbf{0}\rangle, \tag{80}$$

where $\mathsf{H}: \mathbb{B} \to \mathbb{B}$ is the Hadamard gate, represented by the matrix

$$\mathsf{H} = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}.$$

We will come back to this point in Sec. XI when the algorithm is implemented on an actual quantum computer.

Next, recall that the basis functions ψ_j of $\mathcal{H}_n$ have the form $\psi_j = e^{-\tau|j|_p/2}\phi_j$, where the ϕ_j are Fourier functions on the Abelian group $X = \mathbb{T}^d$ (see Sec. III B). Since the Fourier functions are characters of the group, they take the value $\phi_j(e) = 1$ on the identity element $e \in X$ (the point with angle coordinates $\theta = 0$), and thus

$$|\hat{\xi}_{e,n}\rangle = \sum_{\boldsymbol{b}\in\{0,1\}^n}\frac{e^{-\tau|\boldsymbol{\eta}^{-1}(\boldsymbol{b})|_p/2}}{\sqrt{\kappa_n}}|\boldsymbol{b}\rangle.$$

It follows that

$$\||\hat{\xi}_{e,n}\rangle - |\Omega\rangle\|^2_{\mathbb{B}_n} = \sum_{\boldsymbol{b}\in\{0,1\}^n}\left|\frac{1}{\sqrt{N}} - \frac{e^{-\tau|\boldsymbol{\eta}^{-1}(\boldsymbol{b})|_p/2}}{\sqrt{\kappa_n}}\right|^2, \tag{81}$$

and noting that $\lim_{\tau\to 0}\kappa_n = N$ [see (45)], we conclude that, for fixed n, $|\hat{\xi}_{e,n}\rangle$ converges to $|\Omega\rangle$ as $\tau \to 0$. In particular, since $|\Omega\rangle$ can be efficiently prepared via (80), we can efficiently approximate $|\hat{\xi}_{e,n}\rangle$ by $|\Omega\rangle$ to arbitrarily high precision.

We now claim that every state vector $|\hat{\xi}_{x,n}\rangle$ from (79) can be reached efficiently from $|\hat{\xi}_{e,n}\rangle$ by applying a suitable unitary Koopman operator. Indeed, letting $S^x: \mathfrak{A} \to \mathfrak{A}$ be the shift operator by $x = (\theta^1, \ldots, \theta^d) \in \mathbb{T}^d$, i.e.,

$$(S^x f)(y) = f(x+y),$$

we have that $S^x = U^t$, where U^t is the Koopman operator for any time t and rotation frequencies $\alpha_1, \ldots, \alpha_d$ such that $x = (\alpha_1 t, \ldots, \alpha_d t)$. Thus, if $\hat{S}^x_n: \mathbb{B}_n \to \mathbb{B}_n$ is the unitary operator induced at the quantum computational Hilbert space $\mathbb{B}_n$ by S^x [cf. (52)],

$$\hat{S}^x_n = (\mathcal{W}_n \circ \mathbf{\Pi}_n \circ \mathbf{\Pi})S^x,$$

we can implement $\hat{S}^x_n$ with a circuit of size $O(n)$ and depth $O(1)$ using an analogous approach to that used for the Koopman operator. In particular, by translation invariance of the kernel k [see (11)], we have $\xi_x = S^{-x}\xi_e$, and thus $|\hat{\xi}_{x,n}\rangle = \hat{S}^{-x}_n|\hat{\xi}_{e,n}\rangle$. Therefore, the state vector $|\hat{\xi}_{x,n}\rangle$ can be obtained efficiently by application of that circuit to $|\hat{\xi}_{e,n}\rangle$.

Consider now the state vector

$$|\check{\xi}_{x,n}\rangle := \hat{S}_n^{-x}|\Omega\rangle. \tag{82}$$

We have

$$\begin{aligned}\||\check{\xi}_{x,n}\rangle - |\hat{\xi}_{x,n}\rangle\|_{\mathbb{B}_n} &= \left\|\hat{S}_n^{-x}|\Omega\rangle - \hat{S}_n^{-x}|\hat{\xi}_{e,n}\rangle\right\|_{\mathbb{B}_n} \\ &= \||\Omega\rangle - |\hat{\xi}_{e,n}\rangle\|_{\mathbb{B}_n},\end{aligned}$$

where we have used the unitarity of $\hat{S}_n^{-x}$ to obtain the last equality. By (81), it follows that as $\tau \to 0$ at fixed n, $|\check{\xi}_{x,n}\rangle$ converges to $|\hat{\xi}_{x,n}\rangle$. We therefore conclude that for any error tolerance ϵ there exists $\tau > 0$ such that the desired initial state vector, $|\hat{\xi}_{x,n}\rangle$, is approximated by $|\check{\xi}_{x,n}\rangle$ with an error of at most ϵ in the norm of $\mathbb{B}_n$. Moreover, the state vector $|\check{\xi}_{x,n}\rangle$ can be prepared by passing the initial quantum computational state vector $|\mathbf{0}\rangle$ through a circuit of size $O(n)$ and depth $O(1)$. As with the QFT-based measurement scheme (see Sec. VII B and Appendix C), as $n \to \infty$, errors due to approximation of $|\hat{\xi}_{x,n}\rangle$ by $|\check{\xi}_{x,n}\rangle$ can be controlled by taking a decreasing sequence of RKHA parameters τ.

IX. COMPARISON WITH CLASSICAL SIMULATORS OF QUANTUM CIRCUITS

The reader may have noticed that with the state preparation scheme described in Sec. VIII, QECD results in a quantum computational system where entanglement is produced only by a terminating QFT. That is, we start with the state vector $|\mathbf{0}\rangle$ from (78) (which is a tensor product of single-qubit state vectors $|0\rangle$, and thus unentangled), and then

(1) Apply to $|\mathbf{0}\rangle$ a tensor product $\bigotimes_{i=1}^{n} \mathsf{H}$ of Hadamard gates to obtain the state vector $|\Omega\rangle$ in (80).

(2) Apply to $|\Omega\rangle$ the shift operator $\hat{S}_n^{-x}$ to obtain the state vector $|\check{\xi}_{x,n}\rangle$ in (82) encoding the classical initial condition x.

(3) Apply to $|\check{\xi}_{x,n}\rangle$ the Koopman operator $\hat{U}_n^{-t}$ in (63) to obtain the state vector $|\check{\xi}_{x,n}^{(t)}\rangle := \hat{U}_n^{-t}|\check{\xi}_{x,n}\rangle$ representing the quantum state $\hat{\rho}_{x,n}^{(t)} \approx |\check{\xi}_{x,n}^{(t)}\rangle\langle\check{\xi}_{x,n}^{(t)}|$ in (66).

(4) Apply to $|\check{\xi}_{x,n}^{(t)}\rangle$ the QFT to arrive at the state vector $|\check{\xi}_{x,n}^{(t)}\rangle := \mathcal{F}_{n,d}|\check{\xi}_{x,n}^{(t)}\rangle$ representing the state $\tilde{\rho}_{x,n}^{(t)} \approx |\check{\xi}_{x,n}^{(t)}\rangle\langle\check{\xi}_{x,n}^{(t)}|$ in (76).

Note that the relations $\hat{\rho}_{x,n}^{(t)} \approx |\check{\xi}_{x,n}\rangle\langle\check{\xi}_{x,n}|$ and $\tilde{\rho}_{x,n}^{(t)} \approx |\check{\xi}_{x,n}^{(t)}\rangle\langle\check{\xi}_{x,n}^{(t)}|$ in Steps 3 and 4 are approximate since in Steps 1 and 2 we use the approximate state preparation scheme of Sec. VIII, but the approximation errors vanish as $n \to \infty$ along a decreasing sequence of τ. In this procedure, the only operation that produces entanglement is the QFT in Step 4. All operations in Steps 1–3 operate on individual qubits, and result in no entanglement.

It is known [62,63] that under certain conditions, quantum computational algorithms exhibiting no entanglement can be simulated efficiently using randomized classical algorithms. In Ref. [62], Josza and Linden (JL) consider a quantum algorithm that runs in $T(n)$ steps, where T is a polynomial function of the number of qubits n. They show that if the input state is sufficiently unentangled, it is possible to classically sample the measurement distribution in the quantum computational basis at an error tolerance η using a randomized classical algorithm with running time

$$\mathcal{T}(n, \eta) = \text{poly}(T(n), \log(1/\eta)). \tag{83}$$

Here, poly(u) represents a polynomial function of u that asymptotically bounds the quantity of interest (in this case, running time) as $u \to \infty$. Using this result, Browne [64] shows that certain classes of quantum circuits where entanglement is produced only by a terminating QFT can also be efficiently simulated classically (i.e., in a polynomial running time $\mathcal{T}$) using appropriate classical simulators of the QFT. The QFT simulators they propose include the semiclassical QFT algorithm of Griffiths and Niu [79], as well as approximate QFT algorithms [80,81].

We will now compare the computational complexity of our QECD scheme against randomized classical algorithms which rely on the theory of JL (including Ref. [64]).

We begin by noting that the tolerance parameter η in (83) represents errors due to rational approximation of quantum gates, which is necessary for efficient classical simulability. More specifically, η represents the error in the trace norm between the exact quantum state employed in a given quantum computational algorithm and the state obtained by running the algorithm with rational approximations of its quantum gates. In certain classes of problems (e.g., discrete decision problems), a finite tolerance is sufficient and it is not necessary to consider arbitrarily small η. Yet, in the setting of our work, i.e., continuous-time dynamical systems, taking $\eta \to 0$ limits is necessary to attain convergence. In this context, an essential requirement for a consistent and robust simulator of a continuous-time system is that its prediction errors converge to 0 uniformly over time intervals and sets of initial conditions (cf. the convergence results for QECD in Proposition 3 and Corollary 8).

Consider that the evolution time t varies in an interval $[0, t_{\max}]$. As t varies in that interval, the $\mathsf{R}_z(\vartheta_l(t))$ gates implementing the Koopman evolution $\hat{U}_n^{-t}$ of our circuit swipe through continuous intervals of phase angles $\vartheta_l(t)$ [see (64)]. As a result, the collection of the matrix elements $(\mathsf{R}_z(\vartheta_l(t)))_{ij}$, as the $\vartheta_l(t)$ vary in these intervals, is an uncountable set that does not lie in any finite algebraic extension of the rationals. This means that rational approximations of $(\mathsf{R}_z(\vartheta_l(t)))_{ij}$ of the form $a_{ijl}(t)/b_{ijl}(t)$ for complex numbers $a_{ijl}(t)$ and $b_{ijl}(t)$ with integer real and imaginary parts, satisfying

$$\left|(\mathsf{R}_z(\vartheta_l(t)))_{ij} - \frac{a_{ijl}(t)}{b_{ijl}(t)}\right| < \eta$$

uniformly over $t \in [0, t_{\max}]$ and with respect to i, j, l, require numbers $b_{ijl}(t)$ of modulus $\Theta(\eta^{-1/2})$ [82]. The reason for this requirement stems from the fact that there exist badly approximable real numbers, i.e., real numbers u satisfying $|u - a/b| > c/b^2$ for some $c > 0$ and every rational a/b. Thus, in order to uniformly attain a tolerance η, a classical simulator based on the theory of JL utilizes integers of m digits, where

$$m = \Theta(\log \eta). \tag{84}$$

In Appendix A 5 [see, in particular, (A18)], we show that for fixed RKHA parameters $p \in (0, 1)$ and $\tau > 0$, the n-qubit quantum computational state $\hat{\rho}_{x,n}^{(t)}$ employed by QECD approximates the infinite-dimensional quantum state ρ_x (see Sec. IV) at an error of $O(e^{-\tilde{\tau} n^p/2})$ for any $\tilde{\tau} \in (0, \tau)$, uniformly with respect to the evolution time $t \in \mathbb{R}$ and initial condition

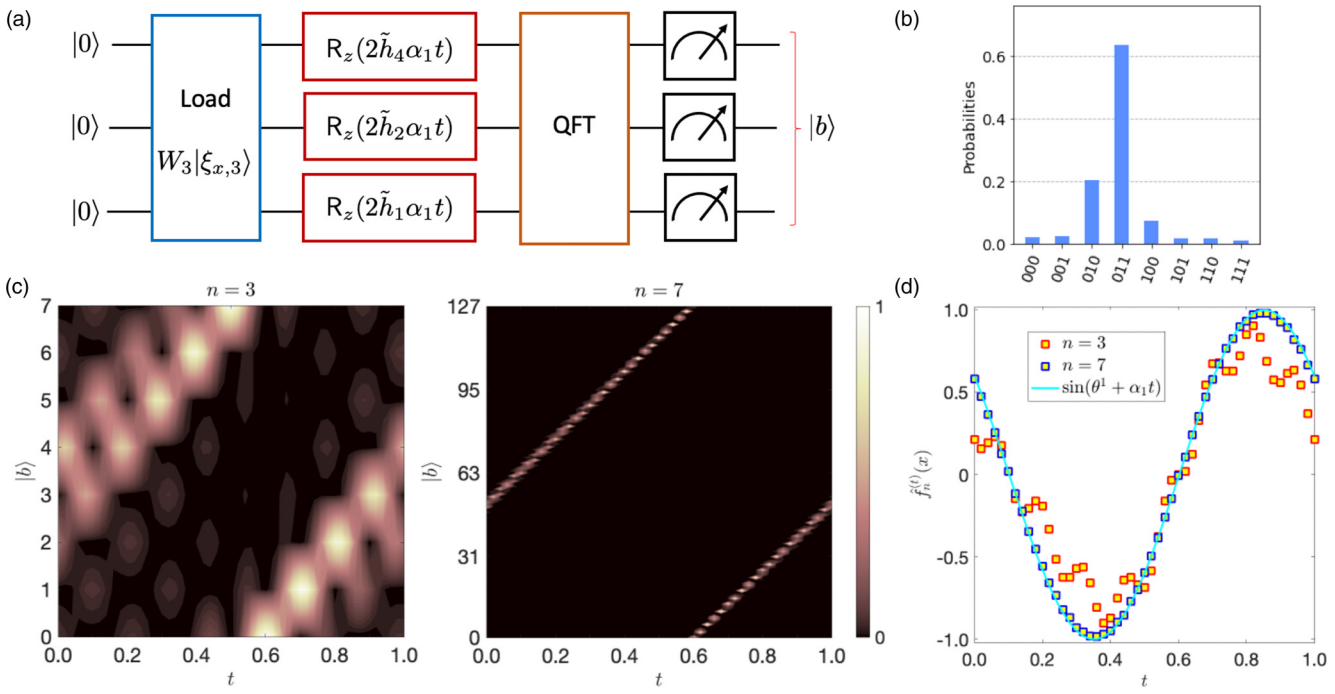

FIG. 6. Quantum circuit implementation of the three- and seven-qubit approximation of a circle rotation with frequency $\alpha_1 = 2\pi$ in the ideal Qiskit Aer environment. (a) Circuit diagram with $n = 3$ qubits, comprising (from left to right) of state vector load, Koopman evolution over time t using R_z gates, quantum Fourier transform (QFT), and measurement. (b) Empirical distribution of an ensemble of $K = 10^6$ projective measurements (shots) of the projection-valued measure (PVM) associated with the computational basis vectors $|\boldsymbol{b}\rangle \equiv |b\rangle$ for $n = 3$ and $t = 0.94$. (c) Temporal evolution of the empirical probability distributions for $n = 3$ and 7. (d) Reconstruction of the classical observable $f^{(t)}(x) = \sin(x(t)) = \sin(\theta^1 + \alpha_1 t)$ from the ensemble means, $\hat{f}_n^{(t)}(x)$. The analytical result $f^{(t)}(x)$ is plotted as a solid line. In panels (b)–(d), the initial condition is $x = \theta^1 = 2.5$ and the reproducing kernel Hilbert algebra (RKHA) parameters are $p = \tau = 1/4$. Measurements are performed at a fixed time step $\Delta t = 0.02$. In panels (b) and (c), the computational basis vectors $|b\rangle$ are indexed by an integer b in the range $0, \ldots, 2^{n-1}$.

$x \in X$. Setting

$$\eta = e^{-\tilde{\tau} n^p/2} \tag{85}$$

in (84), it follows that in order to achieve a comparable error to an n-qubit implementation of QECD, the classical simulator employs integer arithmetic with $m = \Theta(n^p)$ digits.

In order to derive the estimate in (83), JL assume a polynomial complexity of integer arithmetic operations, such as addition and multiplication, with respect to m. For our purposes, it is appropriate to assume that the dominant cost is due to integer multiplication associated with application of (rational approximations of) quantum gates to qubits. The complexity of this operation is $O(m^2)$ using the standard algorithm (long multiplication). Very recently, Harvey and van der Hoeven [83] developed an $O(m \log m)$ algorithm, which is conjectured to be an optimal bound consistent with the true complexity of integer multiplication.

Letting $\mathrm{mult}(m)$ denote the complexity of the multiplication algorithm employed, the running time $\mathcal{T}(n, \eta)$ of a classical simulator of our QECD circuit with the tolerance from (85) satisfies

$$\mathcal{T}(n, \eta) = O(n^2 \mathrm{mult}(m)) = O(n^2 \mathrm{mult}(n^p)). \tag{86}$$

In QECD, the running time $T(n)$ is dominated by the QFT, and is thus $T(n) = O(n^2)$. Therefore, if standard multiplication with $\mathrm{mult}(n^p) = O(n^{2p})$ is used, QECD offers an improved upper bound by a factor of n^{2p} over (86), which is approximately quadratic for $p \approx 1$. Using the Harvey and van der Hoeven algorithm, this factor drops to $n^p \log n$. However, it should be kept in mind that due to various overhead costs, the efficiency of fast multiplication algorithms such as in Ref. [83] is typically not realized since impractically large integers have to be used [84]. Thus, the bound of $\mathcal{T}(n, \eta) = O(n^{2(1+p)})$ associated with standard multiplication is the relevant one for classical simulation of our quantum circuit in practical applications. We should also point out that the analysis above does not take into account any additional costs incurred by the classical simulator in the approximation of the QFT.

X. SIMULATED QUANTUM CIRCUIT EXPERIMENTS

In this section, we demonstrate the performance of the QECD framework with simulated quantum circuit experiments implemented in the ideal Qiskit Aer simulator [61]. We consider a periodic example on the circle (Sec. X A), as well as a quasiperiodic system on the 2-torus (Sec. X B). In both cases, we compare the mean from an ensemble of quantum measurements with the true dynamical evolution of representative classical observables. The numerical results, displayed in Figs. 6 and 8 for the one- and two-dimensional examples, respectively, are in good agreement with the theory developed in Secs. IV–VII.

A. Circle rotation

According to (5), in dimension $d = 1$ the orbits of the dynamics are given by

$$x(t) = \Phi^t(x) = (\theta^1 + \alpha_1 t) \mod 2\pi,$$

where α_1 is the frequency parameter and $x = \theta^1$ the initial condition. We set $\alpha_1 = 2\pi$, so the orbits have period $2\pi/\alpha_1 = 1$. We seek to approximate the evolution of a real-valued observable $f: S^1 \to \mathbb{R}$ on the orbit starting at x, which is represented using the Koopman operator as

$$f^{(t)}(x) = U^t f(x) = f(\Phi^t(x)) = f(\theta^1 + \alpha_1 t).$$

In this experiment, we consider the bandlimited observable $f(x) = \sin x$.

The quantum circuit output by QECD, displayed graphically in Fig. 6(a), consists of the following four logical stages:

(1) A load stage, where the initial quantum state $\hat{\rho}_{x,n} = \hat{\mathcal{F}}_n(x)$ is prepared using the quantum feature map $\hat{\mathcal{F}}_n$ in (51)

(2) A dynamical evolution stage, which evolves $\hat{\rho}_{x,n}$ to the state $\hat{\rho}^{(t)}_{x,n} = \hat{\Psi}^t_n(\rho_{x,n})$ using the evolution operator $\hat{\Psi}^t_n$ in (50)

(3) A QFT stage, rotating $\hat{\rho}^{(t)}_{x,n}$ to the state $\tilde{\rho}^{(t)}_{x,n} = \mathfrak{F}_{n,d}\hat{\rho}^{(t)}_{x,n}$ using the Fourier operator in (72)

(4) A measurement stage, measuring the quantum-computational PVM $\mathcal{E}_n$ on the state $\hat{\rho}^{(t)}_{x,n}$. The quantum mechanical approximation $\hat{f}^{(t)}_n(x)$ of $f^{(t)}(x)$ is then obtained as an ensemble mean of K independent shots using (77).

The circuit is parameterized by three parameters, namely, the RKHA parameters p and τ and the number of qubits n. We set $p = \tau = 1/4$, and consider experiments with $n = 3$ and $n = 7$ qubits, corresponding to the quantum computational Hilbert spaces $\mathbb{B}_3$ and $\mathbb{B}_7$ of dimension $N = 2^3 = 8$ and $N = 2^7 = 128$, respectively. Another input parameter is the evolution time t, which we set to integer multiples of a fixed timestep $\Delta t = 0.02$ for purposes of visualization.

Since all quantum states in the pipeline are pure, in practice we implement the circuit as a sequence of operators on the corresponding state vectors. First, the initial state is given by

$$\hat{\rho}_{x,n} = |\hat{\xi}_{x,n}\rangle\langle\hat{\xi}_{x,n}|,$$

where the state vector $|\hat{\xi}_{x,n}\rangle = W_n \xi_{x,n}$ is obtained by application of the unitary $W_n: \mathcal{H}_n \to \mathbb{B}_n$ from (59) on the normalized RKHS feature vector $\xi_{x,n}$ from (45). See also (79). We note that in these experiments the state vector $|\hat{\xi}_{x,n}\rangle$ is loaded into the quantum register "exactly," using an amplitude encoding scheme applied to the initial state vector $|\mathbf{0}\rangle$ [see Fig. 6(a)], as opposed to the efficient approximate scheme described in Sec. VIII. In particular, we loaded $|\hat{\xi}_{x,n}\rangle$ using the Qiskit function `QuantumCircuit.initialize`. We will discuss experiments utilizing the preparation approach of Sec. VIII in Sec. XI.

The next step is the unitary Koopman evolution, given by

$$\hat{\rho}^{(t)}_{x,n} = \hat{\Psi}^t_n(\hat{\rho}_{x,n}) = \hat{U}^{t*}_n |\hat{\xi}_{x,n}\rangle\langle\hat{\xi}_{x,n}| \hat{U}^t_n.$$

Here, $\hat{U}^t_n = e^{itH_n}$ is the unitary operator in (55), which is generated by the Hamiltonian H_n with the Walsh factorization in (62). We have $\hat{\rho}^{(t)}_{x,n} = |\hat{\xi}^{(t)}_{x,n}\rangle\langle\hat{\xi}^{(t)}_{x,n}|$ with

$$|\hat{\xi}^{(t)}_{x,n}\rangle = \hat{U}^{t*}_n |\hat{\xi}_{x,n}\rangle = e^{-iH_n t}|\hat{\xi}_{x,n}\rangle.$$

Therefore, our circuit implements the transformation $|\hat{\xi}_{x,n}\rangle \mapsto |\hat{\xi}^{(t)}_{x,n}\rangle$, i.e.,

$$\begin{aligned}|\hat{\xi}^{(t)}_{x,n}\rangle &= \sum_{b=0}^{2^n-1} \frac{\psi^*_{o^{-1}(b)}(x)}{\sqrt{\kappa_n}} e^{-itH_n}|b\rangle \\ &= \sum_{b=0}^{2^n-1} \frac{\psi^*_{o^{-1}(b)}(x)}{\sqrt{\kappa_n}} \left[\bigotimes_{l=0}^{n-1} \exp(-it\alpha_1 \tilde{h}_{2^l} Z)\right]|b\rangle,\end{aligned}$$

where $\tilde{h}_{2^l} = \hat{h}_{2^l}/\alpha_1$, and $\hat{h}_{2^l}$ are the Walsh-Fourier coefficients in (62). In more detail, using (B6) with $d = 1$ and $n = 3$, we obtain that all coefficients $\tilde{h}_{2^l}$ are zero except from $\tilde{h}_1 = -5/2$, $\tilde{h}_2 = -1$, and $\tilde{h}_4 = -1/2$. For $n = 7$, the seven nonvanishing coefficients are $\tilde{h}_1 = -65/2$, $\tilde{h}_2 = -16$, $\tilde{h}_4 = -8$, $\tilde{h}_8 = -4$, $\tilde{h}_{16} = -2$, $\tilde{h}_{32} = -1$, and $\tilde{h}_{64} = -1/2$. The implementation of this second step on the quantum computer is done for each qubit channel separately, as seen in Fig. 6(a), by a R_z rotation gate from (65). Specifically, we have

$$\exp(-it\alpha\tilde{h}_{2^l} Z) = \mathsf{R}_z(2\alpha_1 t \tilde{h}_{2^l}).$$

The third step is the application of the QFT, which results to

$$\tilde{\rho}^{(t)}_{x,n} = \mathfrak{F}_{n,1}\hat{\rho}^{(t)}_{x,n} = |\tilde{\xi}^{(t)}_{x,n}\rangle\langle\tilde{\xi}^{(t)}_{x,n}|,$$

where $|\tilde{\xi}^{(t)}_{x,n}\rangle = \mathfrak{F}_{n,1}|\hat{\xi}^{(t)}_{x,n}\rangle$. We again operate at the level of state vectors, effecting the transformation $|\hat{\xi}^{(t)}_{x,n}\rangle \mapsto |\tilde{\xi}^{(t)}_{x,n}\rangle$ using a standard QFT circuit. The subsequent measurement of the PVM $\mathcal{E}_n$ on the state represented by $|\tilde{\xi}^{(t)}_{x,n}\rangle$ for K shots leads to an empirical probability distribution over the binary strings $\boldsymbol{b} \in \{0, 1\}^n$ (which index the basis vectors $|b\rangle \equiv |\boldsymbol{b}\rangle$), depicted in Fig. 6(b) for a representative evolution time t. In Fig. 6(c), we display the time evolution of this probability distribution for $K = 10^6$ shots and $n = 3$ and $n = 7$ qubits. Notice that as n increases, the probability distribution becomes increasingly concentrated around straight lines that periodically wrap around the set $b = 0, \ldots, 2^n$ indexing the $|b\rangle$ vectors. This is a manifestation of the fact that the time-dependent quantum state $\tilde{\rho}^{(t)}_{x,n}$ "tracks" the underlying classical state $x(t)$.

Figure 6(d) displays the true evolution, $f^{(t)}(x)$, and simulated evolution, $\hat{f}^{(t)}_n(x)$, of the observable $f(x) = \sin x$ over the time interval $t \in [0, 1]$ starting from the initial condition $x = \theta^1 = 2.5$. The simulated evolution $\hat{f}^{(t)}_n(x)$, which is again obtained using $K = 10^6$ shots, is seen to be in good agreement with the true signal for $n = 7$ qubits. The simulation fidelity for $n = 3$ qubits is clearly degraded, exhibiting higher variance near the extrema $f^{(t)}(x) = \pm 1$ of the true signal, but nevertheless captures an approximately sinusoidal waveform with the correct frequency.

