

Efficient Measurement-Driven Eigenenergy Estimation with Classical Shadows

Yizhi Shen^{1,*†}, Alex Buzali^{2,†}, Hong-Ye Hu², Katherine Klymko³, Daan Camps³,
Susanne F. Yelin² and Roel Van Beeumen¹

¹*Applied Mathematics and Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

²*Department of Physics, Harvard University, 17 Oxford Street, Cambridge, Massachusetts 02138, USA*

³*NERSC, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*



(Received 18 November 2024; revised 25 July 2025; accepted 23 December 2025; published 10 February 2026)

Quantum algorithms exploiting real-time evolution under a target Hamiltonian have demonstrated remarkable efficiency in extracting key spectral information. However, the broader potential of these methods, particularly beyond ground-state calculations, is underexplored. In this work, we introduce the framework of multiobservable dynamic mode decomposition (MODMD), which combines the observable dynamic mode decomposition (DMD), a measurement-driven eigensolver tailored for near-term implementation, with classical shadow tomography. MODMD leverages random scrambling in the classical shadow technique to construct, with exponentially reduced resource requirements, a signal subspace that encodes rich spectral information. Notably, we replace typical Hadamard-test circuits with a protocol designed to predict low-rank observables, thereby broadening the use of classical shadow tomography for predicting many low-rank observables. We establish theoretical guarantees on the spectral approximation from MODMD, taking into account distinct sources of error. In the ideal case, we prove that the spectral error scales as $\exp(-\Delta E t_{\max})$, where ΔE is the Hamiltonian spectral gap and $t_{\max}$ is the maximal simulation time. This analysis provides a rigorous justification of the rapid convergence observed across simulations. To demonstrate the utility of our framework, we consider its application to fundamental tasks, such as determining the low-lying, i.e., ground or excited, energies of representative many-body systems. Our work paves the path for efficient designs of measurement-driven algorithms on near-term and early fault-tolerant quantum devices.

DOI: [10.1103/74s6-3jsz](https://doi.org/10.1103/74s6-3jsz)

I. INTRODUCTION

Quantum algorithms based on real-time evolution [1–14] have gained increasing popularity due to the unitarity of real-time dynamics and their native implementation on quantum hardware. In particular, real-time eigensolvers have demonstrated impressive efficacy in extracting the ground-state information of physical many-body systems. This success raises a natural question: can the algorithms perform highly accurate excited-state calculations?

The real-time approaches were initially conceptualized from a subspace diagonalization perspective, building on classical techniques that recover eigenvalues through time

correlation functions—an idea explored for decades [15]. Yet their quantum computational counterparts have only seen active development in the past few years: several related methods have since emerged under different names in the literature, including quantum filter diagonalization (QFD) [1,2], variational quantum phase estimation (VQPE) [3,4], and quantum Krylov (QK) [5] approaches. While these subspace methods, in theory, have the potential to extract spectral information beyond the ground state, they typically involve solving ill-conditioned generalized eigenvalue problems [16] and can, hence, suffer from any perturbative error, such as prevalent hardware noise. Alternatively, distinct recent approaches [7,8,10,11,17] draw inspirations from a signal processing perspective, where real-time evolution is primarily employed to generate time series that resonate with the target eigenmodes. The signal processing approaches tend to be resilient to perturbative noise due to their robust regularization. On the other hand, they often rely on state preparation such that the initial state undergoing the Hamiltonian dynamics must overlap dominantly with the eigenstates of interest. Without this key

*Contact author: yizhis@lbl.gov

†These authors contributed equally to this work.

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

assumption, accurately resolving multiple eigenenergies may demand substantial quantum or classical resources.

On the other hand, quantum phase estimation (QPE) and its variants [18–24] stand as the canonical approach to access eigenphases and prepare eigenstates coherently on a quantum computer. The original QPE protocol uses a register of ancilla qubits, a sequence of controlled time evolution acting on system qubits, and the inverse quantum Fourier transform. In particular, QPE provides both energy and eigenstate information in each run, achieving Heisenberg-limited scaling with a total runtime inversely proportional to the desired accuracy. The inverse quantum Fourier transform can be performed semiclassically through single-qubit measurements and controlled rotations [20] whereas the ancilla overhead can be reduced to a single qubit [21,22], making QPE an especially attractive candidate for fault-tolerant devices.

Despite its optimal asymptotic scaling, QPE imposes a requirement on the circuit depth or maximal runtime: unless the initial state is an exact eigenstate, the depth must increase inversely with the energy accuracy, a lower bound intrinsic to the protocol. All known QPE variants share this maximal runtime scaling, with the preconstant unchanged even as the initial state approaches an exact eigenstate [7]. This requirement can constitute a serious bottleneck for implementing QPE on near-term and early fault-tolerant hardware, where circuit depth continues to be a primary constraint.

In this work, we introduce a quantum-classical hybrid framework that combines time evolution and scrambling on quantum hardware with classical data postprocessing to access the ground and low-lying excited-state energies of many-body systems. Our framework prepares a single, simple initial state on the quantum computer with overall resource demands comparable to the leading ground-state subspace algorithms. Specifically, it features a shallower circuit depth [25] than QPE when the initial state is close to an eigenstate. Though exceeding the Heisenberg limit, the empirical total runtime remains modest for retrieving the lowest eigenenergies at physically relevant accuracies. Furthermore, the classical postprocessing in our approach is *ansatz-free* and therefore circumvents the challenges of complex optimization landscapes commonly encountered in parametric methods, particularly as the number of target eigenenergies increases [26–28].

The main technical tools that we leverage are classical shadow tomography [29–35] and observable dynamic mode decomposition (ODMD) [9]. We develop a simple, single-ancilla shadow protocol to collect real-time signals associated with many observables and utilize ODMD to unravel these signals into single-energy modes. Since oscillating time signals with distinct power spectra are linearly independent of the space of all sinusoidal functions, they form the basis of some *signal subspace*, which can be systematically expanded to accommodate the desired

frequencies. Our approach, termed the MODMD, can hence be viewed as a unifying framework that enjoys the strength of both subspace and signal processing algorithms.

Compared to alternative hybrid eigensolvers, our real-time framework has the following notable advantages: it (1) achieves near-exponential convergence of the eigenenergy estimates surpassing the conventional Fourier limit, (2) provides extensive knowledge of a many-body system including eigenstate properties and dynamical responses, and (3) significantly saves quantum resources in terms of the evolution time while showing stability against perturbative noise. These features make our algorithm promising for implementations on emerging quantum platforms, such as near-term analog quantum simulators and early fault-tolerant quantum computers.

The manuscript is organized as follows. In Sec. II, we first overview key concepts underlying recent advances in real-time eigensolvers. We then establish in Sec. III the building blocks of the real-time MODMD framework and detail our core algorithm for eigenenergy and eigenstate estimation. Theoretical guarantees on its convergence and preliminary error analysis are presented within Sec. IV. Finally, we numerically demonstrate our algorithm in Sec. V by focusing on many-body examples from condensed matter physics and quantum chemistry.

II. PRELIMINARIES

In this work, we develop an approach that gives highly accurate eigenenergy estimates beyond the ground state. Specifically, we propose a real-time framework combining the ODMD and classical shadow tomography, where we fully leverage the synergy between these two algorithmic components. We will show that we can design a simple shadow protocol to predict the real-time expectation value $\langle \phi_0 | O e^{-iHt} | \phi_0 \rangle$ for the problem Hamiltonian H and a reference state $|\phi_0\rangle$. In particular, classical shadows enable the simultaneous prediction of expectations for many observables O of our choice. These expectations generate a multivariate time series, whose characteristic frequencies can be efficiently extracted by ODMD (see Fig. 1 for a summary).

From here on, we will use $\mathcal{H} \subset \mathbb{C}^N$ to denote the physical Hilbert space of many-body quantum states and $\mathbb{L}(\mathcal{H})$ to denote the Liouville space, i.e., the space of linear operators acting on $\mathcal{H}$ (we work with finite-dimensional and thus bounded linear operators for simplicity). Additional essential notations are defined self-consistently in Table I and throughout the main text.

A. Real-time quantum eigensolvers

In this subsection, we first highlight the capabilities of real-time evolution in determining the eigenenergies of a target Hamiltonian $H \in \mathbb{C}^{N \times N}$. We consider the spectral

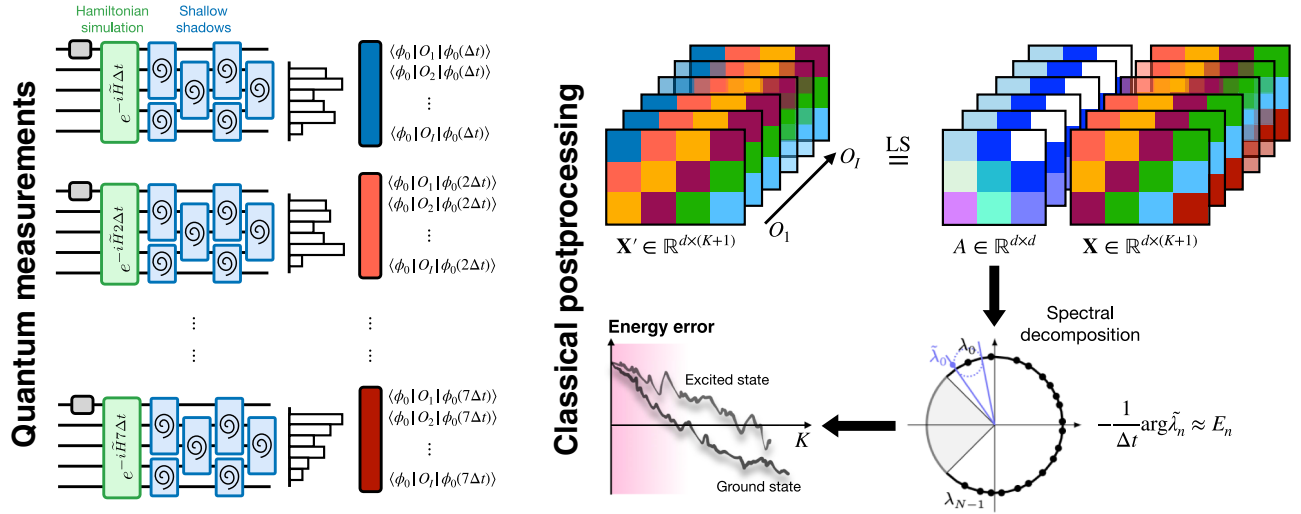


FIG. 1. Multiobservable dynamic mode decomposition for eigenenergy estimation. Multiobservable dynamic mode decomposition collects the expectations, $\text{Tr}[\rho(k\Delta t)\Gamma_O] = \text{Re}\langle\phi_0|Oe^{-iHk\Delta t}|\phi_0\rangle$, with respect to a simple reference state $|\phi_0\rangle$. This data can be measured efficiently on a quantum processor through Hamiltonian simulations combined with shadow tomographic techniques, where the shadow circuits can be shallow with a depth logarithmic in the number of qubits. Multiobservable dynamic mode decomposition then constructs a pair of block Hankel matrices $(\mathbf{X}, \mathbf{X}')$ and computes the dynamic mode decomposition system matrix A that adopts a block companion structure. The eigenvalues $\tilde{\lambda}_n$ of A converge to the true eigenphases λ_n as the size of $\mathbf{X}$ and $\mathbf{X}'$ increases. The low-lying energies E_n are estimated as angles of $\tilde{\lambda}_n$, $\tilde{E}_n = -(1/\Delta t)\arg(\tilde{\lambda}_n)$.

decomposition, $H = \sum_{n=0}^{N-1} E_n |\psi_n\rangle\langle\psi_n|$, of the Hamiltonian with ordered energies $E_0 \leq E_1 \leq \dots \leq E_{N-1}$. The real-time approaches commonly require evaluation of the expectation value [1–3,5,7–13,17],

$$\langle\phi_0|e^{-iHt}|\phi_0\rangle = \sum_{n=0}^{N-1} |\alpha_n|^2 e^{-iE_n t}, \quad (1)$$

where $|\phi_0\rangle \in \mathcal{H}$ is a reference state, $\alpha_n = \langle\psi_n|\phi_0\rangle$ is the overlap between the reference state and eigenstate $|\psi_n\rangle$,

TABLE I. Glossary of key notations.