To gain intuition on the expected fidelity of the quantum computational model as a function of the number of qubits, in Fig. 7 we show the spectra of eigenvalues s_j and representative corresponding eigenfunctions u_j of the self-adjoint operator $S_n := \mathbf{\Pi}_n(Tf)$ from (47) for $n = 3$ and 7 qubits. Recall, in particular, that S_n is an approximation of the multiplication operator by f, with its spectrum of eigenvalues $\sigma(S_n)$

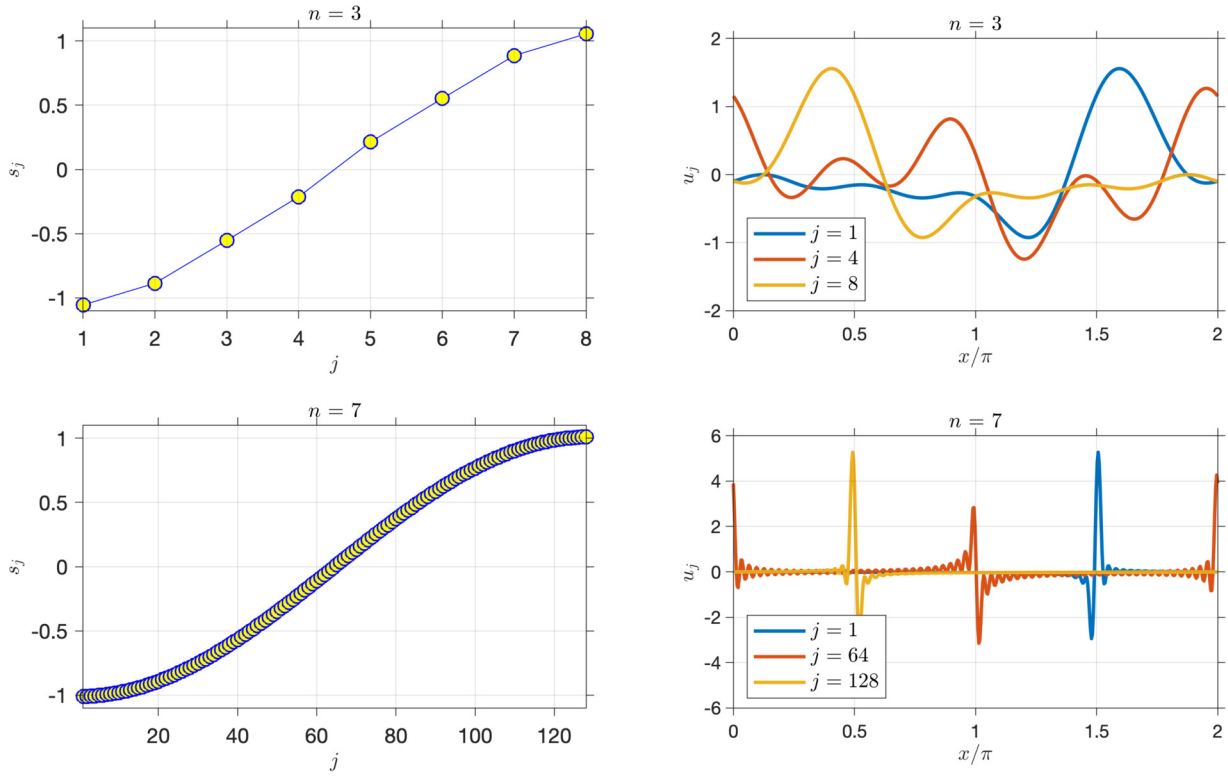

FIG. 7. Eigenvalues s_j (left-hand column) and representative eigenfunctions $u_j : S^1 \to \mathbb{R}$ (right-hand column) of the self-adjoint operators $S_n = T_n f$ representing the classical observable $f(x) = \sin x$ on the circle for the qubit numbers $n = 3$ (top row) and 7 (bottom row). The RKHA parameters are $p = \tau = 1/4$ as in Fig. 6. The index j runs from 1 to 2^n. Notice that as n increases the spectra of S_n provide an increasingly dense sampling of the range of values of f (i.e., the interval $[-1, 1]$), and the eigenfunctions $u_j(x)$ become increasingly localized around values of x for which $f(x) \approx s_j$.

providing a discretization of the (continuous) range of values of f, i.e., in this case the interval $[-1, 1]$. Moreover, S_n is unitarily equivalent to the quantum computational observable $\hat{S}_n = \mathcal{W}_n S_n$, which is in turn approximately unitarily equivalent to the Fourier-transformed observable $\tilde{S}_n = \mathfrak{F}_{n,1}\hat{S}_n$ that our circuit approximately measures. In Fig. 7 it is evident that as n increases, $\sigma(S_n)$ samples the interval $[-1, 1]$ with increasingly high density, exhibiting a clustering of eigenvalues near the boundary points ± 1. This concentration of density is consistent with the distribution of $f(x) = \sin x$ induced by a fixed-frequency rotation on the circle. Meanwhile, as n increases, the eigenfunctions exhibit increasingly high localization, with eigenfunction $u_j(x)$ concentrated on points $x \in S^1$ such that $f(x)$ is close to the corresponding eigenvalue s_j. This is seen in the right-hand column of the figure for representative eigenfunctions u_j. Thus, intuitively, as the number of qubits increases, the PVM associated with $\tilde{S}_n$ (which we approximate by the quantum computational PVM $\mathcal{E}_n$) provides a representation of the classical observable f of increasingly high resolution.

B. Quasiperiodic dynamics on the 2-torus

The two-dimensional case proceeds along similar lines as the one-dimensional example in Sec. X A, so we mainly focus on the points that are different from the one-dimensional example. The classical dynamical orbit on the 2-torus is now given by

$$x(t) = \Phi^t(x) = (\theta^1 + \alpha_1 t, \theta^2 + \alpha_2 t) \mod 2\pi,$$

where α_1 and α_2 are the frequency parameters and $x = (\theta^1, \theta^2)$ is the initial condition. We choose the (rationally independent) values $\alpha_1 = 3\sqrt{2}\pi$ and $\alpha_2 = 2\pi$, leading to an ergodic flow on $\mathbb{T}^2$. We again seek to approximate the evolution of a bandlimited classical observable f, in this case $f(x) = \sin(\theta^1)\cos(\theta^2)$. The evolution of this observable is given by

$$f^{(t)}(x) = U^t f(x) = \sin(\theta^1 + \alpha_1 t)\cos(\theta^2 + \alpha_2 t).$$

To perform quantum simulation, we set the RKHA parameters $p = \tau = 1/4$ as in Sec. X A, and use a total of $n = 8$ qubits, which corresponds to 4 qubits allocated to each torus dimension. The quantum computational Hilbert space, $\mathbb{B}_8$, is thus 256-dimensional, and admits the tensor product factorization

$$\mathbb{B}_8 = \mathbb{B}_4 \otimes \mathbb{B}_4. \tag{87}$$

For convenience in the notation, we will label the basis vectors for each of the $\mathbb{B}_4$ factors in (87) as $|\boldsymbol{k}\rangle$ and $|\boldsymbol{l}\rangle$, where $\boldsymbol{k} = (k_1, k_2, k_3, k_4)$ and $\boldsymbol{l} = (l_1, l_2, l_3, l_4)$ are four-digit binary strings. Note that the factorization in (87) is compatible with the tensor product structure of the infinite-dimensional RKHA $\mathfrak{A}$ in (13), in the sense that each $\mathbb{B}_4$ factor corresponds to the image space under a projection of the $\mathfrak{A}^{(1)}$ spaces in (13). See also Appendix B, and in particular (B6). A similar tensor

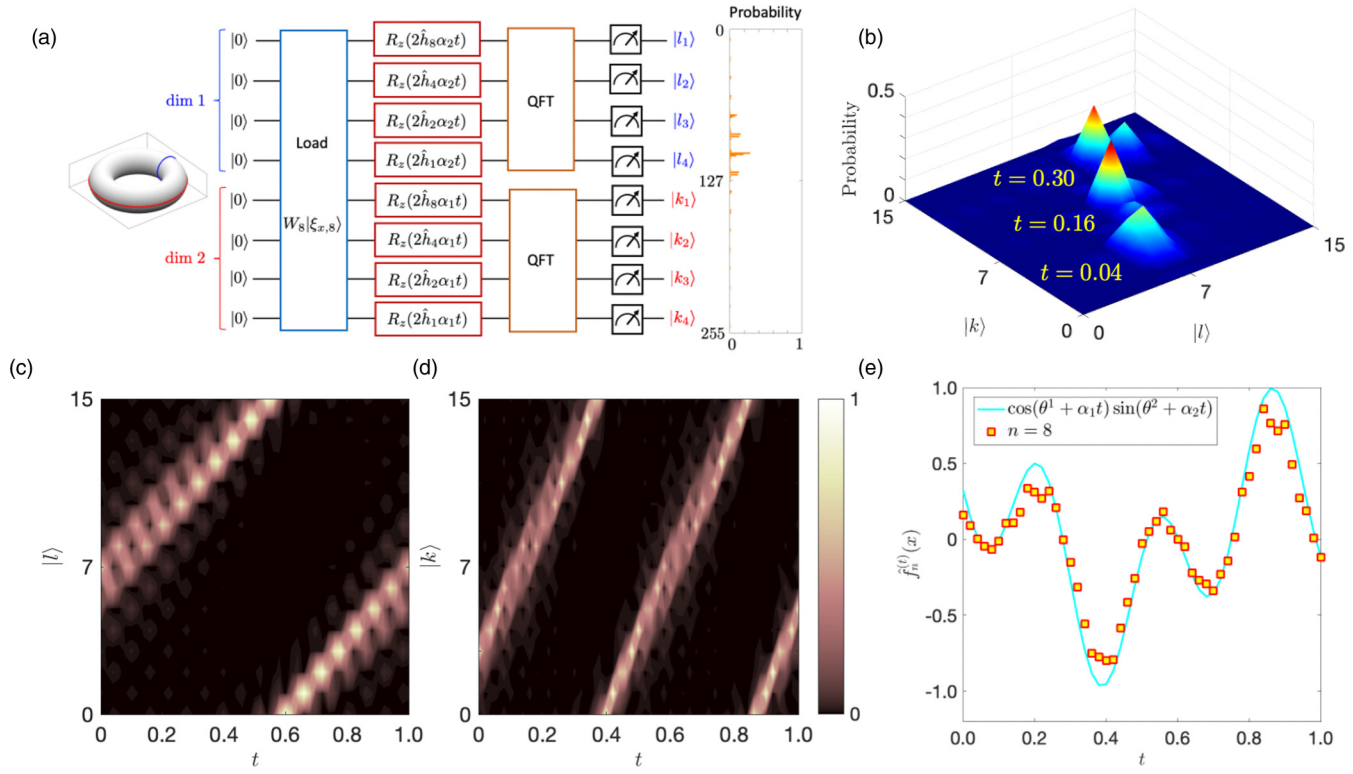

FIG. 8. As in Fig. 6, but for an eight-qubit approximation of a quasiperiodic rotation on the 2-torus with frequency parameters $\alpha_1 = 3\sqrt{2}\pi$ and $\alpha_2 = 2\pi$. (a) Quantum circuit for the quasiperiodic system, composed as two parallel copies of the circuit in Fig. 6(a) for the one-dimensional case, with 4 qubits allocated to each dimension of the 2-torus. An empirical probability distribution obtained from $K = 10^6$ shots is shown to the right of the circuit diagram, where the integers $b = 0, \ldots, 2^8 - 1 = 255$ index the computational basis vectors $|b\rangle$ of the 256-dimensional Hilbert space $\mathbb{B}_n$ with $n = 8$. The RKHA parameters are again $p = \tau = 1/4$. (b) Snapshots of the probability distribution at three representative evolution times, combined in a single surface plot. The horizontal axes labeled $|k\rangle$ and $|l\rangle$ correspond to the basis vector indices for each of the four-qubit spaces associated with each torus dimension through the factorization $\mathbb{B}_8 = \mathbb{B}_4 \otimes \mathbb{B}_4$. Note that the indices k and l range from 0 to $2^4 - 1 = 15$. (c, d) Evolution of the marginal distributions obtained by measurement of the PVMs of each of the two four-qubit spaces, i.e., one of the two torus dimensions only. The initial condition is $x = (\theta^1, \theta^2) = (1.0, 2.5)$, and measurements are performed at a fixed time step $\Delta t = 0.02$. The slopes of the probability contours in panels (c) and (d) are proportional to the frequency parameters α_2 and α_1, respectively. Notice that the slopes in panel (c) are shallower than those in panel (d) since $\alpha_2 < \alpha_1$, and are equal to the corresponding slopes in Fig. 6(c) since α_2 is equal to the frequency parameter of the one-dimensional example. (e) Reconstruction of the classical observable $f^{(t)}(x) = f^{(t)}(x_1, x_2) = \cos(\theta^2 + \alpha_2 t)\sin(\theta^1 + \alpha_1 t)$ from the ensemble means $\hat{f}_n^{(t)}(x)$ output from the quantum computer. The true evolution $f^{(t)}(x)$ is plotted as a solid line.

product structure applies for the quantum feature map, dynamical evolution, and QFT operators,

$$\hat{\mathcal{F}}_n = \hat{\mathcal{F}}_{n/2}^{(1)} \otimes \hat{\mathcal{F}}_{n/2}^{(1)}, \quad \hat{U}_n^t = \left(U_{n/2}^t\right)^{(1)} \otimes \left(U_{n/2}^t\right)^{(1)},$$
$$\mathfrak{F}_{n,2} = \mathfrak{F}_{n/2,1} \otimes \mathfrak{F}_{n/2,1}, \tag{88}$$

so we can form the entire circuit by composing two four-qubit circuits from the one-dimensional case; see Fig. 8(a) for an illustration. In (88), (1)-superscripts and 1-subscripts denote maps inherited from the one-dimensional case.

As in the one-dimensional example of Sec. X A, all quantum states occurring in our scheme are pure, so we implement the circuit in Fig. 8(a) at the level of the vectors $\xi_{x,n}$ (normalized RKHS feature vectors), $|\hat{\xi}_{x,n}\rangle = W_n \xi_{x,n}$ (initial state vectors), $|\hat{\xi}_{x,n}^{(t)}\rangle = \hat{U}_n^{t*}|\hat{\xi}_{x,n}\rangle$ (Koopman-evolved state vectors), and $|\tilde{\xi}_{x,n}^{(t)}\rangle = \mathfrak{F}_{n,2}|\hat{\xi}_{x,n}^{(t)}\rangle$ (state vectors after application of the QFT). Note that the normalized feature vector associated with classical state $x \in \mathbb{T}^2$ takes the form

$$\xi_{x,n} = \sum_{j \in J_{n,2}} \frac{\psi_j^*(x)}{\sqrt{\kappa_n}} \psi_j,$$

with $n = 8$ and

$$\psi_j^*(x) = \exp\left[-\frac{\tau}{2}(|j_1|^p + |j_2|^p)\right] \exp[-i(j_1 x_1 + j_2 x_2)].$$

See again Table II for an example of the ordering of the multi-index j and its mapping to the computational basis in the case $n = 4$ (the table would have 256 rows in the current example). We also note that our $n = 8$ example has 2×4 nonzero Walsh-Fourier expansion coefficients: $\tilde{h}_1 = -9/2$, $\tilde{h}_2 = -2$, $\tilde{h}_4 = -1$, and $\tilde{h}_8 = -1/2$ for each torus dimension.

Figure 8(b) displays snapshots of the empirical joint probability distribution of the $(\boldsymbol{k}, \boldsymbol{l})$ indices at representative evolution times t, obtained from ensembles of $K = 10^6$ measurements of the quantum computational PVM $\mathcal{E}_n$ on the state represented by $|\tilde{\xi}_{x,n}^{(t)}\rangle$ for the initial condition $x = (1.0, 2.5)$.

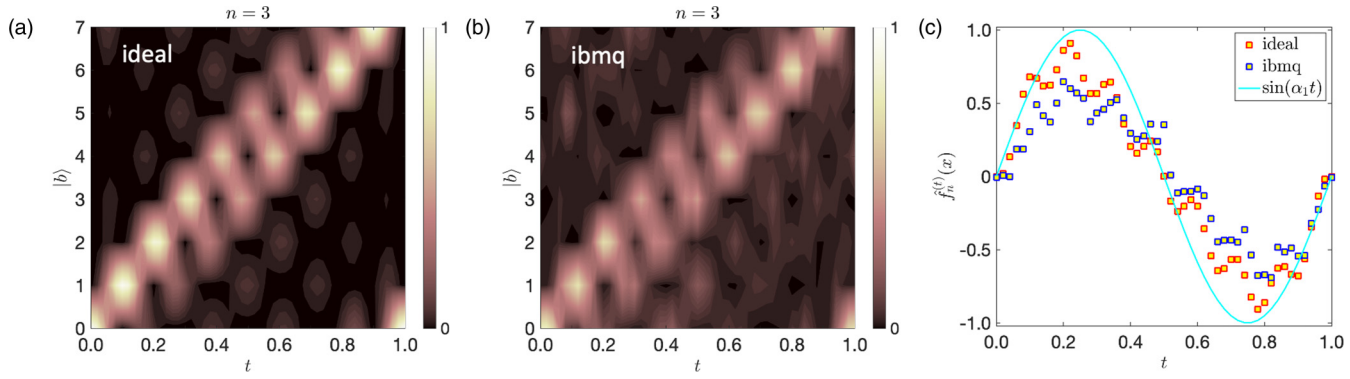

FIG. 9. Comparison of three-qubit approximations of a circle rotation with frequency $\alpha_1 = 2\pi$ from simulated circuit experiments in the ideal Qiskit Aer environment (ideal) and actual quantum computing experiments on the IBM Quantum System One (ibmq). (a) Temporal evolution of the empirical probability similar to Fig. 6(c) in the ideal Qiskit circuit simulation, using amplitude encoding with `QuantumCircuit.initialize` for the state preparation and the RKHA parameters $p = \tau = 1/4$. (b) Temporal evolution of the empirical probability distributions for $n = 3$ on the quantum computer starting with a uniform superposition state $|\Omega\rangle$ at $t = 0$. (c) Reconstruction of the classical observable $f^{(t)}(x) = \sin(x(t)) = \sin(\theta^1 + \alpha_1 t)$ from the ensemble means, $\hat{f}_n^{(t)}(x)$. The analytical result $f^{(t)}(x)$ is plotted as a solid line. In all panels, the initial condition is $x = \theta^1 = 0$. Measurements are performed at a fixed time step $\Delta t = 0.02$. The number of shots is $K = 2^{18} = 262\,144$ in both cases.

The locality of the distributions is indicative of the fact that the quantum computing model successfully tracks the orbit of the underlying classical dynamical system. Figures 8(d) and 8(d) show marginals of these distributions over the $\boldsymbol{k}$ and $\boldsymbol{l}$ index spaces as a function of time t, where periodic evolution at the generating frequencies α_1 and α_2, respectively, is apparent.

In Fig. 8(e) we compare the approximate evolution $\hat{f}_n^{(t)}(x)$ of the observable f computed from the same ensembles of quantum measurements against the true evolution $f^{(t)}(x)$. Despite the modest number of qubits allocated to each torus dimension, $\hat{f}_n^{(t)}(x)$ reproduces the quasiperiodic behavior of $f^{(t)}(x)$ to an adequate degree of accuracy, with more pronounced errors occurring near the extrema of the true signal. As in the one-dimensional example of Fig. 6(d), we expect such discrepancies to rapidly diminish as the number of qubits increases. Similarly, from this example it becomes clear how one can generalize the dynamics to a torus of dimension $d > 2$.

XI. EXPERIMENTS ON THE IBM QUANTUM SYSTEM ONE

The circle rotation algorithm for $n = 3$ qubits was also implemented on the IBM Quantum System One to demonstrate the readiness of QECD on a real NISQ device. This system has a quantum volume (an empirical metric that quantifies the capability and error rates of a quantum device) of 32. The corresponding program was again written in Qiskit (see Sec. X), and then transpiled (translated) into a sequence of appropriate elementary gate operations acting on the physical superconducting qubits via microwave channels at the hardware level. No error correction was used in our simulation. As mentioned in Sec. VIII, the encoding of 2^n (complex) amplitudes that represent the feature vector $|\hat{\xi}_{x,n}\rangle$ associated with classical state $x \in X$ in an n-qubit quantum register can lead to an exponential growth of gates. To give a concrete example for $n = 3$: amplitude encoding using `QuantumCircuit.initialize` with no circuit optimization is transpiled into a sequence of 84 elementary quantum gates. This conversion results to 52 elementary gates for a higher transpiler optimization level of 2.

To circumvent this expensive amplitude encoding of classical data, it was shown in Sec. VIII that the initial state vector $|\hat{\xi}_{x,n}\rangle$ can also be obtained to any degree of accuracy with a circuit of size $O(n)$ and depth $O(1)$. In the particular case $x = e$ (i.e., the point with canonical angle coordinates $\theta^1, \ldots, \theta^d = 0$), the encoding reduces to a uniform superposition state for n-qubits, $|\Omega\rangle$, which is obtained via n Hadamard gates H applied to the standard basis quantum state $|0\rangle^{\otimes n}$ [see (80)]. This step reduces the number of gates, and thus the circuit depth, significantly to 33 and 30 for the transpiler optimization levels 0 and 2, respectively. This depth is close to the quantum volume of the computer.

Figure 9 directly compares the results of an ideal Qiskit Aer simulator for $n = 3$ and $\tau = p = 1/4$ with an experiment on the IBM Quantum System One for the observable $f(x) = \sin x$ and an initial uniform superposition state $|\Omega\rangle$ (approximating $|\hat{\xi}_{e,n}\rangle$). Despite the noise caused by decoherence, the evolution of probability densities [Fig. 9(b)] and expectation values [Fig. 9(c)] obtained from the NISQ device remain consistent with the Qiskit simulation [Figs. 9(a) and 9(c)]. The number of shots, which is limited to 8192 on the Quantum System One, was enhanced to 2^{18} by aggregating results from multiple jobs.

Unfortunately, increasing the number of qubits beyond $n = 3$ led to noticeable degradation of the results on the quantum computer relative to the Qiskit simulations, despite our best efforts to manage noise and decoherence with the tools available to us. Still, to our knowledge, the $n = 3$ results reported in this section constitute the first successful simulation of an observable of a classical dynamical system on a manifold by an actual NISQ device. We expect that as the coherence characteristics, error mitigation and/or circuit optimization schemes for quantum computation improve, the QECD framework presented in this paper will successfully scale to higher qubit numbers.

XII. SUMMARY AND OUTLOOK

We have developed a framework for approximating the evolution of observables of a classical dynamical system by a finite-dimensional quantum system implementable on an actual quantum computer. The procedure, which we refer to as quantum embedding of classical dynamics (QECD), takes the classical system as an input, and passes through intermediate classical statistical, infinite-dimensional quantum mechanical, and finite-dimensional quantum mechanical (matrix-mechanical), representations, ultimately arriving at an n-qubit quantum computational representation of the system. We have thus addressed the full pipeline starting from the classical dynamical system all the way to its experimental verification on a real quantum computer, the IBM Quantum System One.

For the class of dynamical systems under study (i.e., measure-preserving, ergodic dynamical systems with pure point spectra), QECD is able to simulate a 2^n-dimensional Hilbert space of classical observables using circuits of size $O(n^2)$ and depth $O(n)$. This constitutes an exponential advantage over deterministic classical algorithms for Koopman operator approximation, where the cost scales linearly with the subspace dimension, i.e., is $O(2^n)$. In addition, the quantum state encoding of the initial classical state is efficiently prepared, and predictions from the quantum computational system are extracted through projective measurement in the standard computational basis without requiring postprocessing techniques such as quantum state tomography. QECD also has a near-quadratic computational advantage over randomized classical algorithms for quantum circuit simulation [62,64] (see Sec. IX).

One of the mathematical underpinnings of our approach is the theory of reproducing kernel Hilbert spaces (RKHSs). RKHS theory is widely used in kernel methods for machine learning, but was employed here to construct quantum mechanical analogs of feature maps that behave consistently under classical function evaluation and quantum mechanical expectation. A further foundational ingredient is the operator-theoretic description of dynamical systems, which utilizes linear Koopman operators to characterize the action of a (nonlinear) dynamical system on observables.

We described how QECD proceeds along two composite mappings, one taking state variables $x \mapsto \hat{\rho}_{x,n}$ to density operators $\hat{\rho}_{x,n}$ on an n-qubit Hilbert space, $\mathbb{B}_n$, and another one taking classical observables $f \mapsto \hat{S}_n$ to self-adjoint operators $\hat{S}_n$ on $\mathbb{B}_n$. A key aspect of the resulting quantum system is a tensor product factorization of its Hamiltonian in terms of Walsh operators, yielding quantum circuits of low size and depth. In particular, it was shown that for an ergodic dynamical system with finitely generated pure point spectrum, this factorization results in a circuit of size $O(n)$ and no cross-channel communication, implementing unitary Koopman evolution. The QECD framework also includes a state preparation stage of size $O(n)$, as well as a quantum Fourier transform (QFT) stage of size $O(n^2)$ to enable information retrieval through measurement in the computational basis.

The scheme exhibits three types of approximation error, all of which can be controlled, as we have shown, in appropriate asymptotic limits:

(1) Finite-dimensional approximation errors due to projection of the infinite-dimensional quantum system on the RKHS $\mathcal{H}$ to the finite-dimensional quantum computational system on $\mathbb{B}_n$. These errors vanish as $n \to \infty$, and the uniform convergence is unconditional on the defining parameters of $\mathcal{H}$ if idealized state preparation and measurement is employed (see Sec. V).

(2) Bias errors due to preparation of an approximate initial quantum state and measurement of an approximate observable using efficient circuits. These errors vanish in a joint limit of decreasing RKHS parameter τ and increasing n (see Secs. VII, VIII, and Appendix C).

(3) Monte Carlo errors associated with approximation of quantum mechanical expectations with a finite number of measurement shots (see Sec. VII B). These errors vanish as the number of shots, K, increases at fixed n and τ.

We illustrated our approach with periodic and quasiperiodic dynamical systems on the circle and 2-torus, respectively, where many aspects of the quantum embedding of classical dynamics can be directly validated against closed-form solutions. Our numerical experiments were based on simulated quantum circuits of up to $n = 8$ qubits, implemented using the Qiskit framework. In addition we demonstrated the ability of our framework to deal with a classical dynamical system on a real noisy quantum computer. The results demonstrated high-fidelity simulation of the evolution of classical observables through ensemble averages of independent quantum measurements. Our approach is straightforwardly generalizable to quasiperiodic dynamics of arbitrarily large intrinsic dimension through parallel composition of quantum circuits.

The work presented in this paper should be considered a first step, particularly given its focus on systems with pure point spectra. Applications of the procedure to mixing (chaotic) dynamical systems will invariably have to deal with the continuous spectrum of the Koopman operator, potentially generating quantum circuits of higher connectivity than for quasiperiodic dynamics. Studies in this direction are currently underway using RKHS-based spectral discretization approaches for Koopman operators [29] (see Appendix E 3 b), which are able to consistently approximate, in a spectral sense, measure-preserving, ergodic dynamical flows of arbitrary spectral character (pure point spectrum, mixed spectrum, and continuous spectrum) by unitary evolution groups with pure point spectra. A possible route to generalize QECD to this class of systems is to employ the scheme of Ref. [29] to first approximate the Koopman group on $L^2(\mu)$ by a unitary evolution group on an RKHS with a discrete spectrum, and then apply the quantum computational techniques developed in this paper to simulate the discrete-spectrum system.

Another avenue of future research is to develop data-driven formulations of the present quantum embedding framework, using kernel methods to build orthonormal bases from dynamical trajectory data, and employ these bases to represent quantum mechanical states and observables [25] (see Appendix E). This line of research would address the important problem of how errors in data-driven approximation of Koopman operators propagate into errors in the dynamical predictions made by the quantum computer. In addition, insights from classical data-driven approximation techniques for Koopman operators should lead to a systematic development

of quantum machine learning algorithms that can describe classical dynamical systems on NISQ devices. This comprises not only classification and regression tasks [36], but also the development of data-driven quantum algorithms for modeling nonlinear dynamics in high-dimensional phase spaces. A longer-term goal would be to explore applications of quantum mechanical methodologies to perform simulation and forecasting of real-world systems such as climate dynamics [85] and turbulent fluid flows [86].

ACKNOWLEDGMENTS

We wish to thank Sachin Bharadwaj for helpful discussions. We acknowledge the use of IBM Quantum services for this work. The views expressed are those of the authors and do not reflect the official policy or position of IBM or the IBM Quantum team. In this paper we used ibmq_ehningen, which is one of the IBM Quantum Falcon Processors. We thank the Fraunhofer Gesellschaft (Germany) for support. D.G. acknowledges support from the U.S. National Science Foundation under Grants No. 1842538 and DMS-1854383, the U.S. Office of Naval Research under MURI Grant No. N00014-19-1-242, and the U.S. Department of Defense under Vannevar Bush Faculty Fellowship Grant No. N00014-21-1-2946. The work of A.O. was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award No. DE-SC0002164 (underlying dynamical techniques), and by the U.S. National Science Foundation under Awards No. STC 1231306 (underlying data analytical techniques) and DBI-2029533 (underlying analytical models). P.P. and J.S. are supported by the Deutsche Forschungsgemeinschaft with Project No. SCHU 1410/30-1 and by Project No. P2018-02-001 "DeepTurb—Deep Learning in and of Turbulence" of the Carl Zeiss Foundation (Germany). J.Sl. acknowledges support from the Institute for Basic Sciences (IBS), Republic of Korea, under IBS-R028-D1.

APPENDIX A: QUANTUM MECHANICAL REPRESENTATION OF CLASSICAL OBSERVABLES

In this Appendix, we state various properties and results on the representation of classical observables by quantum mechanical operators employed in the main text.

1. Banach *-algebra structure of $\mathfrak{A}$

The fact that the RKHA $\mathfrak{A}$ from Sec. III B is an Abelian, unital, Banach *-algebra under pointwise multiplication of functions means that it has the following defining properties:

(1) $\mathfrak{A}$ is closed under pointwise multiplication of functions, i.e., the function $h: X \to \mathbb{C}$ with $h(x) = f(x)g(x)$ lies in $\mathfrak{A}$ whenever f and g lie in $\mathfrak{A}$. Thus, $\mathfrak{A}$ is an algebra, and is clearly Abelian since $fg = gf$.

(2) $\mathfrak{A}$ is equipped with an antilinear involution operation $^*: \mathfrak{A} \to \mathfrak{A}$ given by complex conjugation of functions, i.e., $(f^*)(x) = f(x)^*$. Thus, $\mathfrak{A}$ is also a *-algebra.

(3) There exists a constant $C > 0$ such that for every $f, g \in \mathfrak{A}$ the relationships

$$\|fg\|_{\mathfrak{A}} \leqslant C\|f\|_{\mathfrak{A}}\|g\|_{\mathfrak{A}}, \quad \|f^*\|_{\mathfrak{A}} = \|f\|_{\mathfrak{A}} \tag{A1}$$

hold. Thus, $\mathfrak{A}$ is a Banach *-algebra.

(4) The function $1_X: X \to \mathbb{C}$ equal everywhere to 1 lies in $\mathfrak{A}$ and satisfies $1_X f = f$ for all $f \in \mathfrak{A}$. Thus, finally $\mathfrak{A}$ is also unital.

More generally, the topic of Banach function algebras on locally compact Abelian groups (with respect to either pointwise multiplication or convolution) has a long history of study, e.g., [49,87–90].

2. Injectivity of the map $\tilde{T}$

We verify the assertion made in Sec. IV C that the map $\tilde{T}: \mathfrak{A} \to B(\mathfrak{A})$ is injective on $\mathfrak{A}_{\mathrm{sa}}$. For that, it is enough to show that if $\tilde{T}f = 0$ for $f \in \mathfrak{A}_{\mathrm{sa}}$, then $f = 0$. By definition of $\tilde{T}$, $\tilde{T}f = 0$ implies that $\pi f = -(\pi f)^*$, or, equivalently

$$\langle \psi_i, f\psi_j\rangle_{\mathfrak{A}} = -\langle f\psi_i, \psi_j\rangle_{\mathfrak{A}}, \quad \forall i, j \in \mathbb{Z}^d. \tag{A2}$$

Expanding $f = \sum_{l\in\mathbb{Z}^d} \tilde{f}_l \psi_l$, and setting $i = 0$ in (A2), we get

$$\tilde{f}_j^* = -c_{j,-j}\tilde{f}_{-j}.$$

However, because f is real, we have $\tilde{f}_j^* = \tilde{f}_{-j}$, and since $c_{j,-j}$ is nonzero we conclude that $\tilde{f}_j = 0$, and thus $f = 0$.

3. Consistency of representations based on the reproducing kernel Hilbert space $\mathcal{H}$

Recall the construction of the RKHS $\mathcal{H}$ in Sec. IV A. Even though $\mathcal{H}$ is a strict subspace of the RKHA $\mathfrak{A}$, the quantum feature map $\mathcal{F}: X \to Q(\mathcal{H})$ from (21) allows us to consistently recover all predictions made for classical observables obtained via the feature map $\tilde{\mathcal{F}}: X \to Q(\mathfrak{A})$ of $\mathfrak{A}$ in (23), as we now describe.

First, observe that by definition of $\varrho_x = \tilde{\mathcal{F}}(x)$ and $\rho_x = \mathcal{F}(x)$, we have

$$\rho_x = \frac{\tilde{\kappa}}{\kappa}\mathbf{\Pi}\varrho_x, \tag{A3}$$

where $\mathbf{\Pi}$ is the projector onto $B(\mathcal{H})$, defined in (28). As a result, if $A \in B(\mathfrak{A})$ is a quantum mechanical observable whose range is included in $\mathcal{H}$ (so that A is well defined as an operator on $\mathcal{H}$), and whose nullspace includes the orthogonal complement $\mathcal{H}^\perp$ in $\mathfrak{A}$, we have

$$\langle A\rangle_{\varrho_x} = \frac{\kappa}{\tilde{\kappa}}\langle A\rangle_{\rho_x}. \tag{A4}$$

Indeed, since every observable A in this class satisfies $\mathbf{\Pi}A = A$, using (A3) and the cyclic property of the trace, we get

$$\begin{aligned}\langle A\rangle_{\varrho_x} &= \mathrm{tr}(\varrho_x A) = \mathrm{tr}(\varrho_x(\mathbf{\Pi}A))\\ &= \mathrm{tr}(\varrho_x \Pi A \Pi) = \mathrm{tr}(\Pi\varrho_x \Pi A)\\ &= \frac{\kappa}{\tilde{\kappa}}\mathrm{tr}(\rho_x A) = \frac{\kappa}{\tilde{\kappa}}\langle A\rangle_{\rho_x},\end{aligned}$$

which verifies (A4). Thus, for all observables $A \in B(\mathfrak{A})$ satisfying

$$\mathrm{ran}A \subseteq \mathcal{H}, \quad \ker A \supseteq \mathcal{H}^\perp, \tag{A5}$$

expectation values with respect to ϱ_x can be recovered from expectation values with respect to ρ_x up to a constant scaling factor. For our purposes, this means that the quantum mechanical observables $\mathbf{\Pi}(\pi f)$ and $\mathbf{\Pi}(\tilde{T}f)$ obtained through

the projections $\mathbf{\Pi}\pi$ and $\mathbf{\Pi}\tilde{T}$ of π and $\tilde{T}$ from (24) and (25), respectively, satisfy (A4).

As noted in Sec. IV C 2, in order to obtain consistency between classical function evaluation and quantum mechanical expectation, analogously to (31), we introduce the modified representation maps $\varpi: \mathfrak{A} \to B(\mathcal{H})$ and $T: \mathfrak{A} \to B(\mathcal{H})$ in (29) to account for scaling errors. We reproduce the definitions here for convenience:

$$\varpi = \mathbf{\Pi}\pi L^{-1}, \quad T = \mathbf{\Pi}\tilde{T}L^{-1}.$$

We define $L: \mathfrak{A} \to \mathfrak{A}$ as the self-adjoint, diagonal operator satisfying the eigenvalue equation

$$L\psi_l = \frac{\eta_l}{\kappa}\psi_l \quad \text{with} \quad \eta_l = \sum_{j \in J_l'} e^{-\tau|j|^p}, \tag{A6}$$

where J_l' is the index set defined as

$$J_l' = \{j \in J : j + l \in J\}.$$

Note that by construction of J_l', the numbers η_l are strictly positive, and have the maximum value $\eta_0 = \kappa$. Moreover, the η_l attain their smallest value, $\kappa - e^{-\tau}$, when $|l| = 1$, i.e., the multi-index $l = (l_1, \ldots, l_d) \in \mathbb{Z}^d$ has exactly one entry l_i equal to ± 1 and all other entries equal to 0. As a result, L is an invertible operator with bounded inverse, satisfying

$$L^{-1}\psi_l = \frac{\kappa}{\eta_l}\psi_l.$$

Since $\kappa/\eta_l \geqslant 1$, we deduce that L^{-1} acts by inflating the expansion coefficients of elements of $\mathfrak{A}$ in the $\{\psi_j\}$ basis.

We then have the following proposition.

Proposition 1. The following classical–quantum consistency relation holds for every $f \in \mathfrak{A}$ and $x \in X$:

$$f(x) = \langle \varpi f \rangle_{\rho_x}.$$

Moreover, if f is a real-valued observable in $\mathfrak{A}_{\mathrm{sa}}$, we have

$$f(x) = \langle Tf \rangle_{\rho_x}.$$

Proof. Suppose that $g = \psi_l$ for some $l \in \mathbb{Z}^d$, and let $A = \mathbf{\Pi}A_g$, where $A_g = \pi\psi_l \in B(\mathfrak{A})$ is the multiplication operator by ψ_l. Then, A satisfies (A5), and using (A4), we get

$$\begin{aligned}
\langle A \rangle_{\rho_x} &= \frac{\tilde{\kappa}}{\kappa}\langle A \rangle_{\varrho_x} = \frac{\tilde{\kappa}}{\kappa}\mathrm{tr}(\varrho_x \Pi A_g \Pi) \\
&= \frac{\tilde{\kappa}}{\kappa}\sum_{j \in \mathbb{Z}^d} \langle \psi_j, \varrho_x \Pi A_g \Pi \psi_j \rangle_{\mathfrak{A}} \\
&= \frac{\tilde{\kappa}}{\kappa}\sum_{j \in J} \langle \psi_j, \varrho_x \Pi A_g \psi_j \rangle_{\mathfrak{A}} \\
&= \frac{\tilde{\kappa}}{\kappa}\sum_{j \in J} \langle \psi_j, \varrho_x \Pi(\psi_l \psi_j) \rangle_{\mathfrak{A}} \\
&= \frac{1}{\kappa}\sum_{j \in J} \langle \tilde{k}_x, \Pi(\psi_l \psi_j) \rangle_{\mathfrak{A}} \langle \psi_j, \tilde{k}_x \rangle_{\mathfrak{A}} \\
&= \frac{1}{\kappa}\sum_{j \in J_l'} \langle k_x, \psi_l \psi_j \rangle_{\mathfrak{A}} \langle k_x, \psi_j \rangle_{\mathfrak{A}}^* \\
&= \frac{1}{\kappa}\sum_{j \in J_l'} \psi_j^*(x)\psi_j(x)\psi_l(x) \\
&= \frac{1}{\kappa}\sum_{j \in J_l'} e^{-\tau|j|^p}|\phi_j(x)|^2\psi_l(x) \\
&= \frac{1}{\kappa}\sum_{j \in J_l'} e^{-\tau|j|^p}\psi_l(x) \\
&= \frac{\eta_l}{\kappa}\psi_l(x) = L\psi_l(x).
\end{aligned} \tag{A7}$$

Meanwhile, an application of (31) for $f = L\psi_l$ gives

$$L\psi_l(x) = \langle \pi(L\psi_l) \rangle_{\varrho_x}, \tag{A8}$$

and combining (A7) and (A8) we arrive at

$$\langle \mathbf{\Pi}(\pi g) \rangle_{\rho_x} = \langle \pi(Lg) \rangle_{\varrho_x}, \tag{A9}$$

where $g = \psi_l$. Since the basis vector ψ_l was arbitrary, it follows by linearity that (A9) holds for every $g \in \mathfrak{A}$. Setting, in particular, $g = L^{-1}f$ yields

$$\langle \mathbf{\Pi}(\pi(L^{-1}f)) \rangle_{\rho_x} = \langle \pi f \rangle_{\varrho_x} \iff \langle \varpi f \rangle_{\rho_x} = f(x),$$

which confirms the first claim of the proposition. The second claim, $f(x) = \langle Tf \rangle_{\rho_x}$, follows similarly under the additional assumption that $f^* = f$. ∎

4. Dynamics on the reproducing kernel Hilbert space $\mathcal{H}$

By construction, the RKHS $\mathcal{H}$ is a Koopman-invariant subspace of $\mathfrak{A}$, i.e., $U^t\mathcal{H} = \mathcal{H}$ for all $t \in \mathbb{R}$. As a result, we can define a generator $V: D(V) \to \mathcal{H}$ with $D(V) \subset \mathcal{H}$, a corresponding Koopman operator $U^t: \mathcal{H} \to \mathcal{H}$, and corresponding evolution maps on observables, $\mathcal{U}^t: B(\mathcal{H}) \to B(\mathcal{H})$, and states, $\Psi^t: \mathcal{Q}(\mathcal{H}) \to \mathcal{Q}(\mathcal{H})$ analogously to the corresponding operators associated with $\mathfrak{A}$. These operators satisfy the compatibility relations [cf. (35) and (37)]

$$\mathcal{U}^t(\varpi f) = \varpi(U^t f), \quad \Psi^t(\mathcal{F}(x)) = \mathcal{F}(\Phi^t(x))$$

for every $f \in \mathfrak{A}$, $x \in X$, and $t \in \mathbb{R}$, where $\varpi: \mathfrak{A} \to B(\mathcal{H})$ is the map on observables in (29) and $\mathcal{F}: X \to \mathcal{Q}(\mathcal{H})$ the quantum feature map in (21). In addition, using the consistency relations in Proposition 1 and (39), we get

$$\begin{aligned}
U^t f(x) &= \langle \mathcal{U}^t(\varpi f) \rangle_{\rho_x} = \langle \varpi f \rangle_{\Psi^t(\rho_x)}, \\
U^t f(x) &= \langle \mathcal{U}^t(Tf) \rangle_{\rho_x} = \langle Tf \rangle_{\Psi^t(\rho_x)},
\end{aligned} \tag{A10}$$

where $T: \mathcal{H} \to B(\mathcal{H})$ was defined in (29), and the equalities in the second line hold for real-valued functions in $\mathcal{H}$. It follows from (A10) that we can consistently represent the evolution of classical observables in $\mathfrak{A}$ [which is a dense subspace of $C(X)$] by quantum mechanical evolution of observables in $B(\mathcal{H})$, even though $\mathcal{H}$ is a nondense subspace of $\mathfrak{A}$.

5. Uniform convergence

In what follows, $\|\cdot\|$ will denote the operator norm of bounded operators on $\mathcal{H}$, and $\|\cdot\|_1$ will denote the trace norm of trace-class operators on $\mathcal{H}$, i.e.,

$$\|A\| = \sup_{f \in \mathcal{H}\setminus\{0\}} \frac{\|Af\|_{\mathcal{H}}}{\|f\|_{\mathcal{H}}}, \quad \|A\|_1 = \mathrm{tr}(\sqrt{A^*A}).$$

Here, $\sqrt{A^*A} \equiv R$ is the positive square root of A^*A, i.e., the unique positive operator $R \in B(\mathcal{H})$ such that $R^2 = A^*A$. We recall that the trace norm of a rank-1 operator of the form $A = \langle f, \cdot\rangle_{\mathcal{H}} g$ with $f, g \in \mathcal{H}$ is given by

$$\|A\|_1 = \|f\|_{\mathcal{H}} \|g\|_{\mathcal{H}}. \tag{A11}$$

Moreover, for any trace-class operator $A: \mathcal{H} \to \mathcal{H}$ and bounded operator $B \in B(\mathcal{H})$ we have

$$|\mathrm{tr}(AB)| \leqslant \|A\|_1 \|B\|. \tag{A12}$$

The following lemma establishes that as $n \to \infty$ the density operators $\rho_{x,n}$ from Sec. V converge to ρ_x in the trace norm, uniformly with respect to the initial condition x.

Lemma 2. With the notation of Sec. V, we have $\lim_{n\to\infty} \|\rho_{x,n} - \rho_x\|_1 = 0$, where the convergence is uniform with respect to $x \in X$.

Proof. Let

$$\breve{\rho}_{x,n} = \mathbf{\Pi}_n \rho_x = \frac{\kappa_n}{\kappa} \rho_{x,n}.$$

We have

$$\begin{aligned} \|\rho_{x,n} - \rho_x\|_1 &= \|\rho_{x,n} - \breve{\rho}_{x,n} + \breve{\rho}_{x,n} - \rho_x\|_1 \\ &\leqslant \|\rho_{x,n} - \breve{\rho}_{x,n}\|_1 + \|\breve{\rho}_{x,n} - \rho_x\|_1 \\ &= \left|1 - \frac{\kappa_n}{\kappa}\right| \|\rho_{x,n}\|_1 + \|\breve{\rho}_{x,n} - \rho_x\|_1 \\ &= 1 - \frac{\kappa_n}{\kappa} + \|\breve{\rho}_{x,n} - \rho_x\|_1. \end{aligned} \tag{A13}$$

Moreover,

$$\begin{aligned} &\|\breve{\rho}_{x,n} - \rho_x\|_1 \\ &= \frac{1}{\kappa} \|\langle k_{x,n}, \cdot\rangle_{\mathcal{H}} k_{x,n} - \langle k_x, \cdot\rangle_{\mathcal{H}} k_x\|_1 \\ &= \frac{1}{\kappa} \|\langle k_{x,n}, \cdot\rangle_{\mathcal{H}} k_{x,n} - \langle k_x, \cdot\rangle_{\mathcal{H}} k_{x,n} + \langle k_x, \cdot\rangle_{\mathcal{H}} k_{x,n} - \langle k_x, \cdot\rangle_{\mathcal{H}} k_x\|_1 \\ &\leqslant \frac{1}{\kappa} (\|\langle k_{x,n} - k_x, \cdot\rangle_{\mathcal{H}} k_{x,n}\|_1 + \|\langle k_x, \cdot\rangle_{\mathcal{H}} (k_{x,n} - k_x)\|_1) \\ &= \frac{1}{\kappa} (\|k_{x,n} - k_x\|_{\mathcal{H}} \|k_{x,n}\|_{\mathcal{H}} + \|k_x\|_{\mathcal{H}} \|k_{x,n} - k_x\|_{\mathcal{H}}) \\ &\leqslant \frac{2}{\kappa} \|k_x\|_{\mathcal{H}} \|k_{x,n} - k_x\|_{\mathcal{H}} \\ &= \frac{2}{\sqrt{\kappa}} \|k_{x,n} - k_x\|_{\mathcal{H}}, \end{aligned}$$

where we used (A11) to obtain the equality in the third to last line and the fact that $\|k_{x,n}\|_{\mathcal{H}} < \|k_x\|_{\mathcal{H}}$ [which follows from (41)] to arrive at the second to last line. Using again (41), we get

$$\begin{aligned} \|k_{x,n} - k_x\|_{\mathcal{H}} &= \left\| \sum_{j \in J \setminus J_n} \psi_j^*(x) \psi_j \right\|_{\mathcal{H}} \\ &= \sqrt{\sum_{j \in J \setminus J_n} |\psi_j(x)|^2} = \sqrt{\sum_{j \in J \setminus J_n} e^{-\tau |j|_p}}, \end{aligned}$$

and inserting this result in (A13) it follows that

$$\|\rho_{x,n} - \rho_x\|_1 \leqslant 1 - \frac{\kappa_n}{\kappa} + \frac{2}{\sqrt{\kappa}} \sqrt{\sum_{j \in J \setminus J_n} e^{-\tau |j|_p}}. \tag{A14}$$

Taking the $n \to \infty$ limit, we obtain

$$\lim_{n\to\infty} \|\rho_{x,n} - \rho_x\|_1 = 0,$$

as claimed, and since the right-hand side of (A14) does not depend on x it follows that the convergence to the limit uniform. ∎

Using Lemma 2, we show that the predictions made by the finite-dimensional quantum systems on $\mathcal{H}_n$ constructed in Sec. V converge, as $n \to \infty$, to those made by the infinite-dimensional system on $\mathcal{H}$, uniformly with respect to the initial condition x and evolution time t.

Proposition 3. With the notation of Sec. V, for any quantum mechanical observable $A \in B(\mathcal{H})$ we have

$$\lim_{n\to\infty} \langle A_n \rangle_{\Psi_n^t(\rho_{x,n})} = \langle A \rangle_{\Psi^t(\rho_x)},$$

where the convergence is uniform with respect to $x \in X$ and $t \in \mathbb{R}$.

Proof. As in Sec. V, we identify $A_n \in B(\mathcal{H}_n)$ with the unique observable $\tilde{A}_n \in B(\mathcal{H})$ such that $\tilde{A}_n f = A_n f$ if f lies in $\mathcal{H}_n$, and $\tilde{A}_n f = 0$ if f lies in the orthogonal complement of $\mathcal{H}_n$ in $\mathcal{H}$. With this identification, we have

$$\begin{aligned} \langle A_n \rangle_{\Psi_n^t(\rho_{x,n})} &= \mathrm{tr}\big(\Psi_n^t(\rho_{x,n}) A_n\big) \\ &= \mathrm{tr}\big(\rho_{x,n} \mathcal{U}_n^t A_n\big) \\ &= \mathrm{tr}(\rho_{x,n} \mathcal{U}^t \mathbf{\Pi}_n A) \\ &= \mathrm{tr}(\rho_{x,n} \mathbf{\Pi}_n \mathcal{U}^t A) \\ &= \mathrm{tr}((\mathbf{\Pi}_n \rho_{x,n}) \mathcal{U}^t A) \\ &= \mathrm{tr}(\rho_{x,n} \mathcal{U}^t A), \end{aligned}$$

where we used the commutative relations in (44) to obtain the equality in the fourth line and the fact that $\rho_{x,n}$ is invariant under $\mathbf{\Pi}_n$ to obtain the final equality. Therefore, using (A12) and the fact that $\mathcal{U}^t$ is an isometry of $B(\mathcal{H})$, i.e., $\|\mathcal{U}^t A\| = \|A\|$, we obtain

$$\begin{aligned} |\langle A_n \rangle_{\Psi_x^t(\rho_{x,n})} - \langle A \rangle_{\Psi^t(\rho_x)}| &= |\mathrm{tr}((\rho_{x,n} - \rho_x) \mathcal{U}^t A)| \\ &\leqslant \|\rho_{x,n} - \rho_x\|_1 \|\mathcal{U}^t A\| \\ &\leqslant \|\rho_{x,n} - \rho_x\|_1 \|A\|. \end{aligned} \tag{A15}$$

It therefore follows by Lemma 2 that

$$\lim_{n\to\infty} |\langle A_n \rangle_{\Psi_x^t(\rho_{x,n})} - \langle A \rangle_{\Psi^t(\rho_x)}| = 0,$$

as claimed. The convergence is uniform with respect to $x \in X$ by Lemma 2, and it is uniform with respect to $t \in \mathbb{R}$ since the right-hand side of (A15) does not depend on t. ∎

Before closing this Appendix, we manipulate the bound in (A14) to bring it in a simplified form for use in the complexity analysis in Sec. IX. First, we have

$$1 - \frac{\kappa_n}{\kappa} = \frac{1}{\kappa} \sum_{j \in J \setminus J_n} e^{-\tau |j|^p},$$

and inserting this expression in (A14) leads to

$$\|\rho_{x,n} - \rho_x\|_1 \leqslant \frac{1}{\kappa} S_n + \frac{2}{\sqrt{\kappa}} \sqrt{S_n}, \tag{A16}$$

where

$$S_n = \sum_{j \in J \setminus J_n} e^{-\tau |j|^p}.$$

Next, for any $\tilde{\tau} \in (0, \tau)$, we have

$$S_n = e^{-\tilde{\tau} d^p n^p} \sum_{j \in J \setminus J_n} e^{\tilde{\tau} d^p n^p - \tau |j|^p}.$$

Observe that since $\tilde{\tau} < \tau$ and $|j|^p > d^p n^p$ for every $j \in J \setminus J_n$, the exponents $\tilde{d}^p n^p - \tau |j|^p$ are all negative. Using concavity arguments for the function $f(u) = u^p$ (as done in Appendix C 4 ahead), it can be shown that the sums $\tilde{S}_n = \sum_{j \in J \setminus J_n} e^{\tilde{\tau} d^p n^p - \tau |j|^p}$ are uniformly bounded with respect to n. Using this result in (A16), we get

$$\|\rho_{x,n} - \rho_x\|_1 \leqslant \frac{\tilde{C}}{\kappa} e^{-\tilde{\tau} d^p n^p} + \frac{\sqrt{\tilde{C}}}{2} e^{-\tilde{\tau} d^p n^p / 2},$$

where $\tilde{C}$ is an upper bound for $\tilde{S}_n$. It therefore follows that there exists a constant C such that

$$\|\rho_{x,n} - \rho_x\|_1 \leqslant C e^{-\tilde{\tau} d^p n^p / 2}. \tag{A17}$$

Since the state $\rho_{x,n}$ is unitarily equivalent to the quantum computational state $\hat{\rho}_{x,n}$ from (66), we can use (A17) as an upper bound of the approximation error of the infinite-dimensional state ρ_x by the n-qubit state employed by QECD. Moreover, since the right-hand side of (A17) is independent of the initial condition x and we have $\Psi^t(\rho_x) = \rho_{\Phi^t(x)}$ and $\Psi^t(\rho_{x,n}) = \rho_{\Phi^t(x),n}$ we can conclude that the bound applies for any evolution time $t \in \mathbb{R}$, i.e.,

$$\left\| \rho_{x,n}^{(t)} - \rho_x^{(t)} \right\|_1 \leqslant C e^{-\tilde{\tau} d^p n^p / 2}. \tag{A18}$$

APPENDIX B: WALSH OPERATOR REPRESENTATION OF THE KOOPMAN GENERATOR

In this Appendix, we lay out the calculation of the discrete Walsh transform of the spectral function $h \in L_n^2([0, 1])$ of the Hamiltonian H_n induced by the Koopman generator of a quasiperiodic dynamical system, defined in (61). In particular, we show that h is expressible as a linear combination of Rademacher functions $R_l = w_{2^l}$ (without contributions from more general Walsh functions), leading to the factorization of H_n in (62).

First, by (14) and (61), for any $m \in \{0, \ldots, 2^n - 1\}$ we have

$$h\left(\frac{m}{2^n}\right) = \omega_j = \alpha_1 j_1 + \alpha_2 j_2 + \ldots + a_d j_d, \tag{B1}$$

where $j_1, \ldots, j_d$ are integers in the set J_1, defined uniquely by the property that the concatenated binary strings $\eta(j_1), \ldots, \eta(j_d)$ give the dyadic decomposition of $m/2^n$,

$$\gamma\left(\frac{m}{2^n}\right) = (\eta(j_1), \ldots, \eta(j_d)). \tag{B2}$$

We can express the left-hand side of (B2) in terms of Rademacher functions using (57), *viz.*,

$$\begin{aligned} \gamma\left(\frac{m}{2^n}\right) &= \left(\gamma_1\left(\frac{m}{2^n}\right), \ldots, \gamma_n\left(\frac{m}{2^n}\right)\right) \\ &= \frac{1}{2} - \frac{1}{2}\left(R_0\left(\frac{m}{2^n}\right), \ldots, R_{n-1}\left(\frac{m}{2^n}\right)\right). \end{aligned} \tag{B3}$$

Meanwhile, setting $m_i = o(j_i)$ and using again (57), the right-hand side of (B2) becomes

$$\begin{aligned} &(\eta(j_1), \eta(j_2), \ldots, \eta(j_d)) \\ &\quad = \left(\gamma\left(\frac{m_1}{2^{n/d}}\right), \ldots, \gamma\left(\frac{m_d}{2^{n/d}}\right)\right) \\ &\quad = \frac{1}{2} - \frac{1}{2}\left(R_0\left(\frac{m_1}{2^{n/d}}\right), \ldots, R_{n/d-1}\left(\frac{m_1}{2^{n/d}}\right), \right. \\ &\qquad\qquad R_0\left(\frac{m_2}{2^{n/d}}\right), \ldots, R_{n/d-1}\left(\frac{m_2}{2^{n/d}}\right), \\ &\qquad\qquad \ldots\ldots, \\ &\qquad\qquad \left. R_0\left(\frac{m_d}{2^{n/d}}\right), \ldots, R_{n/d-1}\left(\frac{m_d}{2^{n/d}}\right)\right). \end{aligned} \tag{B4}$$

Substituting for $\gamma(m/2^n)$ and $(\eta(j_1), \ldots, \eta(j_d))$ in (B2) using (B3) and (B4), respectively, we deduce that for each $i \in \{1, \ldots, d\}$ and $l \in \{0, \ldots, n-1\}$

$$R_l\left(\frac{m}{2^n}\right) = R_{l-(i-1)n/d}\left(\frac{m_i}{2^{n/d}}\right), \tag{B5}$$

for all $m \in \{0, \ldots, 2^n - 1\}$.