Simulation protocol	
Hilbert space	$\mathcal{H}, \dim\mathcal{H} = N = 2^L$
Liouville space	$\mathbb{L}(\mathcal{H}), \dim\mathbb{L}(\mathcal{H}) = N^2$
Hamiltonian in Pauli basis	$H = \sum_{v=1}^M \kappa_v P_v$
Hermitian observables	$O_i = \sum_{v=1}^{M_i} \kappa_{i,v} P_{i,v}$
Hamiltonian eigenstates	$\{ \psi_n\rangle\}_{n=0}^{N-1}$
Pure quantum state	$\rho_\phi = \phi\rangle\langle\phi $
Multiobservable signal	$\{\langle\phi_0 O_i e^{-iHt} \phi_0\rangle\}_{i=1}^I$
Classical shadow dataset	$\{\hat{\sigma}_q = U_q^\dagger b_q\rangle\langle b_q U_q\}_{q=1}^Q$
Algorithmic controls	
Time step	Δt
Number of distinct observables	I
Shape of data slices per observable	$(d, K), \frac{K}{d} = \text{constant}$
Systematic error (e.g., shot noise)	$\epsilon_1(\Delta t, I, \max_{1 \leq i \leq I} \ \vec{\kappa}_i\ _1, Q)$
Algorithmic error from eigensolver	$\epsilon(\langle\phi_0\rangle, \Delta t, I, d, \epsilon_1)$

and $t = k\Delta t$ is typically an integer multiple of some time step Δt . The expectation can be efficiently sampled from repeated quantum measurements through the Hadamard test [36] or mirror fidelity estimation [5,37].

There are several approaches for obtaining energy estimates from data of the type Eq. (1). Quantum subspace diagonalization [1–3,5,12,13] approximates the extremal eigenenergies by forming a projected eigenvalue problem. In particular, the quantum subspace can be built through successive applications of unitary time evolutions, where increasing powers $k = t/\Delta t$ of the time-evolution operator $e^{-iH\Delta t}$ generate additional matrix elements for the projected problem. Earlier work on QFD [1,2] introduced multireference variants of this approach, improving coverage of spectrum at the expense of a quadratic overhead in the number of initial states. More recent approaches, such as VQPE [3,4] and QK algorithms [5,6], have focused on single-reference protocols that remain effective even with relatively coarse time step—going beyond the perturbative regime where Taylor expansions of e^{-iHt} suffice. While VQPE and QK differ in circuit implementations, they share the same underlying philosophy of treating real-time evolution as the generator of subspace. With a single reference state $|\phi_0\rangle$, the methods often struggle to find excited states, though convergence can be potentially facilitated by preparation of multiple references with the aforementioned quadratic cost. Moreover, the real-time states $\{e^{-iHk\Delta t}|\phi_0\rangle\}_{k \geq 0}$ lose orthogonality as a subspace basis, which can thereby lead to ill-conditioning and susceptibility to noise.

Alternatively, signal processing phase estimation algorithms [7,8,10,11,17] capitalize on Fourier or harmonic analysis to resolve the eigenfrequencies or eigenenergies of interest. These approaches mitigate the impacts of noise by minimizing customized objective functions, which can sharpen into robust optima as the number of time steps increases. Despite a resilience to noise, the optimization landscape depends critically on the choice of the reference state $|\phi_0\rangle$. For example, consider the time correlations of Eq. (1) to be a spectral density over the unit circle, where the squared overlaps $|\alpha_n|^2$ define the normalized spectral weights. Techniques such as peak finding on the spectral density may not yield unique solution when the reference state has roughly uniform eigenstate overlaps. To ensure accurate energy estimation, more sophisticated reference state preparations and postprocessing designs need to be accounted for. Among such approaches, quantum complex exponential least squares (QCELS) [7,8] algorithms fit measured time correlations to a linear combination of complex exponentials upon minimizing the residual error, thereby extracting the eigenphases with high resolution. Building on this, quantum multiple eigenvalue Gaussian filtered search (QMEGS) [11] checks a grid of candidate spectral centers in the frequency domain—interpreted as Fourier transform of the time-domain signal—to localize peaks resonant with the eigenenergies. Both methods are contingent on a reference state that overlaps dominantly with the target eigenstates. While QCELS achieves high accuracy at a cost of classical postprocessing that scales exponentially in the number of estimated eigenenergies, QMEGS offers a scalable route but suffers from spectral broadening due to decoherence and other noise sources.

To access both ground- and excited-state properties, we aim for a simple, accurate, and robust real-time protocol. To this end, we will introduce a *quantum signal subspace* approach that utilizes signals of the form $\langle\phi_0|Oe^{-iHt}|\phi_0\rangle$ for general operators O . Importantly, we seek to measure many such operator expectations simultaneously with a cost logarithmic in the number of observables. This can be precisely achieved via classical shadows, which we explore in the following subsection.

B. Efficient measurement with classical shadows

Classical shadow tomography [29,30,38–40] embodies a powerful suite for efficiently measuring expectations of many observables simultaneously,

$$\text{Tr}[\rho_\phi \Gamma_i] = \langle\phi|\Gamma_i|\phi\rangle, \quad \{\Gamma_i\}_{i=1}^I, \quad (2)$$

where $\Gamma_i \in \mathbb{L}(\mathcal{H})$ has an efficient representation on a classical computer. Here $\rho_\phi = |\phi\rangle\langle\phi| \in \mathbb{L}(\mathcal{H})$ is a pure state, though a similar trace evaluation applies to mixed states. Classical shadow tomography consists of two key steps: (1) random quench evolution using $U \in \mathcal{U}$ from a unitary ensemble $\mathcal{U}$ and (2) computational basis measurement.

Upon each measurement, the quantum state collapses to a bitstring $|b_q\rangle$. After repetitive simulations, one obtains a classical *shadow dataset*, $\{U_q^\dagger|b_q\rangle\langle b_q|U_q\}_{q=1}^Q$, which can be viewed as a classical sketch of the quantum state. It is known [29] that $Q = \mathcal{O}(\log(I) \max_i \|\Gamma_i\|_{\text{sh}} \epsilon_1^{-2})$ shadows can predict all the I expectation values given by Eq. (2) to uncertainty ϵ_1 with high probability. Here, the shadow norm, $\|\Gamma\|_{\text{sh}}$, depends on both the unitary ensemble $\mathcal{U}$ and the operator Γ . For example, when the ensemble $\mathcal{U}$ is the L -qubit random Clifford unitaries, i.e., $\mathcal{U} = \text{Cl}(2^L)$, and the operator $\Gamma = \Gamma^\dagger$ is Hermitian, the shadow norm is $\|\Gamma\|_{\text{sh}} = 3\text{Tr}[\Gamma^2]$. This is especially powerful if Γ is *low-rank*, meaning that its operator rank stays independent of the system size. Because $\text{Tr}[\Gamma^2] \leq \text{rank}(\Gamma)\|\Gamma\|_2^2$, one can use classical shadows to predict exponentially many low-rank expectations simultaneously even for large systems. In Sec. III, we demonstrate how to transfer the measurement of many expectations $\{\text{Re}\langle\phi_0|Oe^{-iHt}|\phi_0\rangle\}_i$ to the task of predicting low-rank Hermitian operators.

Overall, the framework presented within this work is distinctive in various crucial aspects. First, our algorithm can directly estimate the individual eigenenergies and the associated energy gaps. Conversely, resolving the energy levels from estimated excitation gaps is nearly infeasible. Second, instead of the standard usage of classical shadows in predicting many local Pauli observables, we introduce a targeted application of classical shadows by replacing the Hadamard-test-like circuits with those that predict many *low-rank* observables, expanding their primary utility beyond entanglement witnesses in quantum information science. Last, our postprocessing scheme yields robust and accurate eigenenergy estimates, achieving an exponential error reduction in the low-noise regime. This surpasses an algebraic error decay in the conventional Fourier limit. Although recent work [41] has begun exploring the use of shadow techniques in spectroscopic calculations of energy gaps, our approach distinguishes itself in the sense we have discussed above.

Table II presents a comprehensive comparison of state-of-the-art real-time methods for extracting valuable spectral information, highlighting the efficiency and accuracy of our eigensolver. In Sec. III we will outline the basic construction of our measurement-driven approach, which deploys classical shadows to evaluate the real-time expectations.

III. MODMD FRAMEWORK

Here we present a general perspective to the problem of eigenenergy estimation, pushing the limits of convergence and robustness via a quantum signal subspace approach. Our quantum signal space is composed of time correlation functions of the form

$$\langle\phi_0|Oe^{-iHt}|\phi_0\rangle, \quad O = O^\dagger \in \mathbb{C}^{N \times N}, \quad (3)$$

TABLE II. Landscape of real-time hybrid eigensolver leveraging quantum measurements of $\langle \phi_0 | O e^{-iHt} | \phi_0 \rangle$ and classical postprocessing schemes. We consider estimating a target set of Hamiltonian eigenergies, $\{E_n : n \in \mathcal{N}_{\text{eig}} \subset [N]\}$, where $|\mathcal{N}_{\text{eig}}| = N_{\text{eig}}$ for $N_{\text{eig}} \ll N$. On quantum hardware, real-time evolution of a reference state $|\phi_0\rangle$ is performed and different expectation values are measured with an uncertainty of ϵ_1 . The classical postprocessing then extracts from these measurements our eigenenergy estimates with an algorithmic error of ϵ . Here $\tilde{O}(\cdot)$ denotes asymptotic upper bound with polylogarithmic factors neglected.