Observe now that for $j_i \in J_1$,

$$\begin{aligned} j_i &= \begin{cases} m_i - 2^{n/d-1} & : \quad 0 \leqslant m_i \leqslant 2^{n/d-1} - 1 \\ m_i - 2^{n/d-1} + 1 & : \quad 2^{n/d-1} \leqslant m_i \leqslant 2^{n/d} - 1 \end{cases} \\ &= \sum_{l=0}^{n/d-1} \frac{1 - R_l\left(\frac{m_i}{2^{n/d}}\right)}{2^{l+2-n/d}} + \frac{1 - R_0\left(\frac{m_i}{2^{n/d}}\right)}{2} - 2^{n/d-1} \\ &= -\sum_{l=0}^{n/d-1} \frac{R_l\left(\frac{m_i}{2^{n/d}}\right)}{2^{l+2-n/d}} - \frac{R_0\left(\frac{m_i}{2^{n/d}}\right)}{2} \\ &= -\sum_{l=0}^{n/d-1} \frac{R_{l+(i-1)n/d}\left(\frac{m}{2^n}\right)}{2^{l+2-n/d}} - \frac{R_{(i-1)n/d}\left(\frac{m}{2^n}\right)}{2}, \end{aligned}$$

where we used (57) and (B5) to obtain the second and last lines, respectively. Substituting the above in (B1), we obtain

$$\begin{aligned} h\left(\frac{m}{2^n}\right) &= -\sum_{i=1}^{d} \alpha_i \left[\sum_{l=0}^{n/d-1} R_{l+(i-1)n/d}\left(\frac{m}{2^n}\right) 2^{-l-2+n/d}\right] \\ &\quad - \sum_{i=1}^{d} \frac{\alpha_i}{2} R_{(i-1)n/d}\left(\frac{m}{2^n}\right) \\ &= -\sum_{i=1}^{d} \frac{\alpha_i}{2} \sum_{l=0}^{n/d-1} \left(2^{-l-1+n/d} + \delta_{l0}\right) \\ &\quad \times R_{l+(i-1)n/d}\left(\frac{m}{2^n}\right). \end{aligned}$$

We therefore conclude that for a quasiperiodic system, the spectral function of the generator h is expressible as a linear combination of Rademacher functions. Explicitly, we have

$$h = \sum_{i=1}^{d} \sum_{l=0}^{n/d-1} \hat{h}_{2^{l+(i-1)n/d}} R_{l+(i-1)n/d}, \tag{B6}$$

with

$$\hat{h}_{2^{l+(i-1)n/d}} = -\alpha_i (2^{-l-1+n/d} + \delta_{l0})/2, \tag{B7}$$

which is consistent with the factorization of the Hamiltonian H_n in (62).

APPENDIX C: APPROXIMATE DIAGONALIZATION OF OBSERVABLES USING THE QUANTUM FOURIER TRANSFORM

In this Appendix, we perform an analysis of approximate diagonalization of quantum mechanical observables induced at the quantum computational level from classical observables through the use of the QFT. In Appendixes C 1 and C 2, we describe how such quantum mechanical observables become increasingly diagonal as the number of qubits n increases, and provide explicit bounds verifying the approximate eigenvalue equation (75). In Appendix C 3, we show that quantum mechanical expectation values of the approximately diagonalized observables converge to the true expectation values in a limit of infinite qubit number n and vanishing RKHA parameter τ. Appendixes C 4 and C 5 contain proofs of two auxiliary lemmas, Lemma 4 and 6, which are stated in Appendixes C 1 and C 3, respectively.

1. Approximate diagonalization in dimension $d = 1$

We begin with the one-dimensional case, $d = 1$, where $X = S^1$. In this case, the index set J_n in (17) becomes $J_n = J_{n,1} = \{-N/2, \ldots, -1, 1, \ldots, N/2\}$ with $N = 2^n$, and the map $\mathfrak{F}_{n,d}$ in (72) reduces to the standard n-qubit QFT, $\mathfrak{F}_{n,d} \equiv \mathfrak{F}_n$. We also recall that $p \in (0, 1)$ and $\tau > 0$ are the parameters associated with the RKHA $\mathfrak{A}$.

a. Diagonalization using the regular representation π

Fixing $m \in \mathbb{Z}$, consider the regular representer (multiplication map) $\pi\psi_m \in B(\mathfrak{A})$ of basis vector ψ_m, the associated quantum computational observable

$$\hat{A}_{m,n} := (\mathcal{W}_n \circ \mathbf{\Pi}_n \circ \mathbf{\Pi} \circ \pi)\psi_m \in B(\mathbb{B}_n),$$

and the Fourier-transformed observable

$$\tilde{A}_{m,n} = \mathfrak{F}_n \hat{A}_{m,n} \equiv \mathfrak{F}_n^* \hat{A}_{m,n} \mathfrak{F}_n. \tag{C1}$$

First, note that by definition of the projection $\mathbf{\Pi}_n$, $\hat{A}_{m,n}$ is the zero operator (and thus trivially diagonal) whenever $|m| > N/2$. This is a manifestation of an effective "Nyquist limit" on the wave number m of classical observables that can be resolved by the finite-dimensional system on $\mathbb{B}_n$. Here, we are interested in characterizing the behavior of $\hat{A}_{m,n}$ in the "well-resolved" regime, $|m| \ll N/2$. The following lemma provides a bound showing that (1) such well-resolved observables $\hat{A}_{m,n}$ are approximately diagonal in the quantum computational basis $\{|0\rangle, \ldots, |N-1\rangle\}$ and (2) the diagonal part approximately recovers the values of ψ_m at particular points on the circle S^1.

Lemma 4. With the notation of (C1), the observable $\tilde{A}_{m,n}$ satisfies

$$(\tilde{A}_{m,n})_{kl} := \langle k|\tilde{A}_{m,n}|l\rangle = \psi_m(\theta_l)\delta_{kl} + \varepsilon_{mnkl},$$

where $\theta_l = 2\pi l/N$, and ε_{mnkl} is a residual obeying the bound

$$|\varepsilon_{mnkl}| \leqslant \frac{C\tau|m|}{N^{1-p}} + \frac{(2|m|+1)e^{-\tau|m|^p}}{N},$$

for a constant C independent of k, l, m, n, p, and τ.

A proof of Lemma 4 will be given in Appendix C 4. Using this basic result, we can derive error estimates for more general quantum mechanical observables than those induced by the individual basis functions ψ_m.

First, note that the terms $(2|m|+1)e^{-\tau|m|^p}$, $m \in \mathbb{Z}$, are bounded by a constant that depends on p and τ (and diverges as either of these parameters tends to 0). Moreover, since $p > 0$, $1/N$ is bounded by a constant times $1/N^{1-p}$. Thus, for every $p \in (0, 1)$ and $\tau > 0$ there exists a constant $C_{p,\tau}$ such that for all $m \in \mathbb{Z}$,

$$\frac{(2|m|+1)e^{-\tau|m|^p}}{N} \leqslant \frac{C_{p,\tau}}{N^{1-p}}.$$

This means that we can simplify the estimate for $|\varepsilon_{mnkl}|$ in Lemma 4 to (the less precise) bound

$$|\varepsilon_{mnkl}| \leqslant \frac{C_{p,\tau} + C\tau|m|}{N^{1-p}}. \tag{C2}$$

Using (C2), we estimate the square norm of the residual

$$|r_{mnl}\rangle := \tilde{A}_{m,n}|l\rangle - \psi_m(\theta_l)|l\rangle$$

as

$$\begin{aligned}
\|r_{mnl}\|_{\mathbb{B}_n}^2 &= \|(\tilde{A}_{m,n} - \psi_m(\theta_l)|l\rangle)\|_{\mathbb{B}_n}^2 \\
&= \sum_{k=0}^{N-1} |\langle k|\tilde{A}_{m,n} - \psi_m(\theta_l)I|l\rangle|^2 \\
&= \sum_{k=0}^{N-1} |\varepsilon_{mnkl}|^2 \leqslant \sum_{k=0}^{N-1} \frac{(C_{p,\tau} + C\tau|m|)^2}{N^{2(1-p)}} \\
&= \frac{(C_{p,\tau} + C\tau|m|)^2}{N^{1-2p}},
\end{aligned}$$

giving

$$\|r_{mnl}\|_{\mathbb{B}_n} \leqslant \frac{C_{p,\tau} + C\tau|m|}{N^{1/2-p}}. \tag{C3}$$

Thus, so long as $p < 1/2$, the norm of the residual converges to zero as $n \to \infty$, uniformly with respect to $l \in \mathbb{N}_0$.

We next generalize to bandlimited observables, i.e., observables $f^{(M)}: X \to \mathbb{C}$ for which there exists $M \in \mathbb{N}$ such that $f^{(M)} = \sum_{m=-M}^{M} \hat{f}_m \phi_m$, where the ϕ_m are the Fourier functions on X, and the $\hat{f}_m$ are complex expansion coefficients. We denote the vector space of such bandlimited observables on X by $\mathfrak{B}$. Note that $\mathfrak{B}$ is a dense subalgebra of $C(X)$, and is also a dense subalgebra of $\mathfrak{A}$ for any $\tau > 0$ and $p \in (0, 1)$ (in the respective norms). In particular, viewed as an element of $\mathfrak{A}$, $f^{(M)} = \sum_{m=-M}^{m} \hat{f}_m \phi_m$ can be equivalently expressed as $f = \sum_{m=-M}^{M} \tilde{f}_m \psi_m$, where $\tilde{f}_m = e^{\tau|m|^p/2}\hat{f}_m$.

By linearity, every such observable $f^{(M)} \in \mathfrak{B}$ is represented at the quantum computational level by

$$\hat{A}_n^{(M)} := (\mathcal{W}_n \circ \mathbf{\Pi}_n \circ \mathbf{\Pi} \circ \pi) f^{(M)} = \sum_{m=-M}^{M} \tilde{f}_m \hat{A}_{m,n},$$

and after application of the QFT by

$$\tilde{A}_n^{(M)} = \mathfrak{F}_n \hat{A}_n^{(M)} = \sum_{m=-M}^{M} \tilde{f}_m \tilde{A}_{m,n}. \tag{C4}$$

Thus, using Lemma 4 and (C3), we obtain

$$\begin{aligned}\langle k|\tilde{A}_n^{(M)}|l\rangle &= \sum_{m=-M}^{M} \tilde{f}_m \langle k|\tilde{A}_{m,n}|l\rangle \\ &= \sum_{m=-M}^{M} \tilde{f}_m \psi_m(\theta_l)\delta_{kl} + \sum_{m=-M}^{M} \tilde{f}_m \varepsilon_{mnkl} \\ &= f^{(M)}(\theta_l)\delta_{kl} + \varepsilon_{nkl}^{(M)},\end{aligned}$$

where the residual $\varepsilon_{nkl}^{(M)} := \sum_{m=-M}^{M} \tilde{f}_m \varepsilon_{mnkl}$ can be estimated as

$$\begin{aligned}|\varepsilon_{nkl}^{(M)}| &= \left|\sum_{m=-M}^{M} \tilde{f}_m \varepsilon_{mnkl}\right| \\ &\leqslant \left(\sum_{m=-M}^{M} |\tilde{f}_m|^2\right)^{1/2} \left(\sum_{m=-M}^{M} |\varepsilon_{mnkl}|^2\right)^{1/2} \\ &\leqslant \|f^{(M)}\|_{\mathfrak{A}} \frac{1}{N^{1-p}} \left(\sum_{m=-M}^{M} (C_{p,\tau} + C\tau|m|)^2\right)^{1/2}.\end{aligned}$$

We thus conclude that for bandlimited observables the off-diagonal residual $\varepsilon_{nkl}^{(M)}$ vanishes as $n \to \infty$ at fixed p and τ, uniformly with respect to $k, l \in \mathbb{N}_0$. For later convenience, we set

$$C_{p,\tau,M}^2 = \sum_{m=-M}^{M} (C_{p,\tau} + C\tau|m|)^2,$$

so that

$$|\varepsilon_{nkl}^{(M)}| \leqslant \frac{C_{p,\tau,M}}{N^{1-p}} \|f\|_{\mathfrak{A}}. \tag{C5}$$

Analogously to (C3) we can bound the norm of the residual $|r_{nl}^{(M)}\rangle := \tilde{A}_n^{(M)}|l\rangle - f^{(M)}(\theta_l)|l\rangle$ as

$$\left\|r_{nl}^{(M)}\right\|_{\mathbb{B}_n} = \left(\sum_{k=0}^{N-1} |\varepsilon_{nkl}^{(M)}|^2\right)^{1/2} \leqslant \|f^{(M)}\|_{\mathfrak{A}} \frac{C_{p,\tau,M}}{N^{1/2-p}}, \tag{C6}$$

and we deduce that the residual vanishes as $n \to \infty$ if $p < 1/2$.

Suppose now that $f = \sum_{m=-\infty}^{\infty} \tilde{f}_m \psi_m \in \mathfrak{A}$ is not bandlimited. Then, for any $\epsilon > 0$ there exists $M \in \mathbb{N}_0$ such that the bandlimited observable $f^{(M)} := \sum_{m=-M}^{M} \tilde{f}_m \psi_m \in \mathfrak{B}$ satisfies

$$\|f - f^{(M)}\|_{\mathfrak{A}} < \epsilon. \tag{C7}$$

Defining

$$\tilde{A}_n := (\mathfrak{F}_n \circ \mathcal{W}_n \circ \mathbf{\Pi}_n \circ \mathbf{\Pi} \circ \pi) f \tag{C8}$$

and $\tilde{A}_n^{(M)}$ by (C4), we get

$$\begin{aligned}|\langle k|\tilde{A}_n|l\rangle - f(\theta_l)\delta_{kl}| &= |\langle k|(\tilde{A}_n - \tilde{A}_n^{(M)})|l\rangle - (f(\theta_l) \\ &\quad - f^{(M)}(\theta_l))\delta_{kl} \\ &\quad + \langle k|\tilde{A}_n^{(M)}|l\rangle - f^{(M)}(\theta_l)\delta_{kl}| \\ &\leqslant |\langle k|(\tilde{A}_n - \tilde{A}_n^{(M)})|l\rangle| \\ &\quad + |f(\theta_l) - f^{(M)}(\theta_l)| \\ &\quad + |\langle k|\tilde{A}_n^{(M)}|l\rangle - f^{(M)}(\theta_l)\delta_{kl}|.\end{aligned}$$

To bound the terms in the right-hand side of the last inequality, note first that the operators $\pi: \mathfrak{A} \to B(\mathfrak{A})$, $\mathbf{\Pi}: B(\mathfrak{A}) \to B(\mathcal{H})$, $\mathbf{\Pi}_n: B(\mathcal{H}) \to B(\mathcal{H}_n)$, $\mathcal{W}_n: B(\mathcal{H}_n) \to B(\mathbb{B}_n)$, and $\mathfrak{F}_n: B(\mathbb{B}_n) \to B(\mathbb{B}_n)$ all have unit norm. Using this fact, it follows that

$$\begin{aligned}|\langle k|\tilde{A}_n - \tilde{A}_n^{(M)}|l\rangle| &= |\langle k|(\mathcal{W}_n \circ \mathbf{\Pi}_n \circ \mathbf{\Pi} \circ \pi)(f - f^{(M)})|l\rangle| \\ &\leqslant \|(\mathcal{W}_n \circ \mathbf{\Pi}_n \circ \mathbf{\Pi} \circ \pi)(f - f^{(M)})\|_{\mathbb{B}_n} \\ &\leqslant \|\mathcal{W}_n\| \|\mathbf{\Pi}_n\| \|\mathbf{\Pi}\| \|\pi\| \|f - f^{(M)}\|_{\mathfrak{A}} \\ &< \epsilon.\end{aligned}$$

Moreover, it follows from the reproducing property of $\mathfrak{A}$ that

$$\begin{aligned}|f(\theta_l) - f^{(M)}(\theta_l)| &= |\langle k_{\theta_l}, f - f^{(M)}\rangle_{\mathfrak{A}}| \\ &\leqslant \|k_{\theta_l}\|_{\mathfrak{A}} \|f - f^{(M)}\|_{\mathfrak{A}} \\ &< \kappa\epsilon.\end{aligned}$$

Using these bounds and (C5), we obtain

$$|\langle k|\tilde{A}_n|l\rangle - f(\theta_l)\delta_{kl}| < \epsilon(1+\kappa) + \frac{C_{p,\tau,M}}{N^{1-p}} \|f^{(M)}\|_{\mathfrak{A}}.$$

In particular, for large-enough N we have

$$\frac{C_{p,\tau,M}}{N^{1-p}} \|f^{(M)}\|_{\mathfrak{A}} < \epsilon,$$

and thus

$$|\langle k|\tilde{A}_n|l\rangle - f(\theta_l)\delta_{kl}| < (2+\kappa)\epsilon. \tag{C9}$$

Since ϵ was arbitrary, we conclude that as $n \to \infty$, $|\langle k|\tilde{A}_n|l\rangle - f(\theta_l)\delta_{kl}|$ converges to 0, i.e., the matrix elements of the quantum mechanical observable $\tilde{A}_n$ are consistently approximated by the matrix elements of the diagonal observable associated with the values $f(\theta_l)$. Note that unlike the bandlimited case we do not have an explicit rate for this convergence.

Consider now the residual

$$|r_{nl}\rangle = \tilde{A}_n|l\rangle - f(\theta_l)|l\rangle. \tag{C10}$$

In order to examine the asymptotic behavior of $|r_{nl}\rangle$ as $n \to \infty$, it is useful to view the spaces $\mathbb{B}_n$ as a nested family of subspaces of the sequence space ℓ^2, i.e., $\mathbb{B}_1 \subset \mathbb{B}_2 \subset \cdots \subset \ell^2$. With this identification, $\{|0\rangle, |1\rangle, \ldots\}$ is an orthonormal basis of ℓ^2, and $|r_{1l}\rangle, |r_{2l}\rangle, \ldots$ is a bounded sequence in ℓ^2. According to (C9), for any $k \in \mathbb{N}_0$, this sequence satisfies

$$\lim_{n\to\infty} \langle k|r_{nl}\rangle_{\mathbb{B}_n} = 0.$$

It then follows from standard Hilbert space results that as $n \to \infty$, $|r_{nl}\rangle$ converges to zero in the weak topology of ℓ^2. That is, for any $u \in \ell^2$, we have

$$\lim_{n\to\infty} \langle u_n|r_{nl}\rangle_{\mathbb{B}_n} = 0, \tag{C11}$$

where u_n is the orthogonal projection of u onto $\mathbb{B}_n$.

In summary, in dimension $d = 1$, the residual $|r_{nl}\rangle$ from (C10) converges weakly to zero as $n \to \infty$ for any $f \in \mathfrak{A}$. Moreover, if f is bandlimited, the convergence is strong (i.e., the residual norm vanishes) with a rate of convergence estimated by (C6).

b. Diagonalization using the self-adjoint representation $\tilde{T}$

Using the estimates obtained in Appendix C 1 a, we now derive approximate diagonalization results for the self-adjoint observables induced by the map $\tilde{T}:\mathfrak{A}\to B(\mathfrak{A})$ in (25). For any $f\in\mathfrak{A}$, consider the self-adjoint observable $\tilde{S}_n\in B(\mathbb{B}_n)$ with

$$\tilde{S}_n=(\mathfrak{F}_n\circ\mathcal{W}_n\circ\mathbf{\Pi}_n\circ\mathbf{\Pi}\circ\tilde{T})f\equiv\frac{\tilde{A}_n+\tilde{A}_n^*}{2},\tag{C12}$$

where $\tilde{A}_n$ is defined in (C8). Then, we have

$$\begin{aligned}&|\langle k|\tilde{S}_n|l\rangle-\mathrm{Re}f(\theta_l)|\\&=\tfrac{1}{2}|\langle k|\tilde{A}_n|l\rangle-f(\theta_l)\delta_{kl}+\langle k|\tilde{A}_n^*|l\rangle-f^*(\theta_l)\delta_{kl}|\\&\leqslant\tfrac{1}{2}|\langle k|\tilde{A}_n|l\rangle-f(\theta_l)\delta_{kl}|+|\langle k|\tilde{A}_n^*|l\rangle-f^*(\theta_l)\delta_{kl}|\\&=\tfrac{1}{2}|\langle k|\tilde{A}_n|l\rangle-f(\theta_l)\delta_{kl}|+|(\langle l|\tilde{A}_n|k\rangle-f(\theta_l)\delta_{kl})^*|\\&=\tfrac{1}{2}|\langle k|\tilde{A}_n|l\rangle-f(\theta_l)\delta_{kl}|+|\langle l|\tilde{A}_n|k\rangle-f(\theta_k)\delta_{lk}|,\end{aligned}$$

and we can use the results of Appendix C 1 a to bound the two terms in the last line. In particular, if $f=\sum_{m=-M}^{M}\tilde{f}_m\psi_m$ is bandlimited, then it follows from (C5) that

$$|\langle k|\tilde{S}_n|l\rangle-\mathrm{Re}f(\theta_l)|\leqslant\frac{\left|\varepsilon_{nkl}^{(M)}\right|+\left|\varepsilon_{nlk}^{(M)}\right|}{2}\leqslant\frac{C_{p,\tau,M}}{N^{1-p}}\|f\|_{\mathfrak{A}},$$

and for general $f\in\mathfrak{A}$ and large-enough N,

$$|\langle k|\tilde{A}_n|l\rangle-\mathrm{Re}f(\theta_l)\delta_{kl}|<(2+\kappa)\epsilon,$$

with the same notation as (C9). Moreover, convergence results for the residual $\tilde{S}_n|l\rangle-\mathrm{Re}f(\theta_l)|l\rangle$ can be derived analogously to those for $|r_{nl}\rangle$ in Appendix C 1 a.

2. Approximate diagonalization in dimension $d>1$

We can extend the results in Appendix C 1 to dimension $d>1$ by taking advantage of the tensor product structure of the RKHA $\mathfrak{A}$ on $\mathbb{T}^d$ and the maps effecting the transformations from the classical to the quantum computational level. Following the notation of Sec. III B, we will use (1)-superscripts to distinguish vector spaces, vectors, and linear maps associated with the circle S^1; see, e.g., (13). With this notation, the representation map $\pi:\mathfrak{A}\to B(\mathfrak{A})$ for dimension d decomposes as $\pi=\bigotimes_{i=1}^{d}\pi^{(1)}$, and similarly we have $\mathbf{\Pi}:B(\mathfrak{A})\to B(\mathcal{H})$, $\mathbf{\Pi}_n:B(\mathcal{H})\to B(\mathcal{H}_n)$, and $\mathcal{W}_n:B(\mathcal{H}_n)\to B(\mathbb{B}_n)$ with $\mathbf{\Pi}=\bigotimes_{i=1}^{d}\mathbf{\Pi}^{(1)}$, $\mathbf{\Pi}_n=\bigotimes_{i=1}^{d}\mathbf{\Pi}_{n/d}^{(1)}$, and $\mathcal{W}_n=\bigotimes_{i=1}^{d}\mathcal{W}_{n/d}^{(1)}$, where we have assumed that the number of qubits n is an integer multiple of d. We also recall the definition of the tensor product QFT operator $\mathfrak{F}_{n,d}:B(\mathbb{B}_n)\to B(\mathbb{B}_n)$ in (73), i.e.,

$$\mathfrak{F}_{n,d}A:=\mathfrak{F}_{n,d}A\mathfrak{F}_{n,d}^*\equiv\left(\bigotimes_{i=1}^{d}\mathfrak{F}_{n/d}\right)A.$$

Given any tensor product element $f=\bigotimes_{i=1}^{d}f^{(i)}\in\mathfrak{A}$, we have

$$\tilde{A}_n:=(\mathfrak{F}_{n,d}\circ\mathcal{W}_n\circ\mathbf{\Pi}_n\circ\mathbf{\Pi}\circ\pi)f=\bigotimes_{i=1}^{d}\tilde{A}_n^{(i)},$$

where

$$\tilde{A}_n^{(i)}=\left(\mathfrak{F}_{n/d}\circ\mathcal{W}_{n/d}^{(1)}\circ\mathbf{\Pi}_{n/d}^{(1)}\circ\mathbf{\Pi}^{(1)}\circ\pi^{(1)}\right)f^{(i)}.$$

Meanwhile, for any binary string $\boldsymbol{b}=(\boldsymbol{b}^{(1)},\ldots,\boldsymbol{b}^{(d)})\in\{0,1\}^n$ with associated evaluation point $x_{\boldsymbol{b}}\in\mathbb{T}^d$ from (74) we have

$$f(x_{\boldsymbol{b}})=\prod_{i=1}^{d}f^{(i)}(\theta_{\boldsymbol{b}^{(i)}}).$$

Thus, for any two computational basis vectors $|\boldsymbol{a}\rangle$ and $|\boldsymbol{b}\rangle$ of $\mathbb{B}_n$ with $\boldsymbol{a}=(\boldsymbol{a}^{(1)},\ldots,\boldsymbol{a}^{(d)})$ and $\boldsymbol{b}=(\boldsymbol{b}^{(1)},\ldots,\boldsymbol{b}^{(d)})$ we have

$$|\langle\boldsymbol{a}|\tilde{A}_n|\boldsymbol{b}\rangle-f(x_{\boldsymbol{b}})\delta_{\boldsymbol{ab}}|=\prod_{i=1}^{d}|\langle\boldsymbol{a}^{(i)}|\tilde{A}_n|\boldsymbol{b}^{(i)}\rangle-f^{(i)}(\theta_{\boldsymbol{b}^{(i)}})\delta_{\boldsymbol{a}^{(i)}\boldsymbol{b}^{(i)}}|,$$

and we can use the results of Appendix C 1 to bound the right-hand side. In particular, it follows from (C9) that $|\langle\boldsymbol{a}|\tilde{A}_n|\boldsymbol{b}\rangle-f(x_{\boldsymbol{b}})\delta_{\boldsymbol{ab}}|$ converges to 0 as $n\to\infty$, so that $\tilde{A}_n$ is consistently approximated by a diagonal observable with eigenvalues equal to the values of f at the points $\boldsymbol{x}_{\boldsymbol{b}}$. Moreover, the residual is $O(N^{1-p})$ analogously to (C6) if f is bandlimited, and converges weakly to zero as n increases in the sense of (C11).

The extension to elements of $\mathfrak{A}$ which are not of tensor product form follows by linearity. We omit the details of these calculations in the interest of brevity.

Note now that for every $f\in\mathfrak{A}$, the spectrum of the corresponding multiplication operator πf consists precisely of the range of values of f, i.e., $\sigma(\pi f)=\mathrm{ran}f$ [50]. In particular since the elements of $\mathfrak{A}$ are all continuous functions, πf has nonempty continuous spectrum, unless f is constant. Define $D_n:\mathbb{B}_n\to\mathbb{B}_n$ and $E_n:\mathbb{B}_n\to\mathbb{B}_n$ as the diagonal operators satisfying

$$D_n|\boldsymbol{b}\rangle=f(x_{\boldsymbol{b}})|\boldsymbol{b}\rangle,\quad E_n|\boldsymbol{b}\rangle=\mathrm{Re}f(x_{\boldsymbol{b}})|\boldsymbol{b}\rangle,\tag{C13}$$

where E_n is self-adjoint. The following theorem summarizes the properties of the quantum computational observables approximating πf and $\tilde{T}f$ obtained in Appendixes C 1 and C 2.

Theorem 5. Let $f\in\mathfrak{A}$ be arbitrary, and consider the operators $\tilde{A}_n$ and $\tilde{S}_n$ defined as in (C8) and (C12) for dimension $d\geqslant 1$. Consider also the diagonal operators in (C13). Then, the following hold as $n\to\infty$.

(1) The matrix elements $\langle k|\tilde{A}_n|l\rangle$ of $\tilde{A}_n$ converge to the matrix elements $\langle k|D_n|l\rangle=f(x_l)\delta_{kl}$ of D_n.

(2) The matrix elements $\langle k|\tilde{S}_n|l\rangle$ of $\tilde{S}_n$ converge to the matrix elements $\langle k|E_n|l\rangle=\mathrm{Re}f(x_l)\delta_{kl}$ of E_n.

(3) For each basis vector $|l\rangle$, the residuals $(\tilde{A}_n-\tilde{E}_n)|l\rangle$ and $(\tilde{S}_n-\tilde{E}_n)|l\rangle$ converge to zero weakly. Moreover, if f is bandlimited, the convergence is strong and the norms of the residuals are $O(N^{1-p})$.

(4) For every element $z\in\mathrm{ran}f$ there exists a sequence of eigenvalues z_n of D_n and a sequence of eigenvalues u_n of E_n such that $z=\lim_{n\to\infty}z_n$ and $\mathrm{Re}z=\lim_{n\to\infty}u_n$.