Real-time methods	Measurement cost	Algorithmic convergence	Target overlaps	System properties
Single-observable signal				
Subspace [1–3,5,12,13]	$\mathcal{O}(\epsilon_1^{-2})$	$\tilde{O}(\epsilon_1^{-2})$ in low- ϵ_1 limit unstable in high- ϵ_1 limit	nonvanishing	✓ eigenstates
<i>Example algorithms</i>				
• QFD [1,2]: multireference; uniform timegrid (with small Δt) • VQPE [3,4]: single-reference (primary focus); flexible timegrid (with large Δt); noise-thresholded convergence • QK [5]: single-reference (primary focus); uniform timegrid; signal sampling with multifidelity estimation protocol				
Signal processing [7,8,10,11]	$\mathcal{O}(\epsilon_1^{-2})$	$\tilde{O}(\epsilon^{-1})$	dominant	✗ eigenstates
<i>Example algorithms</i>				
• (MM)QCELS [7,8]: single-reference; adaptive or randomly sampled timegrid • QMEGS [11]: single-reference; randomly sampled timegrid; improved classical postprocessing				
Multiobservable signal				
Shadow spectroscopy [41]	$\mathcal{O}(\log(I)3^{\text{loc}[O]}\epsilon_1^{-2})$	uncertain possibly ambiguous energy gaps	nonvanishing	✓ dynamics
Signal subspace (this work)	$\mathcal{O}(\log(I)\epsilon_1^{-2})$	$\tilde{O}(\epsilon^{-4/3})$ in some cases stable with respect to ϵ_1 [25,42,43]	nonvanishing	✓ eigenstates ✓ dynamics

which captures the system quantum dynamics. Since the expectation value oscillates over time, it can be uniquely expressed in the natural basis $e^{-iE_n t}$. Representing real-time data in this eigenfrequency basis establishes a clear notion of linear independence in the space of signals. By evaluating the expectation values in Eq. (3) for multiple independent operators $\{O_1, O_2, \dots, O_I\}$, we thus manage to construct a signal subspace from which we can extract spectral information significantly better. We will refer to the vector of expectation values,

$$\vec{s}(t) = \begin{bmatrix} \langle \phi_0 | O_1 e^{-iHt} | \phi_0 \rangle \\ \langle \phi_0 | O_2 e^{-iHt} | \phi_0 \rangle \\ \vdots \\ \langle \phi_0 | O_I e^{-iHt} | \phi_0 \rangle \end{bmatrix}, \quad (4)$$

as a *multiobservable* signal associated with $\{O_i\}_{i=1}^I$. The central ingredient of our real-time framework is hence the efficient collection of a multiobservable signal $\vec{s}(t)$, whose dimensionality I reflects independence or richness of the underlying spectral information. In particular, the state overlap $\langle \phi_0 | e^{-iHt} | \phi_0 \rangle$ in Eq. (1) can be viewed as a simple “one-dimensional” signal which oscillates over time.

To efficiently measure the multiobservable signal $\vec{s}(t)$, we leverage classical shadow tomography, specifically the shallow shadows recently demonstrated on hardware [40]. We will show that one can evaluate all the I expectation values in Eq. (4) up to a small error simultaneously, with each expressed as the expectation value of a

low-rank observable. Before elaborating on the favorable exponential cost reduction, we first introduce our basic measurement-driven framework.

A. Basic construction

1. Primer: ODMD

In recent work [9] we examined the ODMD as a powerful extension of the classical DMD formalism. ODMD exploits quantum resources to efficiently measure the expectations of time evolution operators, $s(t) = \text{Re} \langle \phi_0 | e^{-iHt} | \phi_0 \rangle$, rather than directly tracking the dynamics of the full quantum state $e^{-iHt} |\phi_0\rangle \in \mathcal{H}$. The extremal (both maximum and minimum) eigenphases $e^{-iE_n t}$ and, thereby, the eigenenergies E_n can be inferred via the least-squares (LS) solution to the following system of linear homogeneous equations:

$$\underbrace{\begin{bmatrix} s_1 & s_2 & \cdots & s_{K+1} \\ s_2 & s_3 & \cdots & s_{K+2} \\ \vdots & \vdots & \ddots & \vdots \\ s_d & s_{d+1} & \cdots & s_{K+d} \end{bmatrix}}_{\mathbf{X}' \in \mathbb{R}^{d \times (K+1)}} \stackrel{\text{LS}}{=} \underbrace{\begin{bmatrix} s_0 & s_1 & \cdots & s_K \\ s_1 & s_2 & \cdots & s_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_{d-1} & s_d & \cdots & s_{K+d-1} \end{bmatrix}}_{\mathbf{X} \in \mathbb{R}^{d \times (K+1)}}, \quad (5)$$

where $s_k = s(k\Delta t)$ is sampled at a regular time step Δt , and $(\mathbf{X}, \mathbf{X}')$ are time-shifted data matrices containing the overlap evaluations. We note that the Hankel structure of $\mathbf{X}$ and $\mathbf{X}'$ immediately implies a compact companion structure of

the *system matrix*, $A \in \mathbb{R}^{d \times d}$. The extremal-phase eigenvalues $\tilde{\lambda} \in \mathbb{C}$ of the system matrix A converge rapidly to the extremal eigenphases $\lambda_n = e^{-iE_n \Delta t}$ of the target Hamiltonian as we increase the dimensions d and K of the data matrices.

The ODMD algorithm excels in extracting the ground-state information. As we move toward the interior of the spectrum, its convergence often slows down progressively or even stagnates. To achieve compact and accurate excited-state estimation—a task more intricate than ground-state estimation—we take advantage of classical shadows to generate extensive real-time data with minimal quantum resource overhead.

2. MODMD with classical shadows

Now we generalize the ODMD approach from the time evolution operator e^{-iHt} to an operator pool of arbitrary size. For the signal collection process, we invoke the idea of classical shadow tomography originally introduced [29] to extract arbitrary many-body properties from random projective measurements. Similar to the Hadamard test, we introduce a single ancilla qubit to control elementary operations on the system registers, demonstrating for the first time that this setup can in fact efficiently predict the multiobservable signal $\vec{s}(t)$ in Eq. (4).

With the ancilla initialized in $|0_a\rangle \in \mathcal{H}_a$, we first create a superposition state $|\Phi(t)\rangle \in \mathcal{H}_a \otimes \mathcal{H}$ given by

$$|\Phi(t)\rangle = \frac{|0, \phi_\perp\rangle + |1, \phi_0(t)\rangle}{\sqrt{2}}, \quad (6)$$

where $|1, \phi_0(t)\rangle = |1_a\rangle \otimes |\phi_0(t)\rangle$ denotes the product of the ancillary state and the time-evolved reference state $|\phi_0(t)\rangle = e^{-iHt} |\phi_0\rangle$ and $|\phi_\perp\rangle \in \mathcal{H}$ is any residual state.

For the composite state $\rho(t) := |\Phi(t)\rangle \langle \Phi(t)| \in \mathbb{L}(\mathcal{H}_a \otimes \mathcal{H})$ and the operator $\Gamma = |1, \phi_0\rangle \langle 0, \phi_\perp| + |0, \phi_\perp\rangle \langle 1, \phi_0| \in \mathbb{L}(\mathcal{H}_a \otimes \mathcal{H})$, we recognize the pivotal relation that

$$\text{Re} \langle \phi_0 | \phi_0(t) \rangle = \text{Tr}[\rho(t)\Gamma], \quad (7)$$

which establishes a fundamental connection between the expectation in Eq. (1) and the trace in Eq. (2). Notably, $\Gamma = \Gamma^\dagger$ is Hermitian, and its classical simulability entirely depends on that of $|\phi_0\rangle$ and $|\phi_\perp\rangle$. Following Eq. (7), it is straightforward to show that the density operator $\rho(t)$ in fact encodes the time correlation function $\langle \phi_0 | O | \phi_0(t) \rangle$ of any Hermitian operator $O = O^\dagger \in \mathbb{L}(\mathcal{H})$, since

$$\text{Re} \langle \phi_0 | O | \phi_0(t) \rangle = \text{Tr}[\rho(t)\Gamma_O] = \text{Re} \sum_{n=0}^{N-1} c_n e^{-iE_n t}, \quad (8)$$

where $c_n = \sum_{m=0}^{N-1} \alpha_m^* \alpha_n \langle m | O | n \rangle$ and

$$\Gamma_O = (\text{Id}_a \otimes O) |1, \phi_0\rangle \langle 0, \phi_\perp| + \text{h.c.}, \quad (9)$$

with the identity Id_a acting on the ancilla Hilbert space. The classical simulability of Γ_O now depends on that of $|\phi_0\rangle$, $|\phi_\perp\rangle$, and O . In particular, Γ_O has an efficient classical representation if the two states and the operator are sparse in the computational and Pauli basis, respectively. For instance, if $|\phi_0\rangle$ and $|\phi_\perp\rangle$ are simple computational basis states, as in the context of quantum chemistry, Γ_O can be classically represented for any Pauli string O .

If one measures each term from $\{\text{Re} \langle \phi_0 | O_i | \phi_0(t) \rangle\}_{i=1}^I$ individually with a Hadamard-test circuit, it will take a total of $\mathcal{O}(I\epsilon_1^{-2})$ measurements to achieve a uniform error of ϵ_1 . With classical shadow tomography, we significantly reduce this measurement sample complexity. As shown in Eq. (8), we can rewrite each expectation value as the trace of a Hermitian, *low-rank* observable Γ_O over a state $\rho(t) \in \mathbb{L}(\mathcal{H}_a \otimes \mathcal{H})$. Moreover, we emphasize that the state preparation circuit of $\rho(t)$ only involves a single ancillary qubit, matching the overhead of a Hadamard-test circuit. As shown in Fig. 1, after preparing $\rho(t)$ with the quantum time evolution circuit, we apply a global Clifford shadow protocol on the quantum state and collect shadows $\{\hat{\sigma}_q = U_q^\dagger |b_q\rangle \langle b_q| U_q\}_{q=1}^Q$. On a classical computer, we employ the stored shadows to construct an empirical estimator for the trace,

$$\langle \Gamma_O \rangle_{\text{est}} = \frac{1}{Q} \sum_{q=1}^Q \text{Tr}[\mathcal{M}^{-1}(\hat{\sigma}_q)\Gamma_O], \quad (10)$$

where $\mathcal{M} \in \mathbb{L}(\mathcal{H}_a \otimes \mathcal{H})$ is linear. For classical shadows obtained from global random circuits or shallow random circuits, $\mathcal{M}$ can be calculated efficiently [29,32,39,44]. It can be proven [29] that when $Q = \mathcal{O}(\log(I)\epsilon_1^{-2})$, we can estimate all the I expectation values to ϵ_1 error with high probability using the global Clifford classical shadow, i.e., $|\langle \Gamma_{O_i} \rangle_{\text{est}} - \text{Tr}(\rho(t)\Gamma_{O_i})| < \epsilon_1$ for all Γ_{O_i} . In addition, one can also read out the imaginary part, $\text{Im} \langle \phi_0 | O | \phi_0(t) \rangle$, by setting $\Gamma_O = i(\text{Id}_a \otimes O) |1, \phi_0\rangle \langle 0, \phi_\perp| + \text{h.c.}$

To realize global Clifford random unitaries, one needs linear-depth quantum circuits. This could pose a serious challenge on the near-term quantum platforms due to the severe two-qubit gate errors. Fortunately, recent finding shows that shallow quantum circuits with depth $\mathcal{O}(\log L)$ can form global random unitaries on L qubits [45]. Moreover, experiments have demonstrated that classical shadows using log-depth quantum circuits can be made robust against various quantum errors by ongoing theoretical advancements [40]. These emerging developments enable one to fully leverage the robust shallow shadow technique in experiments to achieve a low measurement overhead. In particular, since Γ_O has the low-rank property, it falls within the class of observables that shallow shadows have been shown to predict reliably, even in the presence of realistic quantum noise. Without loss of generality, we will focus on the global Clifford classical shadow tomography

for MODMD, assuming access to high-fidelity scrambling unitaries.