The approximate diagonalization result in (75) is a consequence of Theorem 5.

3. Convergence of quantum mechanical expectations

Thus far, we have established that every element f of $\mathfrak{A}$ can be consistently approximated in a spectral sense by operators

$D_n \in B(\mathbb{B}_n)$ which are diagonal in the computational basis. By construction (see Theorem 5) the spectra of D_n are subsets of the range of values of f. As a result, quantum measurement of D_n (which can be equivalently realized by measurement of the PVM associated with the computational basis as described in Sec. VII B) yields outcomes consistent with values that f takes on classical states in X. While this is a desirable property to have, it does not in itself guarantee that the quantum mechanical measurements are consistent with the value of f on the particular classical state that the system happen to have. Establishing this type of consistency is the goal of this Appendix.

The convergence results that we derive will turn out to hold for a decreasing sequence of RKHA parameters τ, as opposed to fixed τ values in Appendixes C 1 and C 2. Thus, in what follows, we will use the notation $\mathfrak{A}_\tau \equiv \mathfrak{A}$ to make the dependence of the RKHAs on $\tau > 0$ explicit. By construction, the spaces $\mathfrak{A}_\tau$ form an increasing nested family as τ decreases to 0; that is, for every $0 < \tau < \tau'$ and $f \in \mathfrak{A}$ we have $\mathfrak{A}_\tau \subset \mathfrak{A}_{\tau'}$ and $\|f\|_{\mathfrak{A}_\tau} \geqslant \|f\|_{\mathfrak{A}_{\tau'}}$. We also introduce explicit τ subscripts in our notation for the RKHSs $\mathcal{H}_\tau \subset \mathfrak{A}_\tau$ and $\mathcal{H}_{\tau,n} \subset \mathcal{H}_\tau$ and the operators $L_\tau : \mathfrak{A}_\tau \to \mathfrak{A}_\tau$, $\pi_\tau : \mathfrak{A}_\tau \to B(\mathfrak{A}_\tau)$, $\mathbf{\Pi}_\tau : B(\mathfrak{A}_\tau) \to B(\mathcal{H}_\tau)$, and $\mathbf{\Pi}_{\tau,n} : B(\mathcal{H}_\tau) \to B(\mathcal{H}_{\tau,n})$. τ subscripts will also be introduced in our notation for elements of $\mathfrak{A}_\tau$, $\mathcal{H}_\tau$, and the associated operator spaces as appropriate.

As in Appendixes C 1 and C 2, we consider first the one-dimensional case, $d = 1$, and an observable $f = \psi_{m,\tau}$ equal to a basis vector of $\mathfrak{A}_\tau$. We define the diagonal operator $D_{m,\tau,n} : \mathbb{B}_n \to \mathbb{B}_n$ with

$$D_{m,\tau,n}|l\rangle = \psi_{m,\tau}(\theta_l)|l\rangle$$

analogously to (C13), and also set $\tilde{D}_{m,\tau,n} \in B(\mathcal{H}_{\tau,n})$ with

$$\tilde{D}_{m,\tau,n} = (\mathcal{W}_n^* \circ \mathfrak{F}_n^*) D_{m,\tau,n} = W_n^* \mathfrak{F}_n D_{m,\tau,n} \mathfrak{F}_n^* W_n.$$

We also define

$$\tilde{A}_{m,\tau,n} = (\mathfrak{F}_n \circ \mathcal{W}_n \circ \mathbf{\Pi}_{\tau,n} \circ \mathbf{\Pi}_\tau \circ \pi_\tau)\psi_{m,\tau} \in B(\mathbb{B}_n) \quad \text{(C14)}$$

as in (C1). For any $x \in X = S^1$, we consider the quantum computational state $\hat{\rho}_{x,\tau,n} = \hat{\mathcal{F}}_{\tau,n}(x) \in Q(\mathbb{B}_n)$ and the state $\tilde{\rho}_{x,\tau,n} \in Q(\mathbb{B}_n)$ after application of the QFT,

$$\tilde{\rho}_{x,\tau,n} = \mathfrak{F}_n \hat{\rho}_{x,\tau,n} = (\mathfrak{F}_n \circ \mathcal{W}_n)\rho_{x,\tau,n}. \quad \text{(C15)}$$

We then have the following lemma.

Lemma 6. With notation as above, the $n \to \infty$ limit of the expected difference $\langle \tilde{A}_{m,\tau,n} - D_{m,\tau,n}\rangle_{\tilde{\rho}_{x,\tau,n}}$ between measurements of $\tilde{A}_{m,\tau,n}$ and $D_{m,\tau,n}$ on the state $\tilde{\rho}_{x,\tau,n}$ exists, and satisfies

$$\lim_{n\to\infty} |\langle \tilde{A}_{m,\tau,n} - D_{m,\tau,n}\rangle_{\tilde{\rho}_{x,\tau,n}}| \leqslant 1 - e^{-\tau|m|^p/2}.$$

A proof of Lemma 6 can be found in Appendix C 5. For our purposes, a key implication of the result is that while the bias in measuring $D_{m,\tau,n}$ (instead of $\tilde{A}_{m,\tau,n}$) need not vanish as $n \to \infty$, it can be made arbitrarily small for a suitable choice of τ. In particular, for any $\epsilon > 0$ there exists $\tau_m > 0$ such that for all $\tau \in (0, \tau_m)$ we have $1 - e^{-\tau|m|^p/2} < \epsilon$, and thus

$$\lim_{n\to\infty} |\langle \tilde{A}_{m,\tau,n} - D_{m,\tau,n}\rangle_{\tilde{\rho}_{x,\tau,n}}| < \epsilon. \quad \text{(C16)}$$

Since $1 - e^{-\tau|m|^p/2} \leqslant \tau|m|^p/2$, the choice $\tau_m = 2\epsilon|m|^{-p}$ will suffice for (C16) to hold.

Next, we consider bandlimited observables $f^{(M)} \in \mathfrak{B}$ of the form $f^{(M)} = \sum_{m=-M}^{M} \tilde{f}_{m,\tau}\psi_{m,\tau}$. Let

$$\tilde{A}^{(M)}_{\tau,n} = (\mathfrak{F}_n \circ \mathcal{W}_n \circ \mathbf{\Pi}_{\tau,n} \circ \mathbf{\Pi}_\tau \circ \pi_\tau)f = \sum_{m=-M}^{M} \tilde{f}_{m,\tau}\tilde{A}_{m,\tau,n} \in B(\mathbb{B}_n) \quad \text{(C17)}$$

be the corresponding quantum computational observable, and let $D_{\tau,n} \in B(\mathbb{B}_n)$ be the diagonal observable approximating $\tilde{A}_{\tau,n}$,

$$D_{\tau,n}|l\rangle = f(\theta_l)|l\rangle, \quad D_{\tau,n} = \sum_{m=-M}^{M} \tilde{f}_{m,\tau} D_{m,\tau,n}. \quad \text{(C18)}$$

Using Lemma 6 and following a similar approach as in Appendix C 1, we find

$$\lim_{n\to\infty} |\langle \tilde{A}_{\tau,n} - D_{\tau,n}\rangle_{\tilde{\rho}_{x,\tau,n}}| \leqslant C_{p,\tau,M}\|f\|_{\mathfrak{A}_\tau}, \quad \text{(C19)}$$

where

$$C^2_{p,\tau,M} = \sum_{m=-M}^{M} \left(1 - e^{-\tau|m|^p/2}\right)^2.$$

Again, for any $\epsilon > 0$, there exists $\tau_M > 0$ such that

$$\lim_{n\to\infty} |\langle \tilde{A}_{\tau,n} - D_{\tau,n}\rangle_{\tilde{\rho}_{x,\tau,n}}| < \epsilon, \quad \forall \tau \in (0, \tau_M). \quad \text{(C20)}$$

In this case, the choice $\tau_M = 2\epsilon|M|^{-(p+\frac{1}{2})}$ is sufficient for the bound to hold.

To generalize to nonbandlimited observables, we must take into account the fact that the error bounds in (C16) and (C20) imply convergence on a decreasing sequence of RKHA parameters τ, as opposed to the diagonalization results in Appendix C 1 which hold for fixed τ. With that in mind, we consider a space of classical observables that contains the RKHAs $\mathfrak{A}_\tau$ for all admissible values of the parameters τ and p. In particular, we consider observables in the Wiener algebra of X, i.e., the space of functions $f : X \to \mathbb{C}$ with absolutely convergent Fourier series, which we denote here by $\mathfrak{W}$. The Wiener algebra $\mathfrak{W}$ is a dense subalgebra of $C(X)$. Moreover, the RKHAs $\mathfrak{A}_\tau$ employed in this work are all dense subalgebras of $\mathfrak{W}$. Thus, we have the following relationships between algebras of classical observables (which also hold in dimension $d > 1$):

$$\mathfrak{B} \subset \mathfrak{A}_\tau \subset \mathfrak{W} \subset C(X).$$

Suppose then that $f = \sum_{m=-\infty}^{\infty} \hat{f}_m\phi_m$ is an arbitrary element of $\mathfrak{W}$, where the sum over m converges uniformly on X. Then, for any $\epsilon > 0$ there exists $M_* \in \mathbb{N}$ such that for every $M > M_*$ the bandlimited observable $f^{(M)} = \sum_{m=-M}^{M} \hat{f}_m\phi_m \in \mathfrak{B}$ satisfies

$$\|f - f^{(M)}\|_{C(X)} < \epsilon/3. \quad \text{(C21)}$$

The bandlimited observable $f^{(M)}$ is an element of $\mathfrak{A}_\tau$ for any $\tau > 0$, with RKHA norm satisfying

$$\|f^{(M)}\|_{\mathfrak{A}_\tau} = \left(\sum_{m=-M}^{M} e^{\tau|m|^p}|\hat{f}_m|^2\right)^{1/2} \leqslant e^{\tau M^p/2}\left(\sum_{m=-M}^{M} |\hat{f}_m|^2\right)^{1/2}$$

$$\begin{aligned}
&\leqslant e^{\tau M^p/2} \sum_{m=-M}^{M} |\hat{f}_m| \\
&= e^{\tau M^p/2} \sum_{m=-M}^{M} \left| \frac{1}{2\pi} \int_0^{2\pi} e^{-im\theta} f(\theta)\, d\theta \right| \\
&\leqslant e^{\tau M^p/2} \sum_{m=-M}^{M} \|f\|_{C(X)} \\
&= (2M+1) e^{\tau M^p/2} \|f\|_{C(X)}. \qquad \text{(C22)}
\end{aligned}$$

We will also need the observable

$$f_\tau^{(M)} = L_\tau f^{(M)} = \kappa_\tau \sum_{m=-M}^{M} \frac{\hat{f}_m}{\eta_{m,\tau}} \hat{\phi}_m$$

as an intermediate approximation associated with the bias correction introduced in Sec. IV D and Appendix A 3 to take into account the projection from $\mathfrak{A}_\tau$ to $\mathcal{H}_\tau$. Here, L_τ is the operator defined in (A6) and $\eta_{m,\tau}$ are its eigenvalues, where we have again used τ subscripts to make dependencies on that parameter explicit. We have

$$\begin{aligned}
\|f^{(M)} - f_\tau^{(M)}\|_{C(X)} &= \left\| \sum_{m=-M}^{M} \left(\frac{\eta_{\tau,m}}{\kappa_\tau} - 1 \right) \hat{f}_m \phi_m \right\|_{C(X)} \\
&\leqslant C_\tau \sum_{m=-M}^{M} |\hat{f}_m|,
\end{aligned}$$

where

$$C_\tau = \max_{m \in [-M,M]} \left| \frac{\eta_{\tau,m}}{\kappa_\tau} - 1 \right| = \frac{e^{-\tau}}{\kappa_\tau}.$$

Note that to obtain the last result we used the fact that $\eta_{\tau,m}$ lies in the interval $[\kappa_\tau - e^{-\tau}, \kappa_\tau]$; see Appendix A 3. In particular, as $\tau \to 0$, C_τ converges to 0 since $e^{-\tau}$ converges to 1 and κ_τ tends to infinity. Proceeding as in the derivation of (C22) to bound the sum $\sum_{m=-M}^{M} |\hat{f}_m|$, we arrive at

$$\|f^{(M)} - f_\tau^{(M)}\|_{C(X)} \leqslant C_\tau (2M+1) \|f\|_{C(X)}. \qquad \text{(C23)}$$

Next, define the quantum computational observable $\tilde{A}_{\tau,n}^{(M)} \in B(\mathbb{B}_n)$ as

$$\begin{aligned}
\tilde{A}_{\tau,n}^{(M)} &= (\mathfrak{F}_{\boldsymbol{n}} \circ \mathcal{W}_{\tau,n} \circ \mathbf{\Pi}_{\tau,n} \circ \mathbf{\Pi}_\tau \circ \pi_\tau) f^{(M)} \\
&= \kappa_\tau \sum_{m=-M}^{M} \frac{\tilde{f}_{m,\tau}}{\eta_{m,\tau}} \tilde{A}_{m,\tau,n}, \qquad \text{(C24)}
\end{aligned}$$

where $\tilde{f}_{m,\tau} = e^{\tau |m|^p/2} \hat{f}_m$, and $\tilde{A}_{m,\tau,n}$ are operators defined as in (C14). Define also the diagonal observable

$$D_{\tau,n}^{(M)} |l\rangle = f^{(M)}(\theta_l) |l\rangle = \sum_{m=-M}^{M} \tilde{f}_{m,\tau} \psi_{m,\tau}(\theta_l) |l\rangle. \qquad \text{(C25)}$$

Letting x be an arbitrary point in X, defining the quantum state $\tilde{\rho}_{x,\tau,n} \in Q(\mathbb{B}_n)$ as in (C15), and using (C21) and (C23), we get

$$\begin{aligned}
\left| f(x) - \langle D_{\tau,n}^{(M)} \rangle_{\tilde{\rho}_{x,\tau,n}} \right| &= \left| f(x) - f^{(M)}(x) + f^{(M)}(x) - f_\tau^{(M)}(x) + f_\tau^{(M)}(x) - \langle \tilde{A}_{\tau,n}^{(M)} \rangle_{\tilde{\rho}_{x,\tau,n}} + \langle \tilde{A}_{\tau,n}^{(M)} \rangle_{\tilde{\rho}_{x,\tau,n}} - \langle D_{\tau,n}^{(M)} \rangle_{\tilde{\rho}_{x,\tau,n}} \right| \\
&\leqslant |f(x) - f^{(M)}(x)| + \left| f^{(M)}(x) - f_\tau^{(M)}(x) \right| + \left| f_\tau^{(M)}(x) - \langle A_{\tau,n}^{(M)} \rangle_{\tilde{\rho}_{x,\tau,n}} \right| + \left| \langle \tilde{A}_{\tau,n}^{(M)} - D_{\tau,n}^{(M)} \rangle_{\tilde{\rho}_{x,\tau,n}} \right| \\
&\leqslant \|f - f^{(M)}\|_{C(X)} + \leqslant \left\| f^{(M)} - f_\tau^{(M)} \right\|_{C(X)} + \left| f^{(M)}(x) - \langle \tilde{A}_{\tau,n}^{(M)} \rangle_{\tilde{\rho}_{x,\tau,n}} \right| + \left| \langle \tilde{A}_{\tau,n}^{(M)} - D_{\tau,n}^{(M)} \rangle_{\tilde{\rho}_{x,\tau,n}} \right| \\
&< \frac{\epsilon}{3} + C_\tau (2M+1) \|f\|_{C(X)} + \left| f^{(M)}(x) - \langle \tilde{A}_{\tau,n}^{(M)} \rangle_{\tilde{\rho}_{x,\tau,n}} \right| + \left| \langle \tilde{A}_{\tau,n}^{(M)} - D_{\tau,n}^{(M)} \rangle_{\tilde{\rho}_{x,\tau,n}} \right|.
\end{aligned}$$

We can now bound the second, third, and fourth terms in the right-hand side of the last inequality. In particular, it follows by applying Proposition 1 to the observable $f_\tau^{(M)}$ that

$$\lim_{n \to \infty} \left| f_\tau^{(M)}(x) - \langle \tilde{A}_{\tau,n}^{(M)} \rangle_{\tilde{\rho}_{x,\tau,n}} \right| = 0,$$

and from (C19) and (C22) that

$$\begin{aligned}
\lim_{n \to \infty} \left| \langle \tilde{A}_{\tau,n}^{(M)} - D_{\tau,n}^{(M)} \rangle_{\tilde{\rho}_{x,\tau,n}} \right| &\leqslant C_{p,\tau,M} \|f^{(M)}\|_{\mathfrak{A}_\tau} \\
&\leqslant C_{p,\tau,M} (2M+1) e^{\tau M^p/2} \|f\|_{C(X)}.
\end{aligned}$$

Then, using the above in conjunction with the fact that $\lim_{\tau \to 0} C_\tau = 0$, it follows that for any $M \in \mathbb{N}$ there exists $\tau_M > 0$ such that for all $\tau \in (0, \tau_M)$ we have, simultaneously,

$$\begin{aligned}
C_\tau (2M+1) \|f\|_{C(X)} &< \epsilon/3, \\
C_{p,\tau,M} (2M+1) e^{\tau M^p/2} \|f\|_{C(X)} &< \epsilon/3, \qquad \text{(C26)}
\end{aligned}$$

and thus

$$\lim_{n \to \infty} \left| f(x) - \langle D_{\tau,n}^{(M)} \rangle_{\tilde{\rho}_{x,\tau,n}} \right| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + 0 + \frac{\epsilon}{3} = \epsilon.$$

Since ϵ was arbitrary, we conclude that there exists a decreasing sequence of RKHA parameters τ_M such that the quantum mechanical expectation $\langle D_{\tau_M,n}^{(M)} \rangle_{\tilde{\rho}_{x,\tau_M,n}}$ converges to the classical value $f(x)$ in the iterated limit of $M \to \infty$ (infinite bandwidth) after $n \to \infty$ (infinite qubits), and the convergence is uniform with respect to $x \in X$.

Having established this convergence result in dimension $d = 1$, we can extend it to higher dimensions using tensor product arguments analogous to those in Appendix C 2. It is also straightforward to derive analogous results using the symmetrized map $\tilde{T}_\tau : \mathfrak{A}_\tau \to B(\mathcal{H}_\tau)$, inducing the self-adjoint quantum computational observable [cf. (C24)]

$$\tilde{S}_{\tau,n}^{(M)} = (\mathfrak{F}_{\boldsymbol{n}} \circ \mathcal{W}_{\tau,n} \circ \mathbf{\Pi}_{\tau,n} \circ \mathbf{\Pi}_\tau \circ \tilde{T}_\tau) f^{(M)}$$

and the diagonal observable

$$E_{\tau,n}^{(M)} |\boldsymbol{x}_l\rangle = \mathrm{Re} f^{(M)}(\boldsymbol{x}_l) |\boldsymbol{x}_l\rangle. \qquad \text{(C27)}$$

We do not reproduce the details of these analyses in the interest of brevity. The following theorem summarizes the asymptotic convergence of our approach in these settings.

Theorem 7. Let $f = \sum_{m\in\mathbb{Z}^d} \hat{f}_m \phi_m$ be a classical observable in the Wiener algebra $\mathfrak{W}$ of $X = \mathbb{T}^d$. For $M \in \mathbb{N}$, $\tau > 0$, and $n \in \mathbb{N}$, define the bandlimited observable $f^{(M)} = \sum_{|m|\leqslant M} \hat{f}_m \phi_m$ and the corresponding diagonal quantum mechanical observables $D^{(M)}_{\tau,n}$ and $E^{(M)}_{\tau,n}$ from (C25) and (C27), respectively. Then, there exists a sequence $\tau_1, \tau_2, \ldots$, decreasing to 0, such that for any $x \in X$,

$$\lim_{M\to\infty}\lim_{n\to\infty} \langle D^{(M)}_{\tau_M,n}\rangle_{\tilde{\rho}_{x,\tau,n}} = f(x),$$
$$\lim_{M\to\infty}\lim_{n\to\infty} \langle E^{(M)}_{\tau_M,n}\rangle_{\tilde{\rho}_{x,\tau,n}} = \mathrm{Re} f(x),$$

uniformly with respect to $x \in X$.

The fact that quantum states at the quantum computational level evolve compatibly with the underlying classical dynamics (i.e., $\hat{\Psi}^t_n(\hat{\rho}_{x,\tau,n}) = \hat{\rho}_{\Phi^t(x),n}$) leads, in conjunction with the uniform convergence result in Proposition 3, to the following corollary of Theorem 7, which establishes the asymptotic consistency of QECD in simulating the evolution of classical observables.

Corollary 8. With the notation of Theorem 7 and for any $t \geqslant 0$, let $\tilde{f}^{(t)}_{M,n} \in C(X)$ with

$$\tilde{f}^{(t)}_{M,n}(x) = \langle D^{(M)}_{\tau_M,n}\rangle_{\tilde{\rho}_{\Phi^t(x),\tau,n}},$$

be the function representing the expected value of the time-t simulation of f by the quantum computer, given initial conditions x. Then,

$$\lim_{M\to\infty}\lim_{n\to\infty} \tilde{f}^{(t)}_{M,n}(x) = U^t f(x).$$

where the convergence is uniform with respect to $x \in X$ and $t \in \mathbb{R}$. Moreover, if f is real-valued, the analogous result holds for

$$\tilde{f}^{(t)}_{M,n}(x) = \langle E^{(M)}_{\tau_M,n}\rangle_{\tilde{\rho}_{\Phi^t(x),\tau,n}}.$$

Before closing this section, we note that while the convergence results in Theorem 7 and Corollary 8 hold for observables in the Wiener algebra $\mathfrak{W}$ with absolutely convergent Fourier series, the fact that $\mathfrak{W}$ is a dense subspace of $C(X)$ means that any observable $f \in C(X)$ can be approximated to arbitrarily high precision in uniform norm by an observable $g \in \mathfrak{W}$, whose dynamical evolution can in turn be simulated to arbitrarily high precision using QECD as established in Corollary 8. The function g may be constructed by several means available from signal processing, e.g., by convolution of f by an appropriate smoothing kernel. A detailed study of this topic is beyond the scope of the present work.

4. Proof of Lemma 4

Using the definition of the map W_n in (59) and the QFT in (71), we get

$$W^*_n \mathfrak{F}_n |l\rangle = W^*_n \left(\frac{1}{\sqrt{N}}\sum_{q=0}^{N-1} e^{-2\pi i l q/N}|q\rangle\right)$$
$$= \frac{1}{\sqrt{N}}\sum_{q=0}^{N-1} e^{-2\pi i l q/N}\psi_{o^{-1}(q)} = \frac{1}{\sqrt{N}}\sum_{j\in J_n} e^{-2\pi i l o(j)/N}\psi_j,$$

leading to

$$(\pi\psi_m)W^*_n\mathfrak{F}_n|l\rangle = \frac{1}{\sqrt{N}}\sum_{j\in J_n} e^{-2\pi i l o(j)/N}(\pi\psi_m)\psi_j$$
$$= \frac{1}{\sqrt{N}}\sum_{j\in J_n} e^{-2\pi i l o(j)/N}\psi_m\psi_j$$
$$= \frac{1}{\sqrt{N}}\sum_{j\in J_n} e^{-2\pi i l o(j)/N} c_{mj}\psi_{m+j}.$$

Therefore, the operator $\tilde{A}_{m,n}$ has the matrix elements

$$(\tilde{A}_{m,n})_{kl} = \langle k|\tilde{A}_{m,n}|l\rangle$$
$$= \langle k|\mathfrak{F}^*_n W_n \Pi_n(\pi\psi_m)\Pi^*_n W^*_n \mathfrak{F}_n|l\rangle$$
$$= \langle \Pi^*_n W^*_n \mathfrak{F}_n k, (\pi\psi_m)\Pi^*_n W^*_n\mathfrak{F}_n l\rangle_{\mathfrak{A}}$$
$$= \left\langle \frac{1}{\sqrt{N}}\sum_{j'\in J_n} e^{-2\pi i k o(j')/N}\psi_{j'}, \right.$$
$$\left. \frac{1}{\sqrt{N}}\sum_{j\in J_n} e^{-2\pi i l o(j)/N} c_{mj}\psi_{m+j}\right\rangle_{\mathfrak{A}}$$
$$= \frac{1}{N}\sum_{j',j\in J_n} e^{2\pi i[ko(j')-lo(j)]/N} c_{mj}\delta_{j',m+j}$$
$$= \frac{1}{N}\sum_{j\in J_n} e^{2\pi i[ko(m+j)-lo(j)]} c^{(n)}_{mj},$$

where

$$c^{(n)}_{mj} = \begin{cases} c_{mj}, & m+j \in J_n, \\ 0, & \text{otherwise.}\end{cases}$$

Observe now that if $m + j \in J_n$, then $o(m + j) = m + o(j)$. Therefore, since $c^{(n)}_{mj} = 0$ whenever $m + j \notin J_n$, we get

$$(\tilde{A}_{m,n})_{kl} = \frac{1}{N}\sum_{j\in J_n} e^{2\pi i[(k-l)o(j)+km]/N} c^{(n)}_{mj}.$$

Thus, defining

$$\tilde{c}^{(n)}_{mj} = c^{(n)}_{mj} - e^{-\tau|m|^p/2}$$
$$= \begin{cases} e^{\tau(|m|^p+|j|^p-|m+j|^p)/2} - e^{-\tau|m|^p/2}, & m+j\in J_n, \\ -e^{-\tau|m|^p/2}, & \text{otherwise}\end{cases}$$

and

$$\varepsilon_{mnkl} = \frac{1}{N}\sum_{j\in J_n} e^{2\pi i[(k-l)o(j)+km]/N}\tilde{c}^{(n)}_{mj},$$

we get

$$(\tilde{A}_{m,n})_{kl} = \frac{1}{N}\sum_{j\in J_n} e^{2\pi i[(k-l)o(j)+km]/N} e^{-\tau|m|^p/2} + \varepsilon_{mnkl}$$
$$= \frac{1}{N}\sum_{q=0}^{N-1} e^{2\pi i[(k-l)q+km]/N} e^{-\tau|m|^p/2} + \varepsilon_{mnkl}$$
$$= e^{2\pi i k m/N} e^{-\tau|m|^p/2}\delta_{kl} + \varepsilon_{mnkl}.$$

Note that we used standard properties of discrete Fourier transforms to arrive at the last line. It then follows by definition of the ψ_m basis vectors and θ_l grid points that

$$(\tilde{A}_{m,n})_{kl} = \psi_m(\theta_l)\delta_{kl} + \varepsilon_{mnkl},$$

as claimed in the statement of the lemma.

We now proceed to bound the remainder ε_{mnkl}, assuming, for now, that $m \geqslant 0$. Letting $\tilde{N} = N/2$, we have

$$\begin{aligned}
|\varepsilon_{mnkl}| &= \left| \frac{1}{N} \sum_{j \in J_n} e^{2\pi i[(k-l)o(j)+km]/N} \tilde{c}_{mj}^{(n)} \right| \\
&\leqslant \frac{1}{N} \sum_{j \in J_n} \tilde{c}_{mj}^{(n)} \\
&= \frac{1}{N} \sum_{j=-\tilde{N}}^{-m} c_{mj}^{(n)} + \frac{1}{N} \sum_{j=-m+1}^{-1} e^{-\tau|m|^p} \\
&\quad + \frac{1}{N} \sum_{j=1}^{\tilde{N}-m} c_{mj}^{(n)} + \frac{1}{N} \sum_{\tilde{N}-m+1}^{\tilde{N}} e^{-\tau|m|^p} \\
&= \frac{(2|m|+1)e^{-\tau|m|^p}}{N} + \varepsilon_- + \varepsilon_+, \qquad \text{(C28)}
\end{aligned}$$

where

$$\varepsilon_- = \frac{1}{N} \sum_{j=-\tilde{N}}^{-m} c_{mj}^{(n)}, \quad \varepsilon_+ = \frac{1}{N} \sum_{j=1}^{\tilde{N}-m} c_{mj}^{(n)}.$$

Next, to bound the ε_+ term, consider the function $f(u) = u^p$. Since $p \in (0, 1)$, f is strictly concave on the positive real line. Thus, for $m \geqslant 0$ and $j \geqslant 1$, we have

$$\begin{aligned}
|m+j|^p - |j|^p &= |f(m+j) - f(j)| \\
&\leqslant |f'(j)||m| = pj^{p-1}|m|. \qquad \text{(C29)}
\end{aligned}$$

Consider also the function $g(u) = e^{\tau u/2} - 1$ on the interval $u \in [0, u_{\max}]$ with $u_{\max} = pm$. The function g is strictly convex, so

$$g(u) \leqslant g'(u_{\max})u = \frac{\tau}{2} e^{\tau u_{\max}/2} u = \frac{\tau}{2} e^{\tau pm/2} u.$$

Therefore, for $m \geqslant 0$ and $j \geqslant 1$, we obtain

$$\tilde{c}_{mj}^{(n)} = e^{-\tau|m|^p/2} g(f(m+j) - f(j)) \leqslant \tau p|m| j^{p-1}/2. \qquad \text{(C30)}$$

Note that we have used (C29) and the fact that $f(m+j) - f(j) \leqslant pm$ (which follows from the same equation).