3. Symmetry-simplified time evolution circuit

For many-body systems, we can conveniently initialize the states $|\phi_0\rangle$ and $|\phi_\perp\rangle$ in Eq. (6) according to particular symmetry sectors of the Hamiltonian H . The presence of symmetry may *entirely eliminate* the need for controlled time evolution, enabling the use of standard uncontrolled evolution for significantly more native and, thus, compact circuit implementation on hardware.

For the case of molecular Hamiltonians, we note that it is natural to prepare $|\phi_0\rangle$ and $|\phi_\perp\rangle$ with distinct fermionic occupations which, after second quantization, correspond to two computational basis states with distinct Hamming weights (under Jordan-Wigner encoding). In particular, if we initialize $|\phi_\perp\rangle$ to be the vacuum state $|00\cdots 0\rangle$ with zero particle occupation, it will stay invariant under the time evolution: $e^{-iHt}|\phi_\perp\rangle = |\phi_\perp\rangle$. This allows us to implement uncontrolled time evolutions on the system registries alone in actual experiments, provided that the observables O_i do not induce any transitions between the particle sector of interest and the vacuum state.

For general local Hamiltonians, including those arising in familiar spin systems, the requirement of controlled time evolution can also be relaxed by exploiting canonical anti-commutation relations between Pauli operators and grouped Hamiltonian terms. To illustrate, we restrict ourselves to a Hamiltonian H consisting of a single group of mutually commuting Pauli terms; the same argument generalizes to many Hamiltonian terms under Trotterized dynamics. Notably, conjugating the time evolution with an anti-commuting Pauli operator $\tilde{P}$ (a chiral symmetry) effectively flips the sign of evolution time, i.e.,

$$\begin{bmatrix} \text{Id} & 0 \\ 0 & e^{-iHt} \end{bmatrix} = e^{-iH\frac{t}{2}} \begin{bmatrix} \tilde{P} & 0 \\ 0 & \text{Id} \end{bmatrix} e^{-iH\frac{t}{2}} \begin{bmatrix} \tilde{P} & 0 \\ 0 & \text{Id} \end{bmatrix}, \quad (11)$$

which constructs the controlled unitaries without directly controlling the Hamiltonian. This symmetry-based trick helps simplify circuit implementation; see [46] for details as well as concrete examples.

B. Main algorithm

With our efficient shadow implementation, we can estimate the density operator $\rho(t)$ and the associated expectation values for an arbitrary pool of operators $\{O_i\}_{i=1}^I$. By doing so, we facilitate an exponential expansion of the signal subspace relative to the measurement cost for the shadow reconstruction, as estimating I observables only requires $Q = \mathcal{O}(\log(I))$ samples $\{\hat{\sigma}_q\}_{q=1}^Q$. The collection of

expectation values takes the form,

$$\vec{s}(t) = \begin{bmatrix} \langle \phi_0 | O_1 | \phi_0(t) \rangle \\ \langle \phi_0 | O_2 | \phi_0(t) \rangle \\ \vdots \\ \langle \phi_0 | O_I | \phi_0(t) \rangle \end{bmatrix} = \sum_{n=0}^{N-1} \vec{c}_n e^{-iE_n t}, \quad (12)$$

for $\vec{c}_n \in \mathbb{C}^I$ with $c_{n,i} = \sum_{m=0}^{N-1} \alpha_m^* \alpha_n \langle \psi_m | O_i | \psi_n \rangle$.

Given our access to real-time expectations $\vec{s}_k = \vec{s}(k\Delta t)$ sampled at time step Δt , we formulate a LS problem as the multidimensional variant of Eq. (5),

$$\underbrace{\begin{bmatrix} \vec{s}_1 & \vec{s}_2 & \cdots & \vec{s}_{K+1} \\ \vec{s}_2 & \vec{s}_3 & \cdots & \vec{s}_{K+2} \\ \vdots & \vdots & \ddots & \vdots \\ \vec{s}_d & \vec{s}_{d+1} & \cdots & \vec{s}_{K+d} \end{bmatrix}}_{\mathbf{X}' \in \mathbb{R}^{dl \times (K+1)}} \stackrel{\text{LS}}{=} A \underbrace{\begin{bmatrix} \vec{s}_0 & \vec{s}_1 & \cdots & \vec{s}_K \\ \vec{s}_1 & \vec{s}_2 & \cdots & \vec{s}_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ \vec{s}_{d-1} & \vec{s}_d & \cdots & \vec{s}_{K+d-1} \end{bmatrix}}_{\mathbf{X} \in \mathbb{R}^{dl \times (K+1)}}, \quad (13)$$

where $(\mathbf{X}, \mathbf{X}')$ are time-shifted data matrices containing the observable evaluations through shadows. The system matrix $A \in \mathbb{C}^{dl \times dl}$ is now block companion with dl^2 free parameters, i.e., its last I rows. In the special case where the observable pool contains a single operator O_1 , namely the identity Id on the system Hilbert space, our approach reduces to the original ODMD setting.

The system matrix A captures the evolution of multi-dimensional expectations propagated by unitary dynamics $e^{-iH\Delta t}$. Hence, the eigenenergies and corresponding eigenstates of the Hamiltonian H , as the generator of the dynamics, can be simply estimated via the eigenvalue decomposition of A ,

$$A = \sum_{\ell=0}^{dl-1} \tilde{\lambda}_\ell \Psi_{R,\ell} \Psi_{L,\ell}, \quad (14)$$

where $\tilde{\lambda}_\ell$ give the DMD eigenvalues while $(\Psi_{R,\ell}, \Psi_{L,\ell})$ denote the corresponding right and left eigenvectors. We note that $\tilde{\lambda}_\ell \in \mathbb{C}$ since A , being block companion, is not Hermitian.

We can read off our eigenpair approximations from the ordering of the phases $\arg(\tilde{\lambda}_\ell)$. Without loss of generality, we assume that the phases are arranged in a descending order, $\arg(\tilde{\lambda}_0) \geq \arg(\tilde{\lambda}_1) \geq \cdots$, such that the eigenvalue with the maximal phase, $\tilde{\lambda}_0 := |\tilde{\lambda}_0| e^{-i\tilde{E}_0 \Delta t} \approx e^{-iE_0 \Delta t}$, encodes the DMD approximation $\tilde{E}_0$ to the exact ground-state energy E_0 . Likewise, the eigenvalue $\tilde{\lambda}_1$ provides an approximation to the first excited-state energy E_1 . The eigenstates, on the other hand, can be accessed from the DMD eigenvectors $\Psi_{L,n} = [z_{0,n}, z_{1,n}, \dots, z_{dl-1,n}]$. The left

ALGORITHM 1. MODMD eigenenergy estimation.

Input: Reference state $|\phi_0\rangle$, operator pool $\{O_i\}_{i=1}^I$, time step Δt , noise threshold $\tilde{\delta}$.
Output: Estimated energies $\tilde{E}_n$ and eigenstates $|\tilde{\psi}_n\rangle$.

```

1  $k \leftarrow 0$ 
  while not converged do
2    $\vec{s}_k \leftarrow e^{-iHk\Delta t} |\phi_0\rangle$ ; /* classical shadows */
3    $\mathbf{X}, \mathbf{X}' \leftarrow \text{Hankel}(\vec{s}_0, \vec{s}_1, \dots, \vec{s}_k)$ ; /* Eq. (13) */
4    $\mathbf{X}_{\tilde{\delta}} \leftarrow \mathbf{X}$ ; /* least-squares regularization */
5    $A \leftarrow \mathbf{X}' \mathbf{X}_{\tilde{\delta}}^+$ ; /* update system matrix */
6    $\tilde{E}_n \leftarrow -\text{Im} \log(\tilde{\lambda}_n) / \Delta t$ ; /* energies */
7    $|\tilde{\psi}_n\rangle \leftarrow \sum_{(a,i)} z_{(a,i),n} O_i e^{-iHa\Delta t} |\phi_0\rangle$ ; /* states */
8    $k \leftarrow k + 1$ 
  end

```

eigenvectors satisfy the eigenvalue equation,

$$\Psi_{L,n} \mathbf{X}' = \tilde{\lambda}_n \Psi_{L,n} \mathbf{X}, \quad (15)$$

where Eq. (15) can be viewed as an equality relating the matrix elements of $\mathbf{X}$ and $\mathbf{X}'$. Such an equality restricted to the first columns of the data matrices, for example, implies

$$\Psi_{L,n} \begin{bmatrix} \vec{s}_1 \\ \vec{s}_2 \\ \vdots \\ \vec{s}_d \end{bmatrix} = \tilde{\lambda}_n \Psi_{L,n} \begin{bmatrix} \vec{s}_0 \\ \vec{s}_1 \\ \vdots \\ \vec{s}_{d-1} \end{bmatrix}, \quad (16)$$

which can be expressed in terms of the eigenvector coordinates and real-time observables,

$$\begin{aligned} \langle \phi_0 | e^{iH\Delta t} \left(\sum_{a=0}^{d-1} \sum_{i=1}^I z_{(a,i),n}^* e^{iHa\Delta t} O_i |\phi_0\rangle \right) \\ = \tilde{\lambda}_n^* \langle \phi_0 | \left(\sum_{a=0}^{d-1} \sum_{i=1}^I z_{(a,i),n}^* e^{iHa\Delta t} O_i |\phi_0\rangle \right), \end{aligned} \quad (17)$$

where $z_{(a,i),n} := z_{aI+i-1,n}$ are the vectorized coefficients of $\Psi_{L,n}$. Since $\tilde{\lambda}_n \approx e^{-iE_n\Delta t}$, the dynamic mode above closely follows the eigenstate oscillation driven at desired frequency, $\langle \phi_0 | e^{iH\Delta t} |\psi_n\rangle = e^{iE_n\Delta t} \langle \phi_0 | \psi_n\rangle$. Hence, we can approximate $|\psi_n\rangle$ by Eq. (17),

$$|\tilde{\psi}_n\rangle = \sum_{a=0}^{d-1} \sum_{i=1}^I z_{(a,i),n}^* e^{iHa\Delta t} O_i |\phi_0\rangle, \quad (18)$$

where $z_{(a,i),n}$ are now scaled to give a normalized state. Any eigenstate properties can in turn be derived in terms of the pool of operators $\{O_i\}_{i=1}^I$ and time-evolved states $\{e^{-iHa\Delta t} |\phi_0\rangle\}_a$.

The formal solution to Eq. (13) entails computing the Moore-Penrose pseudo-inverse $\mathbf{X}^+ = (\mathbf{X}^\dagger \mathbf{X})^{-1} \mathbf{X}^\dagger$ of the

data matrix $\mathbf{X}$. To ensure stability and filter out perturbative noise, we employ the following truncated singular value decomposition (SVD) of the data matrix:

$$\mathbf{X}_{\tilde{\delta}} = \sum_{\ell: \sigma_\ell > \tilde{\delta} \sigma_{\max}} \sigma_\ell \mathbf{u}_\ell \mathbf{v}_\ell^\dagger, \quad (19)$$

where σ_ℓ and $(\mathbf{u}_\ell, \mathbf{v}_\ell)$ are the singular values and vectors respectively. Here $\tilde{\delta} > 0$ is a truncation threshold defined relative to the largest singular value $\sigma_{\max} = \max_\ell \sigma_\ell$ of $\mathbf{X}$. This thresholding procedure, which removes smaller singular values associated with noise, serves to regularize the LS problem of Eq. (13).

In summary, our shadow-based algorithm requires as input the selected observables $\{O_i\}_{i=1}^I$, time step Δt , and singular value threshold $\tilde{\delta}$. The algorithm is described in Algorithm 1, which we call the MODMD.

C. Selection of hyperparameters

The performance of our algorithm clearly relies on the choice of the input parameters. We first remark that the convergence of MODMD, just as any subspace method, is influenced by the choice of reference state $|\phi_0\rangle$, where the estimation error scales inversely with the squared overlap with the eigenstates of interest. Although having a larger overlap is ideal, it suffices to prepare $|\phi_0\rangle$ using simplified single-particle calculations, which generally translates to a sparse sum of product states. Another possibility is to prepare $|\phi_0\rangle$ as a matrix product state (MPS), which is a dense sum of product states but with highly structured coefficients that allow efficient classical manipulation and storage. With this foundation, we now turn our attention to the optimization of other key hyperparameters.

Time step. The time step Δt determines algorithmic convergence as it sets the separation of the eigenphases $e^{-iE_n\Delta t}$ over the unit circle. A larger time step is thereby advantageous for better resolving the nearby eigenphases, until an aliasing ambiguity arises when $\Delta_{0,N-1} \Delta t \geq 2\pi$, where the energy gaps are defined as $\Delta_{m,n} = E_n - E_m$ (so $\Delta_{0,N-1}$ is the spectral range). Additionally, Δt must satisfy a further compatibility condition [9],

$$\Delta t \lesssim \frac{2\pi}{\Delta_{0,N-1} + \max_{0 \leq n < N_{\text{eig}}} (\Delta_{n,n+1} - \Delta_{0,n})}, \quad (20)$$

where N_{eig} counts the eigenenergies $\{E_0, \dots, E_{N_{\text{eig}}-1}\}$ of interest. For unambiguous and appropriately ordered (so that $\tilde{E}_n$ approximates E_n) estimation, we suggest bounding the spectral range of the Hamiltonian and then linearly shifting the range to be in $[-C\pi, C\pi]$ for some positive constant $C < 1$. In this case, the time step can be set to 1, uniquely restricting eigenangles $E_n \Delta t$ in the 2π -window $(-\pi, \pi)$. Note that Eq. (20) holds for the relevant energy lower and upper bounds so the exact eigenenergies do not need to be known in advance.

In implementing the time evolution on quantum device via, for example, a Trotter-Suzuki formula, the time step Δt is closely tied to the quantum resources requirements. Since we aim to extract eigenvalues of $e^{-iH\Delta t}$ through the observable time-series data, approximating $e^{-iH\Delta t}$ to a target accuracy ϵ_{TS} sets the required circuit complexity. Access to more observables accelerates our search toward the target eigenspaces, providing the flexibility to employ a potentially smaller Δt . This in turn reduces the overall circuit complexity. We defer detailed analysis of the first-order Trotter scheme to Sec. IV C 2.

Observables. The choice of observables $\{O_i\}_{i=1}^I$ also plays a critical role in steering convergence. Strategically choosing observables helps construct a subspace enriched with low-energy spectral content, which can significantly influence performance of MODMD. While expanding the observable pool improves quality of spectral information, it also raises classical postprocessing costs, since the LS problem in Eq. (13) grows proportionally in size with the number of observables I . Such a trade-off is important to keep in mind, though leveraging many observables incurs little overhead in MODMD and remains highly beneficial for optimizing quantum resource allocation.

Drawing on the signal subspace intuition from Eq. (12), the observables should be spectrally “independent” in the sense that the matrix of oscillation amplitudes,

$$\mathbf{c} = \begin{bmatrix} c_{0,1} & c_{0,2} & \cdots & c_{0,I} \\ c_{1,1} & c_{1,2} & \cdots & c_{1,I} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N-1,1} & c_{N-1,2} & \cdots & c_{N-1,I} \end{bmatrix}, \quad (21)$$

maintains a full column rank of I . Otherwise, the multi-dimensional signals contain redundant information. As a convention, we always fix $O_1 = \text{Id}$ corresponding to the OMD algorithm.

It is worth noting that the shadow reconstruction involves strictly classical computation, thus requiring each observable O_i to have a sparse representation in the Pauli basis,

$$O_i = \sum_{v=1}^{M_i} \kappa_{i,v} P_{i,v}, \quad (22)$$

where $\{P_{i,v}\}_{v=1}^{M_i}$ is a set of $M_i = \mathcal{O}(\text{poly}(L))$ distinct L -qubit Pauli strings with associated weights $\{\kappa_{i,v}\}_{v=1}^{M_i}$. Although it is rather convenient to select the observables O_i randomly from the 4^L Pauli strings, the resulting real-time signals may suffer from diminished utility because of probable suppression of the target oscillation amplitudes $|c_{n,i}| = |\langle \psi_n | \phi_0 \rangle| |\langle \phi_0 | O_i | \psi_n \rangle|$. For ground-state estimate, this occurs when $|\langle \phi_0 | O_i | \psi_0 \rangle| \approx 0$, which deteriorates the quality of the signals. As an example, a 1-local Pauli X operator changes the Hamming weight of reference state,

and can hence lead to zero amplitude if the Hamiltonian preserves the total Z -spin.

Alternatively, we propose the systematic generation of observable pools starting from the problem Hamiltonian,

$$H = \sum_{v=1}^M \kappa_v P_v, \quad (23)$$

where the terms are ordered by the magnitude of their coefficients, $|\kappa_1| \geq \cdots \geq |\kappa_M|$. Such sorting induces a family of partial sums, $\sum_{v \in \mathcal{V}} \kappa_v P_v$, with $\mathcal{V} \subseteq \{1, \dots, M\}$ labeling a subset of Pauli strings. Our observables can be selected from these partial sums based on importance of the M Pauli weights $\{\kappa_v\}_{v=1}^M$. That is, we may consider $O_1 = \sum_{v=1}^M \kappa_v P_v = H$, $O_2 = \sum_{v=1}^{M-1} \kappa_v P_v$, etc. Let us assume that the target Hamiltonian is linearly shifted, as discussed for selection of time step, such that the low-lying energies are large in magnitude. In contrast to randomly selecting a Pauli string, the low energy amplitudes of interest, for example, $|c_{0,1}| = |\langle \psi_n | \phi_0 \rangle|^2 |E_0|$, are effectively “magnified” relative to amplitudes associated with energies interior in the spectrum.

We remark that integer powers of the partial sums and their linear combinations can also be desired additions to the observable pool for generating high-quality real-time signals. This imposes no computational bottleneck since the observable predictions can be performed in a parallel and distributed manner on classical computers, and the variance only depends on the 1-norm, $\max_{1 \leq i \leq I} \|\vec{k}_i\|_1$, of the Pauli weights (see Sec. IV C).

Time-delayed embedding. The dimensions (d, K) define the shape of the block Hankel matrices, $\mathbf{X}$ and $\mathbf{X}'$, assembled from the observable time-series data. To avoid potential instabilities due to rank deficiency, we require that $d \leq K$. For consistency, we fix $K/d = \frac{5}{2}$ throughout our manuscript, which has been observed to yield reliable empirical performance (for example, relative to $K/d = 1$). We note that more refined (d, K) choices may be achieved from minimizing known upper bound on the conditioning number of the Hankel data matrices [47].

SVD threshold. The SVD threshold $\tilde{\delta}$ is subject to the noise level, which can be controlled as we collect real-time data via low-rank shadow techniques. As a practical default, we set $\tilde{\delta} \approx 10\epsilon_{\text{noise}}$, placing it roughly an order of magnitude above the statistical uncertainty ϵ_{noise} due to shot noise. Setting $\tilde{\delta}$ too high (low) risks erasing relevant eigenmode contributions (retaining excessive noise).

Parametric validation. To ensure the robustness of MODMD in different parametric regimes, we conduct a comprehensive sweep over key hyperparameters. Results illustrating the effect of time step, number of observables, and SVD threshold are shown in Appendix H. The final accuracy of our algorithm improves with larger time steps—provided aliasing is suppressed—and larger

observable pools. While accuracy depends nonmonotonically on the SVD threshold as expected, it remains rather insensitive to reasonable (over an-order-of-magnitude) variations in the exact threshold value.

D. Hamiltonian properties beyond energies

The MODMD framework extends beyond estimating the eigenenergies, providing access to useful Hamiltonian properties including eigenstate properties and dynamical responses. To illustrate, we first recall from Sec. III B that the MODMD eigenpairs $(\tilde{\lambda}_\ell, \Psi_{L,\ell})$ of system matrix A can also be leveraged to construct approximations $|\tilde{\psi}_n\rangle$ to the low-lying eigenstates $|\psi_n\rangle$, which we confirm using two fidelity metrics within Appendix I. Such eigenstate information can be explicitly expressed and implemented as a linear combination of time evolutions on quantum hardware and is typically inaccessible from the outputs of recent signal processing phase estimation algorithms. As a consequence, arbitrary eigenstate properties can be predicted as

$$f(|\psi_n\rangle) \approx f(|\tilde{\psi}_n\rangle), \quad (24)$$

for any scalar-valued function $f: \mathcal{H} \rightarrow \mathbb{C}$. This predictive capability straightforwardly applies to any state property within the low-lying energy subspace.

In addition, the state shadows $\{\hat{\sigma}_q\}_{q=1}^Q$ stored on the classical computer can be utilized to calculate the time-dependent expectations $\langle\phi_0(t)|O_i|\phi_0(t)\rangle$ (note that they differ from $\langle\phi_0|O_i|\phi_0(t)\rangle$). Specifically, we recognize that

$$\langle O_i(t) \rangle_{\text{est}} = \frac{2}{Q} \sum_{q=1}^Q \text{Tr}[\mathcal{M}^{-1}(\hat{\sigma}_q)(|1_a\rangle\langle 1_a| \otimes O_i)] \quad (25)$$

gives an unbiased estimate of $\langle O_i(t) \rangle = \langle\phi_0(t)|O_i|\phi_0(t)\rangle$. These additional data can be taken as an input to a separate set of MODMD calculations.

Augmenting the capability of real-time subspace methods to compactly represent the eigenstates, the general framework of DMD moreover enables the prediction of system dynamics over longer timescales [48–50]. A multiobservable signal $\vec{s}(t)$, composed of time correlation functions, contains essential dynamical fingerprint that characterizes, for instance, how a many-body system reacts to an external perturbation in the linear-response regime [51,52]. Here, the response represents a dynamical property distinct from the stationary properties governed by a single eigenstate. To further study these dynamical properties, we analyze how time correlation functions are predicted within the MODMD framework.