Next, let $a_{\tilde{N}}$ be the series

$$a_{\tilde{N}} = \sum_{j=1}^{\tilde{N}} \left(\frac{j}{\tilde{N}}\right)^{p-1} \frac{1}{\tilde{N}}.$$

As $\tilde{N} \to \infty$, $a_{\tilde{N}}$ converges to the integral $\int_0^1 u^{p-1}\, du = 1/p$. Therefore, $a_{\tilde{N}}$ is bounded by a constant, $\tilde{C}$, leading to the bound

$$\frac{1}{\tilde{N}} \sum_{j=1}^{\tilde{N}} j^{p-1} = \tilde{N}^{p-1} \tilde{a}_N \leqslant \tilde{C} \tilde{N}^{p-1}. \qquad \text{(C31)}$$

Using (C30) and (C31), we thus obtain

$$\varepsilon_+ \leqslant \frac{1}{2\tilde{N}} \sum_{j=1}^{\tilde{N}} \tilde{c}_{mj}^{(n)} \leqslant \tilde{C} \tau p |m| \tilde{N}^{p-1}. \qquad \text{(C32)}$$

Moreover, analogous arguments for $j \leqslant -1$ lead to the estimate

$$\varepsilon_- \leqslant \hat{C} \tau p |m| \tilde{N}^{p-1}/2 \qquad \text{(C33)}$$

for a constant $\hat{C}$.

Substituting (C32) and (C33) into (C28), it follows that

$$\begin{aligned}
|\varepsilon_{mnkl}| &\leqslant \frac{(2|m|+1)e^{-\tau|m|^p}}{N} + \varepsilon_+ + \varepsilon_- \\
&\leqslant \frac{(2|m|+1)e^{-\tau|m|^p}}{N} + \frac{C\tau p|m|}{N^{1-p}}
\end{aligned}$$

with $C = \min\{\tilde{C}, \hat{C}\}/2^p$, which verifies the claim of the lemma for $m \geqslant 0$. However, since $\psi_{-m} = \psi_m^*$, repeating the calculation described above for $m < 0$ leads to the same bound, so we conclude that the claim holds for any $m \in \mathbb{Z}$. ■

5. Proof of Lemma 6

We have

$$\begin{aligned}
\langle \tilde{A}_{m,\tau,n} - D_{m,\tau,n} \rangle_{\tilde{\rho}_{x,\tau,n}} &= \mathrm{tr}(\tilde{\rho}_{x,\tau,n}(\tilde{A}_{m,\tau,n} - D_{m,\tau,n})) \\
&= \mathrm{tr}(\rho_{x,\tau,n}(\boldsymbol{\Pi}_{\tau,n}(\boldsymbol{\Pi}(\pi\psi_{m,\tau})) - \tilde{D}_{m,\tau,n})).
\end{aligned}$$

By the results in Sec. V and Appendix A, it follows that

$$\begin{aligned}
&\lim_{n\to\infty} \mathrm{tr}(\rho_{x,\tau,n}\boldsymbol{\Pi}_{\tau,n}(\boldsymbol{\Pi}_\tau(\pi_\tau\psi_{m,\tau}))) \\
&\quad = \mathrm{tr}(\rho_{x,\tau}\boldsymbol{\Pi}_\tau(\pi_\tau\psi_{m,\tau})) \\
&\quad = \frac{\eta_{m,\tau}}{\kappa_\tau}\psi_{m,\tau}(x) \\
&\quad = \frac{\sum_{j\in J'_m} e^{-\tau|j|^p}}{\kappa_\tau}\psi_{m,\tau}(x), \qquad \text{(C34)}
\end{aligned}$$

where we recall the definition of the index set J'_m,

$$J'_m = \{j \in J : j + m \in J\}.$$

Moreover, we have

$$\begin{aligned}
\mathrm{tr}(\rho_{x,\tau,n}\tilde{D}_{m,\tau,n}) &= \langle \xi_{x,\tau,n}, \tilde{D}_{m,\tau,n}\xi_{x,\tau,n}\rangle_{\mathcal{H}_\tau} \\
&= \frac{\langle k_{x,\tau,n}, \tilde{D}_{m,\tau,n})k_{x,\tau,n}\rangle_{\mathcal{H}_\tau}}{\kappa_{\tau,n}} \\
&= \frac{(\tilde{D}_{m,\tau,n}k_{x,\tau,n})(x)}{\kappa_{\tau,n}}.
\end{aligned}$$

In the above, the function $\tilde{D}_{m,\tau,n}k_{x,\tau,n} \in \mathcal{H}_{\tau,n}$ can be expressed as

$$\begin{aligned}
&\tilde{D}_{m,\tau,n}k_{x,\tau,n} \\
&\quad = W_n^*\mathfrak{F}_n D_{m,\tau,n}\mathfrak{F}_n^* W_n\left(\sum_{j\in J_n}\psi_{j,\tau}^*(x)\psi_{j,\tau}\right) \\
&\quad = W_n^*\mathfrak{F}_n D_{m,\tau,n}\mathfrak{F}_n^*\left(\sum_{j\in J_n}\psi_{j,\tau}^*(x)|o(j)\rangle\right) \\
&\quad = W_n^*\mathfrak{F}_n D_{m,\tau,n}\left(\frac{1}{\sqrt{N}}\sum_{l=0}^{N-1}\sum_{j\in J_n}\psi_{j,\tau}^*(x)e^{2\pi i o(j)l/N}|l\rangle\right) \\
&\quad = W_n^*\mathfrak{F}_n\left(\frac{1}{\sqrt{N}}\sum_{l=0}^{N-1}\sum_{j\in J_n}\psi_{j,\tau}^*(x)e^{2\pi i o(j)l/N}\psi_{m,\tau}(\theta_l)|l\rangle\right) \\
&\quad = W_n^*\left(\frac{1}{N}\sum_{k,l=0}^{N-1}\sum_{j\in J_n}\psi_{j,\tau}^*(x)e^{2\pi i[o(j)-k]l/N}\psi_{m,\tau}(\theta_l)|k\rangle\right) \\
&\quad = \frac{1}{N}\sum_{k,l=0}^{N-1}\sum_{j\in J_n}\psi_{j,\tau}^*(x)e^{2\pi i[o(j)-k]l/N}\psi_{m,\tau}(\theta_l)\psi_{o^{-1}(k)} \\
&\quad = \sum_{j,j'\in J_n}\psi_{j,\tau}^*(x)\psi_{j',\tau} \\
&\qquad \times\left(\frac{1}{N}\sum_{l=0}^{N-1}e^{i[o(j)-o(j')](2\pi l/N)}\psi_{m,\tau}(2\pi l/N)\right).
\end{aligned}$$

As $n\to\infty$, the summation in the parentheses in the last line converges to a continuous Fourier transform,

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{N}\sum_{l=0}^{N-1}e^{i[o(j)-o(j')](2\pi l/N)}\psi_{m,\tau}(2\pi l/N) \\
&\quad = \int_{S^1}e^{-i(j-j')\theta}\psi_{m,\tau}(\theta)\,d\theta \\
&\quad = e^{-\tau|m|^p/2}\int_{S^1}e^{i(j-j'+m)\theta}\,d\theta = e^{-\tau|m|^p/2}\delta_{j',j+m}.
\end{aligned}$$

As a result, we have

$$\begin{aligned}
&\lim_{n\to\infty}\frac{\tilde{D}_{m,\tau,n}k_{x,\tau,n}}{\kappa_{\tau,n}} \\
&\quad = \frac{1}{\kappa_\tau}\sum_{j,j'\in J}\psi_{j,\tau}^*(x)\psi_{j'}e^{-\tau|m|^p/2}\delta_{j',j+m} \\
&\quad = \frac{1}{\kappa_\tau}\sum_{j\in J_m'}\psi_{j,\tau}^*(x)\psi_{j+m}e^{-\tau|m|^p/2} \\
&\quad = \frac{1}{\kappa_\tau}\sum_{j\in J_m'}\psi_{j,\tau}^*(x)e^{-\tau|j+m|^p/2}\phi_{j+m}e^{-\tau|m|^p/2} \\
&\quad = \frac{1}{\kappa_\tau}\sum_{j\in J_m'}\psi_{j,\tau}^*(x)\psi_{j,\tau}e^{-\tau(|j+m|^p-|j|^p)/2}\psi_{m,\tau},
\end{aligned}$$

and upon evaluation at x,

$$\begin{aligned}
&\lim_{n\to\infty}\frac{(\tilde{D}_{m,\tau,n}k_{x,\tau,n})(x)}{\kappa_{\tau,n}} \\
&\quad = \frac{1}{\kappa_\tau}\sum_{j\in J_m'}e^{-\tau|j|^p}e^{-\tau(|j+m|^p-|j|^p)/2}\psi_{m,\tau}(x). \qquad \text{(C35)}
\end{aligned}$$

Therefore, combining (C34) and (C35), we obtain

$$\begin{aligned}
&\lim_{n\to\infty}\langle\tilde{A}_{m,\tau,n}-D_{m,\tau,n}\rangle_{\tilde{\rho}_{x,\tau,n}} \\
&\quad = \frac{1}{\kappa_\tau}\sum_{j\in J_m'}e^{-\tau|j|^p}\left(1-e^{-\tau(|j+m|^p-|j|^p)/2}\right)\psi_{m,\tau}(x),
\end{aligned}$$

and thus

$$\begin{aligned}
&\lim_{n\to\infty}|\langle\tilde{A}_{m,\tau,n}-D_{m,\tau,n}\rangle_{\tilde{\rho}_{x,\tau,n}}| \\
&\quad = \frac{e^{-\tau|m|^p/2}}{\kappa_\tau}\left|\sum_{j\in J_m'}e^{-\tau|j|^p}\left(1-e^{-\tau(|j+m|^p-|j|^p)/2}\right)\right| \\
&\quad \leqslant \frac{1}{\kappa_\tau}\sum_{j\in J_m'}e^{-\tau|j|^p}\left|1-e^{-\tau(|j+m|^p-|j|^p)/2}\right|.
\end{aligned}$$

Note now that for fixed $m\in\mathbb{Z}$, the largest value of $e^{-\tau|m|^p/2}|1-e^{-\tau(|j+m|^p-|j|^p)/2}|$ over $j\in\mathbb{Z}$ occurs for $|j|=|m|$. That is, we have

$$\begin{aligned}
&e^{-\tau|m|^p/2}\left|1-e^{-\tau(|j+m|^p-|j|^p)/2}\right| \\
&\quad \leqslant e^{-\tau|m|^p/2}\max\left\{\left|1-e^{\tau|m|^p/2}\right|,\left|1-e^{-\tau|m|^p/2}\right|\right\} \\
&\quad = \max\left\{\left|e^{-\tau|m|^p/2}-1\right|,e^{-\tau|m|^p/2}\left|1-e^{-\tau|m|^p/2}\right|\right\} \\
&\quad \leqslant \max\left\{\left|e^{-\tau|m|^p/2}-1\right|,\left|1-e^{-\tau|m|^p/2}\right|\right\} \\
&\quad = 1-e^{-\tau|m|^p/2},
\end{aligned}$$

so that

$$\begin{aligned}
\lim_{n\to\infty}|\langle\tilde{A}_{m,\tau,n}-D_{m,\tau,n}\rangle_{\tilde{\rho}_{x,\tau,n}}| &\leqslant \frac{1}{\kappa_\tau}\sum_{j\in J_m'}e^{-\tau|j|^p}\left(1-e^{-\tau|m|^p/2}\right) \\
&= \frac{\eta_{m,\tau}}{\kappa_\tau}\left(1-e^{-\tau|m|^p/2}\right) \\
&\leqslant 1-e^{-\tau|m|^p/2},
\end{aligned}$$

proving the lemma. ■

APPENDIX D: RESULTS FROM KOOPMAN OPERATOR THEORY

In this Appendix, we collect results and constructions from Koopman operator theory which are relevant to the discussion in the main text. We refer the reader to one of the many references in the literature, e.g., [44,45,91,92], for more detailed expositions of these topics.

1. Koopman operators on classical function spaces

As in the main text, we consider a continuous-time, continuous flow $\Phi^t: X\to X$, $t\in\mathbb{R}$, on a compact, metrizable space X. We denote the Borel σ-algebra and set of Borel probability measures on X by $\mathcal{B}(X)$ and $\mathcal{P}(X)$, respectively. We recall

that Φ^t has an induced action $\Phi^t_*: \mathcal{P}(X) \to \mathcal{P}(X)$ through the pushforward map on measures, $\Phi^t_*(\nu) = \nu \circ \Phi^{-t}$ [cf. (2)].

Let $\mathbb{A}$ be the vector space over the complex numbers consisting of all complex-valued functions on X. For every time $t \in \mathbb{R}$, the Koopman operator $U^t: \mathbb{A} \to \mathbb{A}$ acts by composition with the dynamical flow map, $U^t f := f \circ \Phi^t$. It follows from its definition that U^t is a linear operator, i.e., $U^t(f + cg) = U^t f + cU^t g$ for any $f, g \in \mathbb{A}$, $c \in \mathbb{C}$. Moreover, U^t is invertible by invertibility of Φ^t, with $(U^t)^{-1} = U^{-t}$. The Koopman operator U^t is also compatible with the structure of $\mathbb{A}$ as an Abelian *-algebra under pointwise function multiplication and complex conjugation. That is, for any $f, g \in \mathbb{A}$, we have

$$U^t(fg) = (U^t f)(U^t g), \quad (U^t f)^* = U^t(f^*). \tag{D1}$$

In other words, U^t acts as a *-isomorphism of $\mathbb{A}$.

In the setting of continuous flows, it is natural to restrict attention to the Banach space of continuous, complex-valued functions on X, $C(X) \subset \mathbb{A}$, on which U^t acts as an isometry, $\|U^t f\|_{C(X)} = \|f\|_{C(X)}$ with $\|f\|_{C(X)} = \max_{x \in X} |f(x)|$. The space $C(X)$ is a C^*-algebra under pointwise function multiplication and conjugation, satisfying

$$\|fg\|_{C(X)} \leqslant \|f\|_{C(X)} \|g\|_{C(X)}, \quad \|f^* f\|_{C(X)} = \|f\|^2_{C(X)}.$$

The Koopman operator is again compatible with this algebraic structure, satisfying (D1) for any $f, g \in C(X)$. In other words, U^t acts as an isometric *-isomorphism of the Abelian C^*-algebra $C(X)$.

Next, let $\mu \in \mathcal{P}(X)$ be a Borel probability measure. The flow Φ^t is said to be

(1) *μ-preserving* if $\Phi^t_*(\mu) = \mu$ for any time $t \in \mathbb{R}$

(2) *μ-ergodic* if for any set $A \in \mathcal{B}(X)$ that satisfies $\Phi^{-t}(A) \subseteq A$ we have that $\mu(A)$ is equal to either 0 or 1

(3) *μ-weak-mixing* if for any two sets $A, B \in \mathcal{B}(X)$ we have

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T |\mu(A \cap \Phi^{-t}(A)) - \mu(A)\mu(B)|\, dt = 0. \tag{D2}$$

It should be kept in mind that μ-preservation and μ-ergodicity are independent notions. Intuitively, measure preservation means that the flow preserves the "size" or "volume" of sets assigned by μ. On the other hand, ergodicity is a form of indecomposability that asserts that a dynamical system has no nontrivial invariant sets with respect to a probability measure (here, μ), irrespective of whether that measure is invariant or not. It is also worthwhile noting that ergodicity, in itself, does not imply that the orbits of the dynamics are dense, i.e., that given an initial condition $x \in X$ the orbit $\mathcal{O}_x := \{\Phi^t(x)\}_{t \in \mathbb{R}}$ samples the state space densely in a topological sense. However, what *can* be said is that for μ-almost every $x \in X$, the orbit $\mathcal{O}_x$ is a dense subset of the support of μ. Intuitively, this means that if $x \in X$ is an initial condition drawn randomly with distribution μ, then with probability 1, the orbit $\mathcal{O}_x$ will well sample any subset in X which has nonzero probability of occurring (again with respect to μ) under the dynamics. Meanwhile, weak mixing encompasses the notion that, under dynamical evolution, sets become statistically independent in a time-averaged sense. This behavior is consistent with chaotic dynamics, but one should keep in mind that mathematically weak mixing is an independent notion from topological definitions of chaos, e.g., [93].

Associated with any Borel probability measure μ are the standard $L^p(\mu)$ spaces of (equivalence classes of) complex-valued functions on X, equipped with the norms $\|f\|_{L^p(\mu)} = (\int_X |f|^p\, d\mu)^{1/p}$ for $1 \leqslant p < \infty$ and $\|f\|_{L^\infty(\mu)} = \lim_{p \to \infty} \|f\|_{L^p(\mu)}$. As in the main text, we use the notation $\langle f, g \rangle_{L^2(\mu)} = \int_X f^* g\, d\mu$ to represent the $L^2(\mu)$ inner product. Note that because μ is a finite measure, we have $L^q(\mu) \subseteq L^p(\mu)$ for all $1 \leqslant p \leqslant q \leqslant \infty$.

For the rest of this Appendix, we will assume that Φ^t is μ-preserving. Then, the Koopman operator is well defined as an isometry $U^t: L^p(\mu) \to L^p(\mu)$ for all $1 \leqslant p \leqslant \infty$. In the Hilbert space case, $p = 2$, U^t is an unitary operator satisfying $U^{t*} = U^{-t}$. In what follows, we will let $\mathbb{F}$ stand for any of the $C(X)$ or $L^p(\mu)$ spaces with $1 \leqslant p \leqslant \infty$, and we will let $\mathbb{F}_0$ stand for any of $C(X)$ or $L^p(\mu)$ with $1 \leqslant p < \infty$. Note that all the $\mathbb{F}_0$ spaces are separable Banach spaces (by the assumed compactness and metrizability of X), but unless X is a finite set, $L^\infty(\mu)$ is not separable.

A cornerstone result in measure-preserving, ergodic dynamical systems is the Birkhoff pointwise ergodic theorem. It states that if Φ^t is μ-preserving and μ-ergodic, then for every observable $f \in L^1(\mu)$ and μ-almost every (a.e.) initial condition $x \in X$, the equality

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T f(\Phi^t(x))\, dt = \int_X f\, d\mu \tag{D3}$$

holds. The pointwise ergodic theorem thus asserts that expectation values (spatial averages) of observables are equal to time averages along typical orbits of the system. In fact, an analog of (D3) also holds for discrete-time subsamplings of the continuous-time flow Φ^t. For any sampling interval $\Delta t > 0$, we let $\hat{\Phi}: X \to X$ be the discrete-time map with $\hat{\Phi} = \Phi^{\Delta t}$. The map $\hat{\Phi}$ is μ-preserving by definition. Moreover, it is μ-ergodic for Lebesgue a.e. sampling interval Δt, satisfying

$$\lim_{\mathcal{N} \to \infty} \frac{1}{\mathcal{N}} \sum_{i=0}^{\mathcal{N}-1} f(\hat{\Phi}^i(x))\, dt = \int_X f\, d\mu. \tag{D4}$$

Besides its theoretical significance, (D4) provides a foundation for many data-driven techniques (as will be discussed in more detail in Appendix E), for it enables approximation of expectation values by finite trajectories sampled in discrete time.

2. Eigenvalues of Koopman operators

Consider the eigenvalue equation for the Koopman operator,

$$U^t \phi = \Lambda_t \phi,$$

where $\phi: X \to \mathbb{C}$ is a nonzero function, and $\Lambda_t \in \mathbb{C}$ an eigenvalue. Without imposing any restrictions on the membership of ϕ in a particular function space, we can deduce from its definition that U^t admits as an eigenfunction the constant function $\phi(x) = 1$ and the corresponding eigenvalue is $\Lambda_t = 1$. Since constant functions lie in all of the $C(X)$ and $L^p(\mu)$ spaces, it follows that the number 1 is an eigenvalue of U^t on $\mathbb{F}$. A fundamental result is that the multiplicity

of that eigenvalue provides an equivalent characterization of ergodicity. In particular, a μ-preserving flow $\Phi^t: X \to X$ is μ-ergodic, if and only if the only functions $\phi \in L^1(\mu)$ which are simultaneously eigenfunctions of U^t for all $t \in \mathbb{R}$ corresponding to eigenvalue $\Lambda_t = 1$ are constants.

Besides providing a spectral characterization of ergodicity, Koopman eigenvalues and eigenfunctions have useful algebraic properties which we have exploited in several instances in the main text:

(1) It follows by definition of U^t that if $\phi, \tilde{\phi} \in \mathbb{A}$ are eigenfunctions corresponding to the eigenvalues $\Lambda_t, \tilde{\Lambda}_t \in \mathbb{C}$, respectively, then $\phi\tilde{\phi}$ is also an eigenfunction, corresponding to eigenvalue $\Lambda_t \tilde{\Lambda}_t$. Applying this result for $\tilde{\phi} = \phi$, we deduce that ϕ^2 is an eigenfunction corresponding to Λ_t^2, whereas the choice $\tilde{\phi} = \phi^*$ yields that $|\phi|^2$ is an eigenfunction corresponding to $|\Lambda_t|^2$.

(2) Since $(U^t\phi)^* = U^t(\phi^*)$, it follows that ϕ^* is an eigenfunction of U^t corresponding to the eigenvalue Λ_t^*.

(3) Since $\|U^t f\|_{\mathbb{F}} = \|f\|_{\mathbb{F}}$ for any $f \in \mathbb{F}$, we can also deduce that all eigenvalues Λ_t lie in the unit circle in the complex plane, $\mathbb{T}^1 \subset \mathbb{C}$, and thus if Λ_t is an eigenvalue, then so is $\Lambda_t^{-1} = \Lambda_t^*$. Therefore, we conclude that the eigenvalues of U^t form a multiplicative subgroup of $\mathbb{T}^1$.

Let $\sigma_e(U^t; \mathbb{F})$ denote the set of eigenvalues of U^t on a Banach space $\mathbb{F}$. In the case of the Koopman operator on $L^2(\mu)$, the spectral theorem for unitary operators implies that all eigenfunctions corresponding to distinct eigenvalues are orthogonal. Thus, by separability of $L^2(\mu)$, it follows that $\sigma_e(U^t; L^2(\mu))$ is a countable set. Moreover, it can be shown [45, Proposition 7.18] that $\sigma_e(U^t; L^p(\mu)) = \sigma_e(U^t; L^q(\mu))$ for all $p, q \in [0, \infty]$. One also readily verifies that $\sigma_e(U^t; C(X)) \subseteq \sigma_e(U^t; L^p(\mu))$, so we conclude that U^t has countably many eigenvalues on any of the $\mathbb{F}$ spaces.

Next, suppose that $\phi \in L^2(\mu)$ is an eigenfunction of U^t corresponding to the eigenvalue Λ_t. Then, $|\phi|^2$ lies in $L^1(\mu)$, and using (D1) we get

$$U^t(|\phi|^2) = U^t(\phi^*\phi) = (U^t\phi^*)(U^t\phi) = (U^t\phi)^*(U^t\phi)$$
$$= (\Lambda_t\phi)^*(\Lambda_t\phi) = |\Lambda_t|^2|\phi|^2 = |\phi|^2.$$

It therefore follows that $|\phi|^2 \in L^1(\mu)$ is an eigenfunction of U^t corresponding to the eigenvalue $\Lambda_t = 1$. In particular, if the system is μ-ergodic, this implies in turn that ϕ can be chosen such that it takes values on the unit circle, i.e., $|\phi| = 1$. This choice implies in turn that ϕ is a unit vector in $L^2(\mu)$, i.e., $\|\phi\|_{L^2(\mu)} = 1$.

Further useful properties follow by continuity of the flow Φ^t with respect to time t. The latter, implies that for any $f \in \mathbb{F}_0$ the map $t \mapsto U^t f$ is continuous in the norm of $\mathbb{F}_0$ [but *not* in the norm of $L^\infty(\mu)$]. This in turn implies that the operator group $\{U^t\}_{t\in\mathbb{R}}$ on $\mathbb{F}_0$ has a generator, i.e., a (generally, unbounded) operator $V: D(V) \to \mathbb{F}_0$ defined on a dense subspace $D(V) \subseteq \mathbb{F}_0$, satisfying

$$\frac{d}{dt}U^t f = VU^t f = U^t V f \tag{D5}$$

for every observable $f \in D(V)$. The generator V can be explicitly obtained by taking the limit

$$Vf = \lim_{t\to 0} \frac{U^t f - f}{t} \tag{D6}$$

in the norm of $\mathbb{F}_0$, as done in (6) for $\mathbb{F}_0 = L^2(\mu)$. Note that (D6) is the defining equation for the domain of the generator, i.e., the domain $D(V)$ consists of all $f \in \mathbb{F}_0$ for which the limit in (D6) exists. The generator completely characterizes the Koopman group on $\mathbb{F}_0$, in the sense that given V, the Koopman operator U^t at any time t can be reconstructed by solving the differential equation (D5); symbolically we write $U^t = e^{tV}$ [cf. (7)].

3. Spectral properties of Koopman operators on $L^2(\mu)$

In this subsection, we focus on Koopman operators on $L^2(\mu)$, whose spectra can be characterized using Hilbert space techniques. By Stone's theorem on strongly continuous unitary evolution groups [76], the generator $V: D(V) \to L^2(\mu)$ of the Koopman group on $L^2(\mu)$ is skew adjoint, $V^* = -V$, which implies that its spectrum is a subset of the imaginary line, $i\mathbb{R}$. By the spectral mapping theorem, for every element $z \in i\mathbb{R}$ of the spectrum of V, $e^{zt} \in \mathbb{T}^1$ is an element of the spectrum of U^t lying in the unit circle. That is, if $z = i\omega$ is an eigenvalue of V (associated with the eigenfrequency $\omega \in \mathbb{R}$) and $\phi \in L^2(\mu)$ is a corresponding eigenfunction, then ϕ is also an eigenfunction of U^t corresponding to the eigenvalue $\Lambda_t = e^{i\omega t}$. We will denote the set of eigenvalues of V by $\sigma_e(V) \equiv \sigma_e(V; L^2(\mu))$. We also let H_e be the Hilbert subspace of $L^2(\mu)$ spanned by the corresponding eigenfunctions. That is, H_e has an orthonormal basis $\{\phi_j\}$ satisfying

$$V\phi_j = i\omega_j\phi_j, \quad i\omega_j \in \sigma_e(V).$$

Since every basis vector ϕ_j is an eigenvector of U^t corresponding to eigenvalue $e^{i\omega_j t}$, it follows that every element $f = \sum_j \hat{f}_j\phi_j$ of H_e, with $\hat{f}_j \in \mathbb{C}$, satisfies

$$U^t f = \sum_j e^{i\omega_j t}\hat{f}_j\phi_j. \tag{D7}$$

The μ-preserving flow $\Phi^t: X \to X$ is said to have *pure point spectrum* if the generator V on $L^2(\mu)$ is diagonalizable, i.e., $H_e = L^2(\mu)$. Two important characteristics of pure point spectrum systems are as follows:

(1) The generator equation (D5) governing the evolution of observables in $L^2(\mu)$ is integrable and has (D7) as its solution.

(2) The evolution of observables does not exhibit long-term decay of correlations. In particular, for $f, g \in L^2(\mu)$ with $f = \sum_j \hat{f}_j\phi_j$ and $g = \sum_j \hat{g}_j\phi_j$, we can define the cross-correlation function $C_{fg}: \mathbb{R} \to \mathbb{C}$ by

$$C_{fg}(t) = \langle f, U^t g\rangle_{L^2(\mu)} = \int_X f^*(x)g(\Phi^t(x))\,d\mu(x), \tag{D8}$$

and it follows from (D7) that for a pure point spectrum system $C_{fg}(t)$ has an oscillatory behavior,

$$C_{fg}(t) = \sum_j e^{i\omega_j t}\hat{f}_j^*\hat{g}_j, \tag{D9}$$

which does not decay to zero as $t \to \infty$.