First, the one-point correlators $\langle\phi_0|O_i e^{-iHt}|\phi_0\rangle$ can be propagated forward in time in increments of Δt : Eq. (13) suggests that integer powers of the system matrix A can be used to (approximately) fast-forward the observables beyond the measurement window. That is, for any integer $k > K + d$, MODMD predicts the dynamics at a later time $k\Delta t$ via

$$\langle\phi_0|O_i e^{-iHk\Delta t}|\phi_0\rangle \stackrel{\text{MODMD}}{\approx} \mathbf{e}_i^\dagger A^{k-K-d+1} \mathbf{x}_K, \quad (26)$$

where $\mathbf{x}_K \in \mathbb{R}^{dl}$ is the last column of the data matrix $\mathbf{X}$ and $\mathbf{e}_i = [0, \dots, 0, 1, 0, \dots, 0]^\top \in \mathbb{R}^{dl}$ labels the standard basis vector with the $[(d-1)I + i]$ th entry equal to 1. The system matrix A functions as the forward Δt -propagator for evolving observables in time, establishing A^{-1} as the respective backward propagator. Such predictive abilities of MODMD are demonstrated in Sec. VD.

Next, the two-point correlators can be approximated, with MODMD eigenenergy and eigenstate estimates, as

$$\langle\phi_0|O_i(k\Delta t)O_j(l\Delta t)|\phi_0\rangle \stackrel{\text{MODMD}}{\approx} \sum_{n=0}^{dI-1} e^{-i\tilde{E}_n(k-l)\Delta t} \langle\phi_0|e^{iHk\Delta t}O_i|\tilde{\psi}_n\rangle \langle\tilde{\psi}_n|O_j e^{-iHl\Delta t}|\phi_0\rangle, \quad (27)$$

where $O(t) = e^{iHt}Oe^{-iHt}$ labels a time-evolved operator in the Heisenberg picture. The behavior of the one-point correlators $\langle\tilde{\psi}_n|O_i e^{-iHk\Delta t}|\phi_0\rangle$ over longer times can then be predicted via short-time snapshots $\langle\tilde{\psi}_n|O_i e^{-iHa\Delta t}|\phi_0\rangle$ for $0 \leq a \leq K + d$. Notice that the inner products involve the approximate eigenstates $|\tilde{\psi}_n\rangle$ alongside $|\phi_0\rangle$. Thus by Eq. (18), measuring these snapshots in general incurs an additional cost of $\mathcal{O}(I^2(K + d)^2)$, which makes predictions of two-point correlators more demanding. However, in the case where $|\phi_0\rangle = |\tilde{\psi}_m\rangle$, we can fully leverage information in the signal subspace with no extra measurements

as, for instance, $e^{-iHa\Delta t}|\tilde{\psi}_0\rangle \approx e^{-i\tilde{E}_0 a\Delta t}|\tilde{\psi}_0\rangle$. This reduces the cost back to $\mathcal{O}(\log(I)(K + d))$. Importantly, the favorable $\log(I)$ scaling is absent in conventional subspace methods where matrix elements of the forms $\langle\phi_0|O_i e^{-iHk\Delta t}O_j|\phi_0\rangle$ or $\langle\phi_0|e^{iHk\Delta t}O_i e^{-iH\Delta t}O_j e^{-iHl\Delta t}|\phi_0\rangle$ can be measured at the costs of at least $\mathcal{O}(I \log(I)(K + d))$ or $\mathcal{O}(I^2(K + d)^2)$, respectively. While MODMD offers versatile capabilities discussed in this section, our work focuses on estimation of eigenenergies, leaving the specific explorations of other applications for future studies.

IV. THEORETICAL GUARANTEES

We establish in this section fundamental connections between MODMD and modern spectral approaches, furnishing a theoretical framework that guarantees its convergence. We first explicitly show a speedup of MODMD over ODMD by having an expansive pool of observables. Next, we exploit the linearity of quantum dynamical evolution and consider MODMD as a multireference scheme within a suitably defined function space through Koopman operator analysis [53–58]. These two analytical viewpoints reinforce each other, underpinning the reliable performance of our algorithm for excited-state problems.

A. Multiobservable dynamic mode decomposition

For the LS problem of Eq. (13) with observables $\{O_i\}_{i=1}^I$, the multivariate solution is

$$A = \left[\begin{array}{c|ccc} 0 & & & \\ \vdots & & \text{Id}_{d-1} \otimes \text{Id}_I & \\ 0 & & & \\ \hline -A_0 & -A_1 & \cdots & -A_{d-1} \end{array} \right], \quad (28)$$

where each $A_\ell \in \mathbb{R}^{I \times I}$ for $\ell = 0, 1, \dots, d-1$ represents a submatrix and Id_{d-1} and Id_I are identity operators of respective dimensions. The block companion structure of the system matrix $A \in \mathbb{C}^{dI \times dI}$ immediately follows from the block Hankel structure of the matrices $\mathbf{X}$ and $\mathbf{X}'$. The multiobservable system matrix has a characteristic polynomial,

$$C_A(z) = \det(z - A) = \det\left(\sum_{\ell=0}^d z^\ell A_\ell\right), \quad (29)$$

where $A_d \equiv \text{Id}_I$. Hence, the roots of $C_A(z)$ correspond precisely to the dI eigenvalues of the system matrix A . To understand the convergence of the A -spectrum to the actual eigenphases, we first consider two limiting cases. For the case $(d, I) = (N, 1)$, we essentially recover the celebrated Prony's method [59] as the single-observable ODMD with $A_\ell \in \mathbb{C}$, such that $C_A(z) = \prod_{n=0}^{N-1} (z - \lambda_n)$ for $\lambda_n = e^{-iE_n \Delta t}$. While for the complementary case of $(d, I) = (1, N)$, $C_A(z) = \det(z - \bar{A})$ where $\bar{A} \in \mathbb{C}^{N \times N}$ satisfies the linear homogeneous equation,

$$\begin{bmatrix} \langle \phi_0 | O_1 | \phi_0(t + \Delta t) \rangle \\ \langle \phi_0 | O_2 | \phi_0(t + \Delta t) \rangle \\ \vdots \\ \langle \phi_0 | O_N | \phi_0(t + \Delta t) \rangle \end{bmatrix} = \bar{A} \begin{bmatrix} \langle \phi_0 | O_1 | \phi_0(t) \rangle \\ \langle \phi_0 | O_2 | \phi_0(t) \rangle \\ \vdots \\ \langle \phi_0 | O_N | \phi_0(t) \rangle \end{bmatrix}, \quad (30)$$

for any time $t \in \mathbb{R}$. Such a matrix $\bar{A}$ indeed exists because the real-time expectations $\langle \phi_0 | O_i | \phi_0(t) \rangle$ reside within the N -dimensional space spanned by single-frequency signals

$e^{-iE_n t}$ driven at the individual eigenfrequencies E_n . To investigate the general case $1 \ll dI \ll N$, we start with a simpler version of Eq. (13), assuming the d matrix blocks to be diagonal,

$$A_\ell = \text{Diag} \begin{bmatrix} A_{\ell,11} & & & \\ & A_{\ell,22} & & \\ & & \ddots & \\ & & & A_{\ell,II} \end{bmatrix}, \quad (31)$$

where $A_{\ell,ij} \equiv 0$ for $i \neq j$. In this case, the individual LS residuals associated with the observables O_i are independent, and we can show that the resulting MODMD estimates are bounded by the I single-observable ODMD estimates, e.g.,

$$\min_{1 \leq i \leq I} |\tilde{E}_{i,0} - E_0| \leq |\tilde{E}_0 - E_0| \leq \max_{1 \leq i \leq I} |\tilde{E}_{i,0} - E_0|, \quad (32)$$

where $\tilde{E}_0$ and $\tilde{E}_{i,0}$ indicate, respectively, the ground-state energy estimate using the entire observable pool (the full system matrix A) and one single observable (only i th row $A_{\ell,ii}$ of the submatrices A_ℓ).

More interesting convergence arises when the I single-observable residuals are coupled to one another, allowing reductions in the individual residuals due to the flexibility of off-diagonal elements in the submatrices A_ℓ . Specific scenarios in which MODMD substantially improves upon the ODMD residuals are considered within Appendix C. Intuitively, we expect a reduced total residual to, in turn, improve the eigenenergy estimates, where Eq. (32) holds accordingly with tighter lower and upper bound—ideally, both approaching zero. Moreover, the reduction in total LS residual from MODMD indicates a more expressive system matrix as a proxy for the underlying dynamics, which is crucial for accurately predicting a multiobservable signal over longer times as discussed in Sec. III D.

B. Koopman operator analysis

We now establish convergence properties of MODMD from a functional-theoretic perspective. Our analysis will revolve around the study of the Koopman operator, a pivotal mathematical object in understanding the complexities of a dynamical system. The Koopman operator $\mathcal{K}_{\Delta t}$ probes the underlying dynamics of a system by acting on scalar-valued functions. For any function $f : \mathcal{H} \rightarrow \mathbb{C}$, its action is

$$\mathcal{K}_{\Delta t}[f](|\phi\rangle) = f(e^{-iH\Delta t}|\phi\rangle), \quad (33)$$

which gives a push-forward of the dynamics via the time evolution operator $e^{-iH\Delta t}$. For a quantum-dynamical system, we first observe that $f_n(|\phi\rangle) = \langle \psi_n | \phi \rangle$ constitutes an

eigenfunction of $\mathcal{K}_{\Delta t}$ since

$$\mathcal{K}_{\Delta t}[f_n] = e^{-iE_n\Delta t}f_n, \quad (34)$$

where $e^{-iE_n\Delta t}$ is the corresponding eigenvalue. Therefore scalar functions of the form $f_{\psi,O}(|\phi\rangle) := \langle\psi|O|\phi\rangle = \sum_{n=0}^{N-1} \langle\psi|O|\psi_n\rangle \langle\psi_n|\phi\rangle$ lie in a $\mathcal{K}_{\Delta t}$ -invariant subspace spanned by these eigenfunctions. By choosing $|\psi\rangle$ and O , we recast the task of identifying Hamiltonian eigenmodes as an equivalent, finite-dimensional Koopman eigenvalue problem. We use $\vec{g} = [g_1, g_2, \dots, g_D]^\top$ to denote a vector of D distinct scalar functions g_i in the invariant function subspace, all taking the form of $f_{\psi,O}$.

We first examine the case where $g_i = (\mathcal{K}_{\Delta t})^{i-1}[f_{\psi,\text{Id}}]$, which gives rise to the ODMD approach when $D = d$.

We seek to determine the closest approximation $\mathbf{K}_{\Delta t} \in \mathbb{C}^{d \times d}$ of the Koopman operator when restricted to the invariant subspace. The closest-fitting problem has a LS formulation,

$$\mathbf{K}_{\Delta t} = \arg \min_{\mathbf{K} \in \mathbb{C}^{d \times d}} \sum_{k=0}^K \|\mathcal{K}_{\Delta t}[\vec{g}](|\phi_k\rangle) - \mathbf{K}\vec{g}(|\phi_k\rangle)\|_2^2, \quad (35)$$

where $\mathcal{K}_{\Delta t}$ acts component-wise on $\vec{g}$, and the ℓ^2 -residual is being minimized with respect to some states $\{|\phi_k\rangle\}_{k=0}^K$ sampled from the Hilbert space. Formally, we can express Eq. (35) as a matrix equation

$$\underbrace{\begin{bmatrix} \vec{g}(e^{-iH\Delta t}|\phi_0\rangle) & \dots & \vec{g}(e^{-iH\Delta t}|\phi_K\rangle) \end{bmatrix}}_{\mathbf{G}' \in \mathbb{C}^{d \times (K+1)}} = \mathbf{K}_{\Delta t} \underbrace{\begin{bmatrix} \vec{g}(|\phi_0\rangle) & \dots & \vec{g}(|\phi_K\rangle) \end{bmatrix}}_{\mathbf{G} \in \mathbb{C}^{d \times (K+1)}}, \quad (36)$$

whose solution $\mathbf{K}_{\Delta t} = \mathbf{G}'(\mathbf{G})^+$ involves computing the pseudo-inverse $\mathbf{G}^+$. Utilizing $\mathbf{G}^+ = (\mathbf{G}^\dagger \mathbf{G})^{-1} \mathbf{G}^\dagger$, we can directly show that $\mathbf{K}_{\Delta t}$ yields an equivalent factorization $\mathbf{K}_{\Delta t} = \mathbf{W}(\mathbf{Z})^+$, where $\mathbf{Z}, \mathbf{W} \in \mathbb{C}^{d \times d}$ are constructed by

$$\mathbf{Z} = \frac{1}{K+1} \sum_{k=0}^K \vec{g}(|\phi_k\rangle) \vec{g}^\dagger(|\phi_k\rangle), \quad (37)$$

$$\mathbf{W} = \frac{1}{K+1} \sum_{k=0}^K \vec{g}(e^{-iH\Delta t}|\phi_k\rangle) \vec{g}^\dagger(|\phi_k\rangle). \quad (38)$$

This alternative factorization is intimately related to the powerful Krylov approaches for operator diagonalization [60–62]. Suppose we sample the states $\{|\phi_k\rangle\}_k$ according to a probability measure μ defined over the Hilbert space. With sufficiently many samples, we have

$$\mathbf{Z}_{ij} \rightarrow \int_{|\phi\rangle \in \mathcal{H}} d\mu g_i g_j^* = \langle g_j, g_i \rangle_\mu, \quad (39)$$

$$\mathbf{W}_{ij} \rightarrow \int_{|\phi\rangle \in \mathcal{H}} d\mu \mathcal{K}_{\Delta t}[g_i] g_j^* = \langle g_j, \mathcal{K}_{\Delta t}[g_i] \rangle_\mu, \quad (40)$$

where the $L^2(\mu)$ inner product gives the continuum limit of Monte-Carlo averages in Eqs. (37) and (38). Therefore eigenpairs $(\lambda_{\mathbf{K}}, \vec{v}_{\mathbf{K}})$ of $\mathbf{K}_{\Delta t}$ should satisfy the generalized

eigenvalue equation,

$$\mathbf{W}\vec{v}_{\mathbf{K}} = \lambda_{\mathbf{K}} \mathbf{Z}\vec{v}_{\mathbf{K}}, \quad (41)$$

where $\mathbf{Z}$ and $\mathbf{W}$ can now be reinterpreted, respectively, as the matrix representation of the functional overlap and Koopman operator in the finite basis $\{g_1, \dots, \mathcal{K}_{\Delta t}^{d-1}[g_1]\}$. This special nonorthogonal basis is known as the order- d Krylov basis, and the associated subspace is the Krylov subspace $\mathcal{K}_d(g_1) \subset \text{span}\{f_n\}_n$.

Projecting the full Koopman eigenvalue problem onto this subspace of size $d \ll N$ allows an efficient retrieval of spectral information, including the extremal eigenvalues. In the ODMD algorithm, we choose $|\phi_k\rangle = e^{-iHk\Delta t}|\phi_0\rangle$ for a fixed state $|\phi_0\rangle$ and take the sampling measure μ to be the empirical measure (a sum of Dirac-delta measures),

$$\mu = \frac{1}{K+1} \sum_{k=0}^K \delta_{|\phi_k\rangle}, \quad (42)$$

along the orbit $(|\phi_0\rangle, e^{-iH\Delta t}|\phi_0\rangle, \dots, e^{-iHK\Delta t}|\phi_0\rangle)$. In this case, $\lim_{K \rightarrow \infty} \|\mathcal{K}_{\Delta t}[f]\|_{L^2(\mu)} = \lim_{K \rightarrow \infty} \|f\|_{L^2(\mu)}$ for a continuous function f (we expect the limit to exist [63] for a generic many-body Hamiltonian H and time step Δt), since

$$\lim_{K \rightarrow \infty} \langle \mathcal{K}_{\Delta t}[f], \mathcal{K}_{\Delta t}[f] \rangle_{\mu} = \lim_{K \rightarrow \infty} \frac{1}{K+1} \sum_{k=1}^{K+1} |f(e^{-iHk\Delta t} |\phi_0\rangle)|^2 = \lim_{K \rightarrow \infty} \frac{1}{K+1} \sum_{k=0}^K |f(e^{-iHk\Delta t} |\phi_0\rangle)|^2 = \lim_{K \rightarrow \infty} \langle f, f \rangle_{\mu}, \quad (43)$$

where the middle equality holds due to a vanishing difference, $\lim_{K \rightarrow \infty} (1/(K+1)) [|f(e^{-iH(K+1)\Delta t} |\phi_0\rangle)|^2 - |f(|\phi_0\rangle)|^2] = 0$. In other words, the Koopman operator $\mathcal{K}_{\Delta t}$, when restricted to the invariant subspace, is isometric and hence normal. By the spectral theorem, the $\mathcal{K}_{\Delta t}$ eigenfunctions $\{f_n\}_{n=0}^{N-1}$ are orthogonal to each other. Under this condition, exponentially rapid convergence of standard Krylov approach has been thoroughly analyzed [64], which aligns with our ODMD observations.

Now we extend the result above to eigenfunctions that are nonorthogonal, a regime of practical interest for finite K . In such general cases, the functional Krylov approach still converges exponentially, but its convergence depends explicitly on the conditioning of the eigenfunctions under the $L^2(\mu)$ inner product. In particular, we define

$$\mathcal{B}_{mn} = \frac{\langle f_m, f_n \rangle_{\mu}}{\|f_m\|_{L^2(\mu)} \|f_n\|_{L^2(\mu)}}, \quad (44)$$

where $\{f_n\}_{n=0}^{N-1}$ are mutually orthogonal if and only if $\mathcal{B}$ is the identity. Departures from orthogonality are captured by the condition number $\text{cond}(\mathcal{B}) = \|\mathcal{B}\|_2 \|\mathcal{B}^{-1}\|_2$ [65,66]. In practice, $K+1 < N$ always holds. Since the empirical $L^2(\mu)$ inner product only resolves functions through their values on the sampled orbit, the Gram matrix $\mathcal{B}$ has rank of at most $K+1$ and is thus rank deficient. In this regime, the relevant notion of conditioning is the one defined on the identifiable subspace, given by the intrinsic condition number $\text{cond}(\mathcal{B}) \mapsto \|\mathcal{B}\|_2 \|\mathcal{B}^+\|_2$. It reduces to the usual condition number when $\mathcal{B}$ is full rank and measures how stably a finite- K signal distinguishes resolvable Koopman channels.

In the single observable setting, we invoke the following theorem for the ground-state energy error.

Theorem 1. Let $\tilde{E}_0(d)$ be the approximate ground-state energy extracted in the d -dimensional function subspace spanned by $\{g_1, \mathcal{K}_{\Delta t}[g_1], \dots, \mathcal{K}_{\Delta t}^{d-1}[g_1]\}$. For $d \geq 1$, there exists time step Δt so that the error $\delta E_0(d) = \tilde{E}_0(d) - E_0$ is bounded by

$$\delta E_0(d) \leq \frac{\text{cond}(\mathcal{B})^{1/2} |\sin[(E_{N-1} - E_0)\Delta t]|}{\tilde{\epsilon}_{0 \rightarrow 1}^{2(d-1)} \Delta t} \tan^2 \Xi, \quad (45)$$

where $\cos^2 \Xi = |\langle f_0, g_1 \rangle_{\mu}|^2 / (\|f_0\|_{L^2(\mu)}^2 \|g_1\|_{L^2(\mu)}^2)$ specifies the squared overlap between the reference function g_1 and the true ground-state $\mathcal{K}_{\Delta t}$ eigenfunction, while

$\tilde{\epsilon}_{0 \rightarrow 1} = 1 + 3(E_1 - E_0)\Delta t / 2\pi \in [1, 2]$ characterizes the normalized spectral gap of the Hamiltonian H .

Proof. The proof is provided in Appendix D 1. ■

Note that $\text{cond}(\mathcal{B}) \rightarrow 1$ as $K \rightarrow \infty$ when eigenenergies are nondegenerate. By construction, $\mathcal{B}_{nn} \equiv 1$ and

$$\begin{aligned} |\mathcal{B}_{mn}| &= \frac{1}{K+1} \left| \sum_{k=0}^K e^{i(E_n - E_m)k\Delta t} \right| \\ &= \frac{|\sin(\frac{K+1}{2}\theta_{mn})|}{(K+1)|\sin(\frac{1}{2}\theta_{mn})|} \rightarrow 0, \end{aligned} \quad (46)$$

for $\theta_{mn} = (E_n - E_m)\Delta t \neq 0$. Consequently, $\mathcal{B}$ converges asymptotically to the identity operator and $\text{cond}(\mathcal{B}) \rightarrow 1$. To build intuition at $K+1 < N$, we turn to the harmonic oscillator where $E_n - E_m = (n - m)\Delta E$, yielding a clean toy model with a semidense spectrum. As an illustrative choice, setting $\Delta t = 2\pi / N\Delta E$ makes $\mathcal{B}$ circulant, allowing its eigenvalues $\{0, N/(K+1)\}$ to be computed exactly. In this case, $\|\mathcal{B}\|_2 \|\mathcal{B}^+\|_2 \equiv 1$.

In contrast, setting $\Delta t < 2\pi / N\Delta E$ leads to a less favorable conditioning: $\mathcal{B}$ loses its circulant symmetry and is only Toeplitz, causing its nonzero eigenvalues to disperse. As Δt keeps shrinking, the matrix becomes closer to rank-1. In this case, the conditioning deteriorates, demonstrating how near-resonances or overly fine time steps can impair mode separability.

On the other hand, if the ground-state energy admits a degeneracy, we can replace $\mathcal{B}$ by an effective Gram matrix constructed over distinct energy subspaces, ensuring that the conditioning factor remains finite. Degeneracy does not affect energy convergence, but it precludes resolution of individual eigenstates. Further details are provided in Appendix E.