Next, by definition of V [see (6) and (D6)], 0 is always an eigenvalue with a constant corresponding eigenfunction, $\phi(x) = 1$. It is a fundamental result that the simplicity of this eigenvalue provides a characterization of ergodicity, stated

in stated in terms of the generator: A μ-preserving flow is μ-ergodic if and only if the 0 eigenvalue of the generator on $L^2(\mu)$ is simple (i.e., the nullspace of the generator is one-dimensional and consists of constant functions). In addition, we have an equivalent, spectral characterization of weak mixing: A μ-preserving flow is μ-weak mixing if and only if 0 is the only eigenvalue of the generator. Therefore, for weak-mixing systems, the generator on L^2 (and thus the associated Koopman operators) are manifestly nondiagonalizable, and have nontrivial continuous spectrum. Moreover, it can be shown that observables of weak-mixing systems exhibit a long-term loss of correlation in the following time-averaged sense:

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T |C_{fg}(t)-\bar{f}^*\bar{g}|\,dt=0, \tag{D10}$$

where $\bar{f}=\int_X f\,d\mu$ and $\bar{g}=\int_X g\,d\mu$. As with (D2), the above is consistent with what one would expect for the statistical behavior of chaotic systems.

As a third type of ergodic behavior, we mention the so-called *mixed-spectrum systems*. These systems sit between pure-point-spectrum systems and weak-mixing systems in the sense that the generator has nonconstant eigenfunctions, but the point spectrum subspace H_e is a strict subspace of $L^2(\mu)$. Defining $H_c=H_e^{\perp}$ as the orthogonal complement of H_e in $L^2(\mu)$ associated with the continuous spectrum of V, we have the Koopman-invariant orthogonal decomposition

$$L^2(\mu)=H_e\oplus H_c.$$

Thus, for a mixed-spectrum system, every observable $f\in L^2(\mu)$ has the orthogonal decomposition $f=f_e+f_c$, where $f_e\in H_e$ has the integrable evolution from (D7) and oscillatory cross-correlation from (D9) with any $g\in H_e$, whereas $f_c\in H_c$ exhibits the loss of correlation in (D10) for any $g\in L^2(\mu)$.

We close this subsection by describing how V behaves analogously to a first-order differential operator (even though the state space X need not have differentiable structure), and examining the group structure of the point spectrum $\sigma_e(V)$ that can be deduced from this property. In particular, it can be shown that $\mathfrak{V}:=D(V)\cap L^\infty(\mu)$ is an algebra with respect to pointwise multiplication of functions, and V acts on this algebra as a derivation, i.e., a linear map obeying the Leibniz rule,

$$V(fg)=(Vf)g+f(Vg),\quad \forall f,g\in\mathfrak{V}. \tag{D11}$$

Recently, it has been shown [94] that (D11) is a necessary and sufficient condition for a skew-adjoint operator V to be the generator of a group of unitary Koopman operators on $L^2(\mu)$. That is, we have $U^t=f\circ\Phi^t$ for a dynamical flow $\Phi^t{:}X\to X$ if and only if (D11) holds. Since vector fields on manifolds can be identified with first-order differentiation operators on functions, we can intuitively think of V as a generalization of a vector field field generating a dynamical system on a differentiable manifold. This intuition can be made precise in the case that X *is* a differentiable manifold (with tangent bundle TX), and $\Phi^t{:}X\to X$ is generated by a vector field $\vec{V}{:}X\to TX$, i.e., the orbit $x(t)=\Phi^t(x)$ is the solution of the initial-value problem

$$\dot{x}(t)=\vec{V}(x(t)),\quad x(0)=x.$$

In this setting, we can identify $\vec{V}$ with the directional derivative operator $\mathcal{V}{:}C^1(X)\to C(X)$ acting on continuously differentiable functions on X as $\mathcal{V}f=\vec{V}\cdot\nabla f$. Then, the generator V is an extension of the vector field, in the sense that $C^1(X)$ is a subspace of the domain of V, and $Vf=\mathcal{V}f$ for $C^1(X)$ functions.

Using (D11), we can deduce that the eigenvalues of V form an additive group (in contrast to the eigenvalues of U^t, which form a multiplicative group). That is, if $i\omega$ and $i\tilde{\omega}$ are eigenvalues of V corresponding to eigenfunctions ϕ and $\tilde{\phi}$, respectively, then $i(\omega+\tilde{\omega})$ is also an eigenvalue, corresponding to the eigenfunction $\phi\tilde{\phi}$. An implication of this fact is that if the set of eigenvalues $\sigma_e(V)$ contains two rationally independent elements (i.e., two incommensurate eigenfrequencies), then it is a dense subset of the imaginary line. In that case (since the spectrum of an operator on a Banach space includes the closure of the set of its eigenvalues), the spectrum of V on $L^2(\mu)$ is the entire imaginary line.

An important consequence of the group structure of the point spectrum of V is that it admits *generating sets*. Specifically, a set $iA\subseteq\sigma_e(V)$ is a generating set if for every $i\omega\in\sigma_e(V)$ there exists a finite collection of eigenfrequencies $\{\alpha_1,\ldots,\alpha_q\}\subseteq A$ and integers $j_1,\ldots,j_q$ such that $\omega=\sum_{k=1}^q j_k\alpha_k$. The *rank* of $\sigma_e(V)$, denoted by $\mathrm{rank}\sigma_e(V)$, is defined as the minimal number of elements of its generating sets.

Characterizing the rank of the point spectra of measure-preserving ergodic flows is a challenging problem which has not been completely solved. It is known that if the state space X is a differentiable closed manifold, then the rank of the point spectrum of any smooth measure-preserving ergodic flow $\Phi^t{:}X\to X$ with differentiable Koopman eigenfunctions is bounded by the dimension of X [95]. Meanwhile, smooth, measure-preserving, ergodic flows on manifolds with discontinuous Koopman eigenfunctions and point spectra of arbitrarily large rank, including infinity, can been constructed (as suspension flows generated by suitable diffeomorphisms [96]). The form of these systems is, however, not typical of dynamical systems encountered in physical applications.

4. Conjugacy with torus rotations

We now come to an important property of Koopman eigenfunctions of measure-preserving, ergodic dynamical systems with pure point spectrum, namely, that they provide measure-theoretic, and if continuous, topological isomorphisms with ergodic rotations on tori.

First, let $\phi\in L^2(\mu)$ be an eigenfunction of the generator corresponding to a nonzero eigenfrequency $\alpha\in\mathbb{R}$. As in Sec. D 2, we normalize ϕ such that $|\phi|=1$, so that we can view $\phi{:}X\to\mathbb{T}^1$ as a map from X into the unit circle in the complex plane, $\mathbb{T}^1\subset\mathbb{C}$. Let $R_\alpha^t{:}\mathbb{T}^1\to\mathbb{T}^1$ be the circle rotation with frequency α, i.e.,

$$R_\alpha^t(\theta)=\theta+\alpha t\mod 2\pi,$$

where θ is the argument (phase angle) of the complex number $e^{i\theta}$. Note that $e^{iR_\alpha^t(\theta)}=e^{i\alpha t}e^{i\theta}$, so R_α^t can be identified with the multiplicative action $z\mapsto e^{i\alpha t}z$ of the complex number $e^{i\alpha t}$ on $\mathbb{C}$. Then, since $U^t\phi=\phi\circ\Phi^t=e^{i\alpha t}\phi$, it follows that for

μ-a.e. $x \in X$,

$$\phi(\Phi^t(x)) = R^t_\alpha(\phi(x)). \tag{D12}$$

Thus, the following diagram commutes in a measure-theoretic sense (μ-a.e.):

$$\begin{array}{ccc} X & \xrightarrow{\Phi^t} & X \\ {\scriptstyle\phi}\downarrow & & \downarrow{\scriptstyle\phi} \\ \mathbb{T}^1 & \xrightarrow{R^t_\alpha} & \mathbb{T}^1 \end{array}.$$

If, in addition, ϕ is continuous, then $\phi : X \to \mathbb{C}$ is a surjective (onto) map, and (D12) holds for every $x \in X$. Such a continuous map ϕ which intertwines the flows Φ^t and R^t_α is called a topological semiconjugacy. In particular, for each point $x \in X$, ϕ provides a corresponding angle coordinate through its argument,

$$\theta(x) = \arg(\phi(x)) \in [0, 2\pi). \tag{D13}$$

We can extend this construction to higher dimensions by considering multiple eigenfunctions and their corresponding eigenfrequencies. Specifically, each collection $\vec{\alpha} = (\alpha_1, \ldots, \alpha_d)$ of distinct nonzero eigenfrequencies $\alpha_j \in \sigma_e(V)$ induces a map $\phi_{\vec{\alpha}}: X \to \mathbb{T}^d$ with

$$\phi_{\vec{\alpha}}(x) = (\phi_1(x), \ldots, \phi_d(x)),$$

where $\phi_j \in L^2(\mu)$ is an eigenfunction corresponding to ϕ_j (normalized such that $|\phi_j| = 1$). This map provides a measure-theoretic semiconjugacy between Φ^t and the torus rotation $R^t_{\vec{\alpha}}: \mathbb{T}^d \to \mathbb{T}^d$ with rotation frequencies $\vec{\alpha}$, i.e.,

$$\phi_{\vec{\alpha}}(\Phi^t(x)) = R^t_{\vec{\alpha}}(\phi_{\vec{\alpha}}(x)) \tag{D14}$$

for μ-a.e. $x \in X$, where

$$R^t_{\vec{\alpha}}(\theta^1, \ldots, \theta^d) = (\theta^1 + \alpha_1 t, \ldots, \theta^d + \alpha_d t) \mod 2\pi.$$

Moreover, $\phi_{\vec{\alpha}}$ becomes a topological semiconjugacy if $\phi_1, \ldots \phi_d$ are continuous, and it is an onto map if, in addition, the $\alpha_1, \ldots, \alpha_d$ are rationally independent. In the latter case, the ϕ_j induce d canonical angle coordinates θ_j on X analogously to (D13).

By a classical result of von Neumann [97], for a pure point spectrum system with d generating frequencies $\vec{\alpha} = (\alpha_1, \ldots, \alpha_d)$ [i.e., $\text{rank}\sigma_e(V) = d$], the map $\phi_{\vec{\alpha}}: X \to \mathbb{T}^d$ is a measure-theoretic isomorphism, i.e., it is an invertible, measure-preserving transformation, compatible with the dynamics in the sense of (D14). If, in addition, the associated eigenfunctions ϕ_j are continuous, then $\phi_{\vec{\alpha}}$ is a topological isomorphism (conjugacy), i.e., it is a continuous map with continuous inverse, satisfying the dynamical compatibility condition in (D14) for every $x \in X$.

Since the construction of the QECD scheme described in the main text is based on spaces of continuous functions on X, for our purposes a pure point spectrum system with finitely generated spectrum and continuous corresponding eigenfunctions can be identified with an ergodic rotation on a torus. It is important to note that the transformation $\phi_{\vec{\alpha}}: X \to \mathbb{T}^d$ is based entirely on intrinsic spectral objects (i.e., eigenfunctions and eigenvalues of V), and and does not require *a priori* knowledge of coordinates on X. In particular, the Koopman eigenfunctions ϕ_j *define* canonical angle coordinates θ_j via (D13), which one can then use for other purposes (e.g., to define kernels as in Sec. III B).

It should be kept in mind that by assuming that (1) the eigenvalue spectrum has finite rank and (2) the Koopman eigenfunctions are continuous we are not treating the most general class of systems with pure point spectrum (see Appendix D 2). These assumptions could be potentially relaxed at the expense of increasing the technical complexity of the analysis [if $\sigma_e(V)$ has infinite rank] and/or weakening the form of convergence in the infinite qubit limit (e.g., from uniform convergence with respect to $x \in X$ to μ-a.e convergence if the eigenfunctions are discontinuous). Arguably, however, typical systems encountered in applications have finite-rank eigenvalue spectra and continuous eigenfunctions, so our assumptions are not too restrictive.

APPENDIX E: NUMERICAL APPROXIMATION SCHEMES

There are several mature algorithms for data-driven approximation of the eigenvalues and eigenfunctions of the Koopman operator. Examples include Fourier-based techniques [27,39], dynamic mode decomposition (DMD) [40,98], extended DMD [41], and RKHS-based techniques [24,99]. There is also extensive literature on spectral approximation of the transfer operator [38,42], which is dual to the Koopman operator and acts on Banach spaces of measures (see Sec. II A). Note that if $\mathbb{H}$ is a Hilbert space of observables on which the Koopman operator acts as a unitary map $U^t: \mathbb{H} \to \mathbb{H}$, then the transfer operator can be identified with the adjoint of U^t, i.e., $U^{t*} = U^{-t}$. In such cases, which include the spaces $\mathbb{H} = L^2(\mu)$ and $\mathbb{H} = \mathfrak{A}$ employed in the main text, working with the Koopman vs. transfer operator is merely a matter of convention.

In this Appendix, we provide a brief outline of a class of methods for spectral approximation of Koopman/transfer operators for measure-preserving, ergodic systems [23,25,29,100,101] which are based on kernel techniques and should thus be well-suited for integration with the quantum computational techniques described in this paper.

We consider that available to us is a time series $y_0, y_1, \ldots, y_{\mathcal{N}-1}$ of observations taken along an orbit of the dynamical system under a map $Y: X \to \mathbb{R}^m$. That is, we have $y_i = Y(x_i)$, where $x_i = \Phi^{i\,\Delta t}(x_0) \in X$ are (unobserved) states, $x_0 \in X$ is an arbitrary initial condition, and $\Delta t > 0$ is a fixed sampling interval. We assume throughout that $\Phi^t: X \to X$ is a measure-preserving, ergodic flow on a compact, differentiable manifold X for a Borel probability measure μ (with compact support by compactness of X). Given the data $\{y_i\}_{i=0}^{\mathcal{N}-1}$, we compute a collection of eigenvectors $\boldsymbol{\phi}_0, \boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_{L-1} \in \mathbb{C}^{\mathcal{N}}$ with $L \leqslant \mathcal{N}$ and associated eigenfrequencies $\hat{\omega}_0, \ldots, \hat{\omega}_{L-1}$ which are approximations of the eigenvalues and eigenfrequencies of the Koopman generator V on $L^2(\mu)$.

In particular, the eigenvectors $\boldsymbol{\phi}_j = (\phi_{0j}, \ldots, \phi_{\mathcal{N}-1,j})^\top$ and eigenfrequencies $\hat{\omega}_j$ have the properties that

$$\phi_{ij} \approx \phi_j(x_i), \quad \hat{\omega}_j \approx \omega_j, \tag{E1}$$

where ϕ_j and ω_j are eigenfunctions of the generator, i.e., $V\phi_j = i\omega_j\phi_j$. Under the assumption that the eigenfunctions ϕ_j are continuous, the convergence of this approximation

holds in the limit of large data, $\Delta t \to 0$ after $\mathcal{N} \to \infty$, in the $C(X)$ (uniform) norm. Importantly, the computation of the $\boldsymbol{\phi}_j$ and $\hat{\omega}_j$ is entirely based on the time series data y_i, and does not require a priori knowledge of coordinates for X. It should also be noted that the methods described below for a single training time series $y_0, y_1, \ldots, y_{\mathcal{N}-1}$ can be readily generalized to ensembles of time series with equidistributed initial conditions with respect to μ.

1. Data-driven basis

To compute the eigenvectors $\boldsymbol{\phi}_j$, we first compute a set of basis vectors $\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{L-1} \in \mathbb{C}^{\mathcal{N}}$ by solving the eigenvalue problem for an $\mathcal{N} \times \mathcal{N}$ kernel matrix $\boldsymbol{K}$ with positive spectrum, constructed from the data y_i. That is, we solve the eigenvalue problem

$$\boldsymbol{K}\boldsymbol{u}_l = \hat{\Lambda}_l \boldsymbol{u}_l,$$

where the eigenvalues $\hat{\Lambda}_l$ are positive, and the entries $K_{ij} = k^{(Y)}(y_i, y_j)$ of $\boldsymbol{K}$ are obtained from a kernel function $k^{(Y)}: \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ on data space. The basic requirements on the kernel $k^{(Y)}$ are that as $\mathcal{N} \to \infty$ the eigenvectors $\boldsymbol{u}_j$ should converge to an orthonormal basis $\{u_0, u_1, \ldots\}$ of $L^2(\mu)$, such that each basis vector u_j lies in the domain of the generator V and the corresponding eigenvalues $\hat{\Lambda}_j$ have a strictly positive limit Λ_j. In addition, we will normalize $\boldsymbol{K}$ such that it is a row-stochastic matrix with strictly positive elements (i.e., $K_{ij} > 0$ and $\sum_{j=0}^{\mathcal{N}-1} K_{ij} = 1$). This ensures that $\hat{\Lambda}_l \in [0, 1]$ and there is a simple eigenvalue $\hat{\Lambda}_0 = 1$ corresponding to the constant eigenvector $\boldsymbol{u}_0 = (1, 1, \ldots, 1)^\top \in \mathbb{C}^{\mathcal{N}}$. By convention, we order the eigenvalues $\hat{\Lambda}_l$ in decreasing order, i.e., $0 = \hat{\Lambda}_0 > \hat{\Lambda}_1 \geqslant \hat{\Lambda}_2 \geqslant \cdots \geqslant \hat{\Lambda}_{\mathcal{N}-1}$.

There are several commonly used kernels in the machine learning literature that meet the requirements stated above (see Refs. [25,29]), but as a concrete example we mention here the class of Markov-normalized radial basis function kernels proposed in the diffusion maps algorithm [102]:

$$k^{(Y)}(y, y') = \frac{\tilde{k}^{(Y)}(y, y')}{w_2(y) w_1(y')}. \tag{E2}$$

Here,

$$\tilde{k}^{(Y)}(y, y') = \exp\left(-\frac{\|y - y'\|_2^2}{\epsilon^2}\right)$$

is the radial basis function kernel with bandwidth parameter $\epsilon > 0$, and $w_1, w_2: X \to \mathbb{R}$ are "right" and "left" normalization functions, respectively, defined by

$$w_1(y') = \sum_{j=0}^{\mathcal{N}-1} k^{(Y)}(y', y_j), \qquad w_2(y) = \sum_{j=0}^{\mathcal{N}-1} \frac{k^{(Y)}(y, y_j)}{w_1(y_j)}.$$

As in the case of the generator eigenvectors $\boldsymbol{\phi}_j$, we think of $\boldsymbol{u}_l = (u_{0l}, \ldots, u_{\mathcal{N}-1,l})^\top$ as representing the values $\hat{u}_l(x_i)$ of a function $\hat{u}_l: X \to \mathbb{R}$ sampled along the trajectory $x_0, \ldots, x_{\mathcal{N}-1}$. If desired, the function $\hat{u}_l$ can be constructed by means of a suitable out-of-sample extension technique, such as the Nyström method (e.g., [103]).

The asymptotic behavior of the eigenvalues and eigenvectors of $\boldsymbol{K}$ as the dataset size $\mathcal{N}$ increases can be characterized through a diverse range of approaches available in the literature [104–110]. At a general level, if the kernel $k^{(Y)}$ is continuous, there is an associated compact integral operator $\mathcal{K}: C(X) \to C(X)$ defined by

$$\mathcal{K}f = \int_X k^{(Y)}(Y(\cdot), Y(x)) f(x)\, d\mu(x). \tag{E3}$$

Then, using the results of Ref. [106] in conjunction with ergodicity, it can be shown that for every nonzero eigenvalue Λ_l of $\mathcal{K}$ there is a sequence of eigenvalues $\hat{\Lambda}_l$ of $\boldsymbol{K}$ that converges to it as $\mathcal{N} \to \infty$ (including multiplicities), and there is an associated notion of convergence of the eigenvectors $\boldsymbol{u}_l$ to eigenfunctions $u_l \in C(X)$. The eigenfunctions u_l corresponding to nonzero eigenvalues are continuously differentiable if $k^{(Y)}$ is a C^1 kernel (and thus u_l lies in the domain of the generator), and can be chosen to be orthonormal in $L^2(\mu)$. Moreover, if all eigenvalues Λ_l are strictly positive, the u_l form an orthonormal basis of $L^2(\mu)$. The strict positivity of Λ_l holds, e.g., if Y is an injective map and $k^{(Y)}$ is the normalized Gaussian kernel from (E2). If Y is not injective, then Takens embedding theory [111] can be employed to construct an injective map $\tilde{Y}: X \to \mathbb{R}^{qm}$ (q here being the number of delays), which can in turn be used to build kernels with strictly positive corresponding eigenvalues, e.g., [100].

Under the additional assumption that the support $M \subseteq X$ of μ is a Riemannian manifold, more specialized results [104,105,107–110] connect the asymptotic spectral behavior of $\mathcal{K}$ in the limit of vanishing bandwidth parameter, $\epsilon \to 0$, with the spectrum of the Laplace-Beltrami operator, Δ, on M. In such cases, the eigenfunctions u_l converge as $\epsilon \to 0$ to eigenfunctions of Δ, which are extremizers of the Rayleigh quotient $\mathcal{E}(f)/\|f\|_{L^2(\mu)}$ associated with the Dirichlet energy functional

$$\mathcal{E}(f) = \int_M \|\nabla f\|^2\, d\mu. \tag{E4}$$

Here, ∇ and $\|\cdot\|$ are the Riemannian gradient and norm on tangent vectors, respectively. Intuitively, we think of $\mathcal{E}(f)$ as a measure of roughness of functions induced by the Riemannian metric of M. Correspondingly, the basis $\{u_0, u_1, \ldots\}$ can be thought of as having optimal regularity, in the sense that for any $l \in \mathbb{N}$, u_l has the smallest possible energy $\mathcal{E}(u_l)$ while having unit norm and being orthogonal to $u_0, \ldots, u_{l-1}$.

In summary, given the observed data y_i, we can construct by means of several kernel algorithms basis vectors $\boldsymbol{u}_l$ with the appropriate regularity and asymptotic behavior for representing the Koopman generator V. In the ensuing subsections, we outline the structure of the resulting data-driven approximations of V and their spectral convergence.

2. Finite-difference approximation

Recall from (D6) and (D11) that the generator V behaves as a directional derivative operator on observables associated with the flow on $L^2(\mu)$ induced by the Koopman operator. Based on that, we can approximate the action of V on an observable $f: X \to \mathbb{C}$ in its domain by a finite-difference

approximation,

$$\begin{aligned} Vf(x_i) &= \lim_{t\to\infty} \frac{U^t f(x_i) - f(x_i)}{t} \\ &= \lim_{t\to\infty} \frac{f(\Phi^t(x_i)) - f(x_i)}{t} \\ &\approx \frac{f(x_{i+1}) - f(x_i)}{\Delta t}, \end{aligned} \tag{E5}$$

where x_i is a state underlying our training dataset with $i < \mathcal{N} - 1$. Note that the last expression in the right-hand side can be evaluated given only the values of f on the trajectory x_i, without requiring knowledge of the states x_i (which we consider to be unknown). If f is continuously differentiable, the approximation converges as $\Delta t \to 0$, uniformly with respect to $x_i \in X$. Moreover, other variants of this approximation such as backward, central, and/or higher-order finite-difference schemes can be employed.

We can represent the finite-difference approximation in (E5) by means of a linear operator $\mathsf{V}\colon \mathbb{C}^{\mathcal{N}} \to \mathbb{C}^{\mathcal{N}}$ defined as $\mathsf{V}\boldsymbol{f} = \boldsymbol{g}$, where

$$\boldsymbol{f} = (f_0, \ldots, f_{\mathcal{N}-1})^\top, \quad \boldsymbol{g} = (g_0, \ldots, g_{\mathcal{N}-1})^\top,$$
$$g_i = \begin{cases} (f_{i+1} - f_i)/\Delta t, & 0 \leqslant i \leqslant \mathcal{N} - 2, \\ 0, & i = \mathcal{N} - 1. \end{cases}$$

Note that if $\boldsymbol{f}$ is the vector storing the values of f on the trajectory, $f_i = f(x_i)$, then the elements g_i of $\boldsymbol{g}$ with $i \leqslant \mathcal{N} - 2$ are equal to the terms of the right-hand side of (E5). Moreover, the definition $g_{\mathcal{N}-1} = 0$ was arbitrary; one could set $g_{\mathcal{N}-1}$ to any constant without affecting the asymptotic behavior of the scheme as $\mathcal{N} \to \infty$.

Next, fixing a parameter $L \leqslant \mathcal{N} - 1$, we form the $L \times L$ antisymmetric matrix $\hat{\boldsymbol{V}} = [\hat{V}_{ij}]_{i,j=0}^{L-1}$ with elements

$$\hat{V}_{ij} = \frac{1}{2\mathcal{N}} \left(\boldsymbol{u}_i^\top \mathsf{V} \boldsymbol{u}_j - \boldsymbol{u}_j^\top \mathsf{V} \boldsymbol{u}_i \right). \tag{E6}$$

Using the pointwise ergodic result in (D4) in conjunction with the convergence of the eigenvectors $\boldsymbol{u}_j$ to the eigenfunctions $u_j \in C^1(X)$ (see Appendix E 1) it can be shown (e.g., [29,100]) that $\hat{V}_{ij}$ converges to matrix element $V_{ij} = \langle u_i, Vu_j \rangle_{L^2(\mu)}$ of the generator with respect to the $\{u_l\}_{l=0}^{\infty}$ basis of $L^2(\mu)$ in the iterated limit of $\mathcal{N} \to \infty$ followed by $\Delta t \to 0$. Thus, for any *fixed* $L \in \mathbb{N}$, the $L \times L$ data-driven matrix $\hat{\boldsymbol{V}}$ converges (in any matrix norm) to the matrix representation $\boldsymbol{V} = [V_{ij}]_{i,j=0}^{L-1}$ of the projected generator on the L-dimensional subspace of $L^2(\mu)$ spanned by $u_0, \ldots, u_{L-1}$. Note that the approximation error $\hat{V}_{ij} - V_{ij}$ of the matrix elements is not uniform with respect to $i, j \in \mathbb{N}$, so we fix L to an $\mathcal{N}$-independent value in order to control the error of the operator approximation of $\hat{\boldsymbol{V}}$ by $\boldsymbol{V}$ as $\mathcal{N} \to \infty$. Moreover, by employing an antisymmetric approximation from (E6), $\hat{V}_{ij} = -\hat{V}_{ji}$, our approximation scheme is structurally compatible with the antisymmetry of the generator, $V_{ij} = -V_{ji}$.

3. Regularization

Despite the fact that $\hat{\boldsymbol{V}}$ converges to $\boldsymbol{V}$ at fixed L, this matrix is not suitable by itself for spectral approximation of the generator V. First, V may be a nondiagonalizable operator with nontrivial continuous spectrum. Since every skew-adjoint finite-rank approximation of V, including $\hat{\boldsymbol{V}}$ and $\boldsymbol{V}$, is diagonalizable and has discrete spectrum, there are no *a priori* guarantees that the eigenvalues of the approximate operators converge in some sense to the spectrum of V, nor is it clear that the corresponding eigenvectors have a meaningful relation with properties of V. Even in the case of systems with pure point spectrum, if V has a dense set of eigenfrequencies (as is typically the case; see Appendix D 3) the eigenvalues of $\hat{\boldsymbol{V}}$ and $\boldsymbol{V}$ will not behave stably as L increases (see, e.g., Fig. 5 in Ref. [29]). Essentially, the only case (up to conjugacies) where stable numerical approximation of the spectrum of V is possible by the "raw" generator matrices $\hat{\boldsymbol{V}}$ and $\boldsymbol{V}$ is that of a circle rotation, $X = \mathbb{T}^1$, where the spectrum $\sigma(V; L^2(\mu))$ is a discrete set containing all integer multiples of the rotation frequency. When dealing with systems of higher complexity than this basic case, numerical approximation schemes require some type of regularization in order to ensure spectral convergence.

a. Diffusion regularization for systems with pure-point or mixed spectrum

Arguably, the simplest class of systems to analyze beyond circle rotations is that of finitely generated pure point spectrum systems with smooth Koopman eigenfunctions and smooth manifold structure of the support $M \subseteq X$ of the invariant measure. In such cases, it is possible to equip M with a smooth Riemannian metric whose corresponding Laplace-Beltrami operator Δ commutes with V [25], i.e.,

$$[V, \Delta] f := V\Delta f - \Delta V f = 0, \quad \forall f \in C^\infty(M). \tag{E7}$$

Moreover, the eigenvalues and eigenfunctions of Δ can be consistently approximated by a kernel integral operator $\mathcal{K}$ from (E3), where the kernel $k^{(Y)}$ is constructed using delay-coordinate maps. In more detail, in a limit of infinitely many delays and vanishing kernel bandwidth parameter ($\epsilon \to 0$) the eigenfunctions u_l of $\mathcal{K}$ converge to eigenfunctions of a Laplace-Beltrami operator Δ satisfying (E7), and the quantities $\eta_l = (\Lambda_l^{-1} - 1)/\epsilon$ converge to corresponding eigenvalues. Equivalently, we have that the operator $\mathcal{D} = (\mathcal{K}^{-1} - I)/\epsilon$ [which is self-adjoint on a dense subspace of $L^2(\mu)$] is an unbounded operator that spectrally approximates Δ. The restriction of $\mathcal{D}$ to the L-dimensional subspace of $L^2(\mu)$ spanned by $u_0, \ldots, u_{L-1}$ is represented by an $L \times L$ diagonal matrix, $\boldsymbol{D}$, with diagonal entries $D_{ii} = \langle u_i, \mathcal{D} u_i \rangle_{L^2(\mu)} = \eta_i$. In the data-driven setting, we approximate $\boldsymbol{D}$ by the $L \times L$ diagonal matrix $\hat{\boldsymbol{D}}$ with $\hat{D}_{ii} = \hat{\eta}_i := (\Lambda_l^{-1} - 1)/\epsilon$. Note that the eigenvalues of $\boldsymbol{D}$ are ordered in increasing order, $0 = \eta_0 < \eta_1 \leqslant \eta_2 \leqslant \cdots \eta_{L-1}$, and we employ a similar ordering for the eigenvalues of $\hat{\boldsymbol{D}}$. Moreover, $\hat{\boldsymbol{D}}$ converges to $\boldsymbol{D}$ in the large-data limit, $\mathcal{N} \to \infty$ for any fixed L.

Next, for a parameter $\tau > 0$, we introduce the regularized generator, $W\colon C^\infty(M) \to C^\infty(M)$, defined as

$$W_\tau = V - \tau\Delta. \tag{E8}$$

Let $E_j = \mathcal{E}(\phi_j)$ be the Dirichlet energy of Koopman eigenfunction ϕ_j from (E4). One readily verifies that for any $\tau > 0$, W_τ has the same eigenfunctions ϕ_j as V, and a discrete

spectrum of corresponding eigenvalues γ_j, where

$$\gamma_j = -\tau E_j + i\omega_j.$$

Observe, in particular, that the effect of adding a commuting diffusion operator to V is to shift the eigenvalues along the negative real line by an amount proportional to the Dirichlet energy of the corresponding eigenfunctions. This results in a spectrum of isolated eigenvalues for W_τ, even if the set of eigenfrequencies ω_j is dense in the imaginary line $i\mathbb{R}$. In particular, W_τ has a compact resolvent, and its eigenvalues and eigenfunctions can be approximated using Galerkin methods. Moreover, the operators W_τ converge to V in a spectral sense, since each γ_j clearly converges to ω_j as $\tau \to 0$.

In applications, we approximate the eigenvalues and eigenfunctions of W_τ through the following steps:

(1) Approximation of the Laplacian: Approximate W_τ by $\mathcal{W}_\tau := V - \tau\mathcal{D}$.

(2) Finite-rank projection (Galerkin approximation): Approximate $\mathcal{W}_\tau$ by $\boldsymbol{W} = \boldsymbol{V} - \tau\boldsymbol{D}$.

(3) Data-driven approximation: Approximate $\boldsymbol{W}$ by $\hat{\boldsymbol{W}} = \hat{\boldsymbol{V}} - \tau\hat{\boldsymbol{D}}$.

Having made these approximations, we solve the eigenvalue problem for the $L \times L$ matrix $\boldsymbol{W}$,

$$\hat{\boldsymbol{W}}\boldsymbol{c}_j = \hat{\gamma}_j\boldsymbol{c}_j,$$

normalizing the eigenvectors $\boldsymbol{c}_j = (c_{0j}, \ldots, c_{L-1,j})^\top$ to unit 2-norm, $\|\boldsymbol{c}_j\|_2 = 1$. The eigenvalues $\hat{\gamma}_j \in \mathbb{C}$ approximate γ_j from (E9). In particular, the imaginary part $\hat{\omega}_j := \mathrm{Im}\gamma_j$ approximates the generator eigenfrequency ω_j, whereas the real part approximates the corresponding Dirichlet energy via $\hat{E}_j := -\mathrm{Re}\hat{\gamma}_j/\tau$. Moreover, the elements of the eigenvectors $\boldsymbol{c}_j$ are expansion coefficients of vectors $\boldsymbol{\phi}_j \in \mathbb{C}^{\mathcal{N}}$,

$$\boldsymbol{\phi}_j = \sum_{i=0}^{L-1} c_{ij}\boldsymbol{u}_j,$$

which represent discretely sampled functions approximating the eigenfunctions ϕ_j in accordance with (E1). By convention, we order the eigenpairs $(\omega_j, \boldsymbol{\phi}_j)$ in order of increasing corresponding Dirichlet energy $\hat{E}_j$. Effectively, this means that we order numerical eigenfunctions in order of decreasing regularity (see Appendix E 1). This is a natural choice since the eigenfunctions ϕ_j with high regularity (low Dirichlet energy) are expected to be less sensitive to sampling and/or finite-difference errors.

The scheme described above converges in an iterated limit which parallels the sequence of approximations leading to $\hat{\boldsymbol{W}}$ from W; that is, $\mathcal{N} \to \infty$, $\Delta t \to 0$, $L \to \infty$, and $\tau \to 0$, taken in that order. We refer the reader to Refs. [25,100] for additional details. It should be noted that the C^∞ assumption on the eigenfunctions can be relaxed to mere continuity, $\phi_j \in C(M)$, as long as the eigenvalue spectrum $\sigma_e(V)$ has finite rank and the observation map Y is continuously differentiable [100]. Moreover, the same class of techniques can be employed to approximate the point spectrum of mixed-spectrum systems [100], though in this case the eigenfunctions ϕ_j provide an orthonormal basis of H_e only, which is a strict subspace of $L^2(\mu)$ (see Appendix D 3). It is also worthwhile noting that relaxing the continuity assumption on the ϕ_j is nontrivial. Indeed, there are known examples [112] of advection-diffusion flows, generated by operators of the form $W_\tau = V - \tau\Delta$, where the advection operator V has discontinuous eigenfunctions, and the small-viscosity ($\tau \to 0$) spectral behavior of W_τ as a function of τ is highly singular.

b. Spectral approximation by generator compactification

When the generator V on $L^2(\mu)$ has nontrivial continuous spectrum, the behavior of regularization schemes based on addition of diffusion is significantly more challenging to characterize. First, in many applications of interest the support of the invariant measure μ is not a differentiable manifold, even if the flow Φ^t on X is smooth. A classical example is the Lorenz 63 system [113], which is generated by a smooth vector field on $X = \mathbb{R}^3$, but due to dissipative dynamics the invariant measure μ is concentrated on a fractal attractor (the famous "butterfly" attractor) of zero Lebesgue measure. To our knowledge, in such systems with fractal attractors the construction of a diffusion operator Δ with compatible domain to that of the generator V on $L^2(\mu)$ is an open problem. As mentioned in Appendix E 3 a, even if μ is supported on a manifold, the $\tau \to 0$ limit of W_τ can be challenging to characterize when V has no eigenfunctions of sufficient regularity (which includes the case of mixing systems, where V has no nonconstant eigenfunctions).

As an effort to address these issues, Ref. [29] proposed an alternative approach, where instead of addition of diffusion, V is regularized by *composition* with appropriate smoothing operators. In brief, for each $\tau > 0$, we construct a family of kernel integral operators $G_\tau : L^2(\mu) \to L^2(\mu)$ with $\tau > 0$ such that (1) the associated kernel is C^1, (2) G_τ is an ergodic Markov operator, and (3) as $\tau \to 0$, G_τ converges strongly to the identity. We then have that for any $\tau > 0$, the operator $W_\tau : L^2(\mu) \to L^2(\mu)$ define as

$$\tilde{W}_\tau = G_\tau V G_\tau \tag{E9}$$

is skew-adjoint and compact. As a result, $\tilde{W}_\tau$ has a discrete, bounded spectrum of eigenfrequencies $i\omega_{j,\tau}$, and an associated orthonormal basis of eigenfunctions, $\phi_{j,\tau}$, i.e.,

$$\tilde{W}_\tau\phi_{j,\tau} = i\omega_{j,\tau}\phi_{j,\tau}.$$

Moreover, as $\tau \to 0$, there is a notion of spectral convergence of $\tilde{W}_\tau$ to V, even if V has nontrivial continuous spectrum. More specifically, for every element $i\omega$ of the spectrum of V, there is a sequence of eigenvalues $i\omega_{j_\tau,\tau}$ converging to $i\omega$ as $\tau \to 0$, and the spectral measures of $\tilde{W}_\tau$ (which are purely atomic by compactness) converge to the spectral measure of V in a suitable sense. While quantum simulation of systems with continuous spectrum is beyond the scope of this work, it is possible that the approach of Ref. [29] could be employed as an initial step to approximate the Koopman group generated by V by the unitary evolution group generated by $\tilde{W}_\tau$, and then employ the quantum computational framework described in the paper to simulate that system.

[1] R. P. Feynman, Simulating physics with computers, Int. J. Theor. Phys. **21**, 467 (1982).

[2] S. Lloyd, Universal quantum simulators, Science **273**, 1073 (1996).

[3] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders, Efficient quantum algorithms for simulating sparse Hamiltonians, Commun. Math. Phys. **270**, 359 (2007).

[4] S. Barnett, *Quantum Information*, Oxford Master Series in Atomic, Optical and Laser Physics (Oxford University Press, Oxford, 2009)

[5] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2010).

[6] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[7] I. H. Deutsch, Harnessing the Power of the Second Quantum Revolution, PRX Quantum **1**, 020101 (2020).

[8] L. K. Grover, From Schrödinger's equation to quantum search algorithm, Am. J. Phys. **69**, 769 (2001).

[9] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell *et al.*, Quantum supremacy using a programmable superconducting processor, Nature (London) **574**, 505 (2019).

[10] Y. Zhou, E. M. Stoudenmire, and X. Waintal, What Limits the Simulation of Quantum Computers? Phys. Rev. X **10**, 041038 (2020).

[11] D. A. Meyer, Quantum computing classical physics, Philos. Trans. R. Soc. London A **360**, 395 (2002).

[12] G. Benenti, G. Casati, S. Montangero, and D. L. Shepelyansky, Efficient Quantum Computing of Complex Dynamics, Phys. Rev. Lett. **87**, 227901 (2001).

[13] S. K. Leyton and T. J. Osborne, A quantum algorithm to solve nonlinear differential equations, arXiv:0812.4423 (2008).

[14] D. Giannakis, Quantum mechanics and data assimilation, Phys. Rev. E **100**, 032207 (2019).

[15] I. Joseph, Koopman–von Neumann approach to quantum simulation of nonlinear classical dynamics, Phys. Rev. Research **2**, 043102 (2020).

[16] S. Bharadwaj and K. R. Sreenivasan, Quantum computation of fluid dynamics, Indian Acad. Sci. Conf. Ser. **3**, 77 (2020).

[17] F. Gaitan, Finding flows of a Navier-Stokes fluid through quantum computing, npj Quantum Inf. **6**, 61 (2020).

[18] A. Mezzacapo, M. Sanz, L. Lamata, I. Egusquiza, S. Succi, and E. Solano, Quantum simulator for transport phenomena in fluid flows, Sci. Rep. **5**, 13153 (2015).

[19] M. Lubasch, J. Joo, P. Moinier, M. Kiffner, and D. Jaksch, Variational quantum algorithms for nonlinear problems, Phys. Rev. A **101**, 010301(R) (2020).

[20] T. J. Elliot and M. Gu, Superior memory efficiency of quantum devices for the simulation of continuous-time stochastic processes, npj Quantum Inf. **4**, 18 (2018).

[21] T. J. Elliott, C. Yang, F. C. Binder, A. J. P. G. Garner, J. Thompson, and M. Gu, Extreme Dimensionality Reduction with Quantum Modeling, Phys. Rev. Lett. **125**, 260501 (2020).

[22] T. Berry, D. Giannakis, and J. Harlim, Nonparametric forecasting of low-dimensional dynamical systems, Phys. Rev. E **91**, 032915 (2015).

[23] D. Giannakis, J. Slawinska, and Z. Zhao, Spatiotemporal feature extraction with data-driven Koopman operators, in *Proceedings of the 1st International Workshop on Feature Extraction: Modern Questions and Challenges at NIPS 2015*, edited by D. Storcheus, A. Rostamizadeh, and S. Kumar (PMLR, 2015), Vol. 44, pp. 103–115.

[24] Y. Kawahara, Dynamic mode decomposition with reproducing kernels for Koopman spectral analysis, in *Advances in Neural Information Processing Systems*, edited by D. D. Lee, M. Sugiyama, U. von Luxburg, I. Guyon, and R. Garnett (Curran Associates, 2016), pp. 911–919.

[25] D. Giannakis, Data-driven spectral decomposition and forecasting of ergodic dynamical systems, Appl. Comput. Harmon. Anal. **47**, 338 (2019).

[26] T. Berry, D. Giannakis, and J. Harlim, Bridging data science and dynamical systems theory, Notices Am. Math. Soc. **67**, 1336 (2020).

[27] S. Das and D. Giannakis, Koopman spectra in reproducing kernel Hilbert spaces, Appl. Comput. Harmon. Anal. **49**, 573 (2020).

[28] S. Klus, F. Nüske, S. Peitz, J.-H. Niemann, C. Clementi, and C. Schütte, Data-driven approximation of the Koopman generator: Model reduction, system identification, and control, Physica D **406**, 132416 (2020).

[29] S. Das, D. Giannakis, and J. Slawinska, Reproducing kernel Hilbert space quantification of unitary evolution groups, Appl. Comput. Harmon. Anal. **54**, 75 (2021).

[30] M. Schuld, I. Sinayskiy, and F. Petruccione, An introduction to quantum machine learning, Contemp. Phys. **56**, 172 (2015).

[31] J. Biamonte, P. Wittek, N. Pancotti, N. Wiebe, and S. Lloyd, Quantum machine learning, Nature (London) **549**, 195 (2017).

[32] C. Ciliberto, M. Herbster, A. D. Ialongo, M. Pontil, A. Rocchetto, S. Severini, and L. Wossnig, Quantum machine learning: A classical perspective, Proc. R. Soc. A **474**, 20170551 (2018).

[33] M. Schuld and N. Killoran, Quantum Machine Learning in Feature Hilbert Spaces, Phys. Rev. Lett. **122**, 040504 (2019).

[34] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, Supervised learning with quantum-enhanced feature spaces, Nature (London) **567**, 209 (2019).

[35] C. Blank, D. K. Park, J.-K. K. Rhee, and F. Petruccione, Quantum classifier with tailored quantum kernel, npj Quantum Inf. **6**, 41 (2020).

[36] M. Schuld, Supervised quantum machine learning models are kernel methods, arXiv:2101.11020 (2021).

[37] M. Dellnitz and O. Junge, On the approximation of complicated dynamical behavior, SIAM J. Numer. Anal. **36**, 491 (1999).

[38] M. Dellnitz and G. Froyland, On the isolated spectrum of the Perron-Frobenius operator, Nonlinearity **13**, 1171 (2000).

[39] I. Mezić, Spectral properties of dynamical systems, model reduction and decompositions, Nonlinear Dyn. **41**, 309 (2005).

[40] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, Spectral analysis of nonlinear flows, J. Fluid Mech. **641**, 115 (2009).

[41] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition, J. Nonlinear Sci. **25**, 1307 (2015).

[42] S. Klus, C. Koltai, and P. Schütte, On the numerical approximation of the Perron-Frobenius and Koopman operator, J. Comput. Dyn. **3**, 51 (2016).

[43] S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, and J. N. Kutz, Chaos as an intermittently forced linear system, Nat. Commun. **8**, 19 (2017).

[44] V. Baladi, *Positive Transfer Operators and Decay of Correlations*, Advanced Series in Nonlinear Dynamics, Vol. 16 (World Scientific, Singapore, 2000).

[45] T. Eisner, B. Farkas, M. Haase, and R. Nagel, *Operator Theoretic Aspects of Ergodic Theory*, Graduate Texts in Mathematics, Vol. 272 (Springer, Cham, 2015).

[46] J. C. Ferreira and V. A. Menegatto, Positive definiteness, reproducing kernel Hilbert spaces, and beyond, Ann. Funct. Anal. **4**, 64 (2013).

[47] V. I. Paulsen and M. Raghupathi, *An Introduction to the Theory of Reproducing Kernel Hilbert Spaces*, Cambridge Studies in Advanced Mathematics, Vol. 152 (Cambridge University Press, Cambridge, 2016).

[48] H. G. Feichtinger, S. S. Pandey, and T. Werther, Minimal norm interpolation in harmonic Hilbert spaces and Wiener amalgam spaces on locally compact Abelian groups, J. Math. Kuoto Univ. **47**, 65 (2007).

[49] Y. Kuznetsova and C. Molitor-Braun, Harmonic analysis of weighted L^p-algebras, Expo. Math. **30**, 124 (2012).

[50] S. Das and D. Giannakis, On harmonic Hilbert spaces on compact abelian groups, arXiv:1912.11664 (2019).

[51] D. Mauro, On Koopman–von Neumann waves, Int. J. Mod. Phys. A **17**, 1301 (2002).

[52] U. Klein, From Koopman–von Neumann theory to quantum theory, Quantum Stud.: Math. Found. **5**, 219 (2018).

[53] D. I. Bondar, F. Gay-Balmaz, and C. Tronci, Koopman wavefunctions and classical-quantum correlation dynamics, Proc. R. Soc. A **475**, 20180879 (2019).

[54] P. Morgan, An algebraic approach to Koopman classical mechanics, Ann. Phys. **414**, 168090 (2020).

[55] D. Giannakis, Quantum dynamics of the classical harmonic oscillator, J. Math. Phys. **62**, 042701 (2021).

[56] J. Welch, D. Greenbaum, S. Mostame, and A. Aspuru-Guzik, Efficient quantum circuits for diagonal unitaries without ancillas, New J. Phys. **16**, 033040 (2014).

[57] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, Parametrized quantum circuits as machine learning models, Quantum Sci. Technol. **4**, 043001 (2019).

[58] D. Marković and J. Grollier, Quantum neuromorphic computing, Appl. Phys. Lett. **117**, 150501 (2020).

[59] D. Coppersmith, An approximate Fourier transform useful in quantum factoring, Technical Report, IBM, arXiv:quant-ph/0201067 (1994).

[60] C. Moore and M. Nilsson, Parallel computation and quantum codes, SIAM J. Comput. **31**, 799 (2001).

[61] M. S. Anis, H. Abraham, R. Agarwal, AduOffei, G. Agliardi, M. Aharoni, I. Y. Akhalwaya, G. Aleksandrowicz, T. Alexander, M. Amy *et al.*, Qiskit: An open-source framework for quantum computing (2021).

[62] R. Josza and N. Linden, On the role of entanglement in quantum-computational speed-up, Proc. R. Soc. London A **459**, 2011 (2003).

[63] G. Vidal, Efficient Classical Simulation of Slightly Entangled Quantum Computations, Phys. Rev. Lett. **91**, 147902 (2003).

[64] D. E. Browne, Efficient classical simulation of the quantum Fourier transform, New J. Phys. **9**, 146 (2007).

[65] B. O. Koopman, Hamiltonian systems and transformation in Hilbert space, Proc. Natl. Acad. Sci. USA **17**, 315 (1931).

[66] B. O. Koopman and J. von Neumann, Dynamical systems of continuous spectra, Proc. Natl. Acad. Sci. USA **18**, 255 (1932).

[67] X. Wang, J. Slawinska, and D. Giannakis, Extended-range statistical ENSO prediction through operator-theoretic techniques for nonlinear dynamics, Sci. Rep. **10**, 2636 (2020).

[68] B. Schölkopf, A. Smola, and K. Müller, Nonlinear component analysis as a kernel eigenvalue problem, Neural Comput. **10**, 1299 (1998).

[69] A. Peres, *Quantum Theory: Concepts and Methods*, Fundamental Theories of Physics, Vol. 72 (Kluwer Academic Publishers, New York, 2002).

[70] C. Grebogi, E. Ott, and J. A. York, Attractors on an n-torus: Quasiperiodicity versus chaos, Physica D **15**, 354 (1985).

[71] R. E. Ecke, R. Mainieri, and T. S. Sullivan, Universality in quasiperiodic Rayleigh-Bénard convection, Phys. Rev. A **44**, 8103 (1991).

[72] D. Weixing, H. Wei, W. Xiaodong, and C. X. Yu, Quasiperiodic Transition to Chaos in a Plasma, Phys. Rev. Lett. **70**, 170 (1993).

[73] A. Kalogirou, E. E. Keaveny, and D. T. Papageorgiou, An in-depth numerical study of the two-dimensional Kuramoto-Sivashinsky equation, Proc. R. Soc. A **471**, 20140932 (2015).

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

[75] This implies (1) $U^s \circ U^t = U^{s+t}$, (2) $(U^t)^{-1} = U^{-t}$, and (3) $\lim_{t\to 0} U^t f = f$ for all $s, t \in \mathbb{R}$ and $f \in L^2(\mu)$.

[76] M. H. Stone, On one-parameter unitary groups in Hilbert space, Ann. Math. **33**, 643 (1932).

[77] F. Cucker and S. Smale, On the mathematical foundations of learning, Bull. Am. Math. Soc. **39**, 1 (2001).

[78] The qubit ordering in Qiskit is reverse to that in most textbooks on quantum computing.

[79] R. B. Griffiths and C.-S. Niu, Semiclassical Fourier Transform for Quantum Computation, Phys. Rev. Lett. **76**, 3228 (1996).

[80] N. Yoran and A. J. Short, Efficient classical simulation of the approximate quantum Fourier transform, Phys. Rev. A **76**, 042321 (2007).

[81] D. Aharonov, Z. Landau, and J. Makowsky, The quantum FFT can be classically simulated, arXiv:quant-ph/0611156 (2018).

[82] The notation $f(u) = \Theta(g(u))$ means that $f(u)$ is asymptotically bounded above and below by $g(u)$; i.e., there exists u_* and constants C_1, C_2 such that for all $u > u_*$, we have $C_1 g(u) < f(u) < C_2 g(u)$. Equivalently, we have $f(u) = O(g(u))$ and $g(u) = O(f(u))$.

[83] D. Harvey and J. van der Hoeven, Integer multiplication in time $O(n \log n)$, Ann. Math. **193**, 563 (2021).

[84] E. Klarreich, Multiplication hits the speed limit, Commun. ACM **63**, 11 (2020).

[85] J. Slawinska, A. Ourmazd, and D. Giannakis, A quantum mechanical approach for data assimilation in climate dynamics, in *Workshop on "Climate Change: How Can AI Help?"*, Proceedings of the 36th International Conference on Machine Learning, Long Beach, California (2019).

[86] D. Giannakis, A. Kolchinskaya, D. Krasnov, and J. Schumacher, Koopman analysis of the long-term evolution in a turbulent convection cell, J. Fluid Mech. **847**, 735 (2018).

[87] J. Wermer, On a class of normed rings, Ark. Mat. **2**, 537 (1954).
[88] L. Brandenburg, On identifying the maximal ideals in Banach algebras, J. Math. Anal. App. **50**, 489 (1975).
[89] H. G. Feichtinger, Gewichtsfunktionen auf lokalkompakten Gruppen, Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber. II **188**, 451 (1979).
[90] K. Gröchenig, Weight functions in time-frequency analysis, in *Pseudodifferential Operators: Partial Differential Equations and Time-Frequency Analysis*, Fields Inst. Commun., Vol. 52, edited by L. Rodino *et al.* (American Mathematical Society, Providence, RI, 2007), pp. 343–366.
[91] P. Walters, *An Introduction to Ergodic Theory*, Graduate Texts in Mathematics, Vol. 79 (Springer, New York, 1981).
[92] A. Katok and J.-P. Thouvenot, Spectral properties and combinatorial constructions in ergodic theory, in *Handbook of Dynamical Systems*, Vol. 1B, edited by B. Hasselblatt and A. Katok (North-Holland, Amsterdam, 2006), pp. 649–743.
[93] T.-T. Li and J. A. Yorke, Period three implies chaos, Am. Math. Mon. **82**, 985 (1975).
[94] A. F. M. ter Elst and M. Lemańczyk, On one-parameter Koopman groups, Ergod. Theor. Dyn. Sys. **37**, 1635 (2017).
[95] D. A. Lind, Spectral invariants and smooth ergodic theory, in *Dynamical Systems, Theory and Applications*, Lecture Notes in Physics, edited by J. Moser (Springer, Berlin, 1975), Vol. 38, pp. 296–308.
[96] D. V. Anosov and A. B. Katok, New examples in smooth ergodic theory. Ergodic diffeomorphisms, Trans. Moscow Math. Soc. **23**, 1 (1970).
[97] J. v. Neumann, Zur Operatorenmethode in der klassischen Mechanik, Ann. Math. **33**, 587 (1932).
[98] P. J. Schmid, Dynamic mode decomposition of numerical and experimental data, J. Fluid Mech. **656**, 5 (2010).
[99] S. Klus, I. Schuster, and K. Muandet, Eigendecomposition of transfer operators in reproducing kernel Hilbert spaces, J. Nonlinear Sci. **30**, 283 (2020).
[100] S. Das and D. Giannakis, Delay-coordinate maps and the spectra of Koopman operators, J. Stat. Phys. **175**, 1107 (2019).
[101] D. Giannakis, Delay-coordinate maps, coherence, and approximate spectra of evolution operators, Res. Math. Sci. **8**, 8 (2021).
[102] R. R. Coifman and S. Lafon, Diffusion maps, Appl. Comput. Harmon. Anal. **21**, 5 (2006).
[103] R. R. Coifman and S. Lafon, Geometric harmonics: A novel tool for multiscale out-of-sample extension of empirical functions, Appl. Comput. Harmon. Anal. **21**, 31 (2006).
[104] M. Hein, J.-Y. Audibert, and U. von Luxburg, From graphs to manifolds—Weak and strong pointwise consistency of graph Laplacians, in *Learning Theory*, Lecture Notes in Computer Science, Vol. 3559 (Springer, Berlin, 2005), pp. 470–485.
[105] M. Belkin and P. Niyogi, Convergence of Laplacian eigenmaps, in *Advances in Neural Information Processing Systems* (MIT Press, Cambridge, MA, 2007), pp. 129–136.
[106] U. von Luxburg, M. Belkin, and O. Bousquet, Consitency of spectral clustering, Ann. Stat. **26**, 555 (2008).
[107] Z. Shi, Convergence of Laplacian spectra from random samples, arXiv:1507.00151 (2015).
[108] N. G. Trillos and D. Slepčev, A variational approach to the consistency of spectral clustering, Appl. Comput. Harmon. Anal. **45**, 239 (2018).
[109] N. G. Trillos, M. Gerlach, M. Hein, and D. Slepčev, Error estimates for spectral convergence of the graph Laplacian on random geometric graphs towards the Laplace–Beltrami operator, Found. Comput. Math. **20**, 827 (2020).
[110] C. L. Wormell and S. Reich, Spectral convergence of diffusion maps: Improved error bounds and an alternative normalisation, SIAM J. Numer. Anal. **59**, 1687 (2021).
[111] T. Sauer, J. A. Yorke, and M. Casdagli, Embedology, J. Stat. Phys. **65**, 579 (1991).
[112] P. Constantin, A. Kiselev, L. Ryzhik, and A. Zlatoš, Diffusion and mixing in fluid flow, Ann. Math. **168**, 643 (2008).
[113] E. N. Lorenz, Deterministic nonperiodic flow, J. Atmos. Sci. **20**, 130 (1963).