Building on the single-observable Krylov idea (ODMD) above, the MODMD approach can thus be viewed as an enriched extension, where we allow for an extra degree of freedom in selecting multiple functions

$$\vec{g} = \begin{bmatrix} f_{\phi_0, \text{Id}} \\ f_{\phi_0, O_2} \\ \vdots \\ f_{\phi_0, O_I} \end{bmatrix} \implies \vec{g}(e^{-iHt} |\phi_0\rangle) = \vec{s}(t). \quad (47)$$

By leveraging the classical shadows, each quantum circuit $e^{-iHk\Delta t}$ originally capable of computing a single overlap

can now simultaneously compute the Koopman action on a vector $\vec{g}$ of scalar functions. Each function corresponds to a unique choice of observable. This key algorithmic improvement enables a *block* Krylov scheme, significantly accelerating the energy convergence. The rate of convergence in the MODMD setting is described by the theorem below.

Theorem 2. Let $\tilde{E}_n(d)$ be the approximate n th eigenenergy extracted in the dI -dimensional function subspace $\text{span}\{\vec{g}, \mathcal{K}_{\Delta t}[\vec{g}], \dots, \mathcal{K}_{\Delta t}^{d-1}[\vec{g}]\}$, and $\delta E_n(d) = \tilde{E}_n(d) - E_n$ be the approximation error. Consider the diagonal error matrix

$$\Delta_I(d) = \begin{bmatrix} \delta E_0(d) & & \\ & \ddots & \\ & & \delta E_{I-1}(d) \end{bmatrix}, \quad (48)$$

which contains approximations to the lowest I energies. For $d \geq 1$, there exists a time step Δt so that the spectral approximation $\Delta_I(d)$ is bounded by

$$\|\Delta_I(d)\|_2 \leq \frac{\text{cond}(\mathcal{B})^{1/2} |\sin[(E_{N-1} - E_0)\Delta t]|}{\tilde{\epsilon}_{I-1 \rightarrow I}^{2(d-1)} \Delta t} \|\tan^2 \Theta\|_2, \quad (49)$$

for the operator norm $\|\cdot\|_2$. Here Θ denotes the canonical angle between the two subspaces $\text{span}\{f_n : 0 \leq n \leq I-1\}$ and $\text{span}\{g_i : 1 \leq i \leq I\}$, which generalizes the squared overlap in Theorem D1 (see Appendixes D2 and D3). In the denominator, $\tilde{\epsilon}_{I-1 \rightarrow I} \in [1, 2]$ depends on the I th spectral gap $(E_I - E_{I-1})$ of the Hamiltonian H .

Proof. The proof is provided in Appendix D3. ■

Despite the formal similarity between the bounds from Theorems 1 and 2, we highlight that convergence in multiobservable setting offers a more favorable scaling for excited-state calculations. The ODMD bound for the I th lowest eigenenergy, as in standard subspace methods employing the reference state $|\phi_0\rangle$, includes an additional multiplicative factor of $\Omega(2^{I-1} \tilde{\epsilon}_{I-1 \rightarrow I}^{2I-2})$ [4]. Notably, this prefactor grows exponentially as we approach the higher excited states, counteracting the single-observable error decay from Theorem 1 unless $d > I$. While it is natural to extend the standard subspace methods using I reference states $|\phi_i\rangle \propto O_i |\phi_0\rangle$, the quantum cost of measuring the expectations $\langle \phi_i | e^{-iHt} | \phi_j \rangle$ is at least $\mathcal{O}(I \log(I))$. This is because each state $|\phi_i\rangle$ must be time-evolved, making it much more expensive than MODMD.

Beyond the sharper asymptotic error bound, the multiobservable structure of MODMD introduces qualitative advantages that single-observable ODMD cannot deliver. First, single-observable methods can recover at most one representative in a degenerate or near-degenerate energy

cluster, which often result in stalled convergence due to the collapse of spectral gap between closely spaced levels. By contrast, the block Krylov nature of MODMD allows multiple eigenmodes to be resolved simultaneously within a single invariant subspace, thus revealing richer spectral information [67]. Second, block subspace approaches are known to accelerate convergence for the excited states by approximating invariant subspaces rather than individual eigenstates [64], a property that MODMD inherits in the low-noise regimes. Finally, the use of many noncollinear observables improves the conditioning of the LS problem, effectively attenuating the regression errors and reducing the nonasymptotic measurement overhead for any desired precision. Since classical shadows let us acquire multiple observables at virtually the same quantum circuit depth, with scrambling unitaries adding a precision-independent depth increment, such gains come at no cost to maximal runtime. Together, these advantages show that MODMD offers not only quantitative error improvements but also qualitatively more expressive, rapid, and stable spectral resolution.

C. Error analysis

To account for noisy quantum hardware, we present in this subsection a preliminary error analysis. For purpose of basic premises, we examine error components of three distinct types, (1) the statistical uncertainty arising from the randomizing shadow protocol, (2) the deterministic error due to approximate compilation of the time evolution, and (3) the device-level error that originates in noise on near-term hardware. We show that the first two error types remain completely independent of the problem size across a practical range of observable selections. Specifically, the circuit approximation error grows polynomially with the maximal evolution time when we consider Trotterized evolution as viable proxy for exact dynamics. For device noise, we focus on the depolarization channel—an elementary error model applicable across many quantum platforms—and discuss how its rate can be quantified by the MODMD algorithm.

1. Statistical noise

First, we note that the prediction error associated with randomized measurements in the classical shadow techniques admits, in our case, a variance independent of the system size L . The classical shadows can predict observables, $\gamma_i(t) = \text{Tr}[\rho(t)\Gamma_{O_i}]$, of the form Eq. (8), where we recall that $\text{rank}(\Gamma_{O_i}) \leq 2$ even when $\text{rank}(O_i) = 2^L = N$. The low-rank property implies [29]

$$\mathbb{V}[\gamma_i] \leq \left\| \Gamma_{O_i} - \frac{\text{Tr}[\Gamma_{O_i}]}{N} (\text{Id}_a \otimes \text{Id}) \right\|_{\text{sh}}^2, \quad (50)$$

$$\stackrel{\mathcal{U}=\text{Cl}(N)}{=} 3\text{Tr}[\Gamma_{O_i}^2], \quad (51)$$

where $\|\cdot\|_{\text{sh}}$ is the shadow norm conditional on the measurement primitive and the trace $\text{Tr}[\Gamma_{O_i}^2]$ on the RHS of Eq. (51) can be unfolded by the defining relation $\Gamma_{O_i} = \text{Id}_a \otimes O_i |1, \phi_0\rangle \langle 0, \phi_\perp| + |0, \phi_\perp\rangle \langle 1, \phi_0| \text{Id}_a \otimes O_i$. Hence, the prediction variance for a Pauli string, $P = \bigotimes_{\ell=1}^L \sigma_\ell$ with $\sigma_\ell \in \{\text{Id}_\ell, X_\ell, Y_\ell, Z_\ell\}$, can be uniformly bounded through the Cauchy–Schwarz inequality, $\text{Tr}[\Gamma_P^2] \leq 2 \|P\|_2^2 \leq 2$. This immediately implies

$$O_i = \sum_{v=1}^{M_i} \kappa_{i,v} P_v \implies \mathbb{V}[\gamma_i] \leq \mathcal{O}(\|\vec{\kappa}_i\|_1^2), \quad (52)$$

for a general Hermitian operator $O_i = O_i^\dagger$. Observe that the variance remains rather insensitive to the operator locality or the system size. Accordingly, each data matrix element in the MODMD setting incurs an additive error of $\epsilon_1 = \epsilon_{\text{noise}}$ if we sample $Q = \mathcal{O}(\log(I) \max_{1 \leq i \leq I} \|\vec{\kappa}_i\|_1^2 \epsilon_{\text{noise}}^{-2})$ quantum measurements.

2. Trotter error

A second error source pertains to inexact implementation of the unitary evolution e^{-iHt} , which perturbs both the eigenfrequencies $e^{-iE_n t}$ and the time-evolved states $\rho(t)$. For near-term implementation, we assume query access to an approximate compilation of the evolution, for example, through Trotter–Suzuki factorization [68,69] or linear combination of unitaries (LCU) [70]. We consider specifically the Trotter scheme for illustration and comment that similar analysis should hold for other schemes. For $H = \sum_{v=1}^M H_v = \sum_{v=1}^M \kappa_{H,v} P_v$, a first-order Trotter formula gives

$$\epsilon_1 = \left\| e^{-iHt} - \left(\prod_{v=1}^M e^{-iH_v t/r} \right)^r \right\|_2 = \mathcal{O} \left(\frac{\|\vec{\kappa}_H\|_1^2 t^2}{r} \right), \quad (53)$$

where the Trotterized Hamiltonian simulation can be performed efficiently on the quantum computer if $M = \mathcal{O}(\text{polylog}(N))$. Equivalently, for

$$r = \mathcal{O}(\|\vec{\kappa}_H\|_1^2 t^2 \epsilon_1^{-1}) \quad (54)$$

Trotterized blocks with a time discretization $\Delta t^{\text{Trotter}} = t/r$, we incur a Hamiltonian simulation error $\mathcal{O}(\epsilon_1)$ in the operator norm. For a p th-order Trotter scheme, the error for simulating e^{-iHt} is $\mathcal{O}((\|\vec{\kappa}_H\|_1 t)^{p+1} r^{-p})$ [71]. We will work with this general p th-order setting, as tracking the resulting error contributions is straightforward.

For any maximal runtime $t_{\text{max}} = k_{\text{max}} \Delta t$, $e^{-iHt_{\text{max}}}$ can be Trotterized with depth $\mathcal{O}(M(\|\vec{\kappa}_H\|_1 t_{\text{max}})^{1+1/p} \epsilon_1^{-1/p})$, where we omit the $\mathcal{O}(5^{p/2-1})$ dependence due to Suzuki recursion, which builds a p th-order product formula from five copies of the $(p-2)$ th-order formula. This factor is

constant for a fixed p . As preserving the Hankel structure of data matrices $\mathbf{X}$ and $\mathbf{X}'$ is essential, we pick the same Trotter step to implement e^{-iHt_k} for $t_k \leq t_{\text{max}}$; changing the Trotter step size across time steps k would break the Hankel structure.

Let $\gamma_i(t) := \text{Re} \langle \phi_0 | O_i e^{-iHt} | \phi_0 \rangle$. To achieve a sampling uncertainty ϵ_{noise} when estimating the expectation values $\gamma_i(t_k)$ under Trotterized evolution up to t_{max} , the number of operator exponentials required is

$$\tilde{\mathcal{O}} \left(\frac{\log(I) k_{\text{max}} M(\|\vec{\kappa}_H\|_1 t_{\text{max}})^{1+1/p}}{\epsilon_1^{1/p} \epsilon_{\text{noise}}^2} \right), \quad (55)$$

because $\sum_{k=1}^{k_{\text{max}}} t_k / \Delta t^{\text{Trotter}} = \mathcal{O}(k_{\text{max}} r_{\text{max}})$ while at each time step k , we simulate e^{-iHt_k} with $\tilde{\mathcal{O}}(\log(I) \epsilon_{\text{noise}}^{-2})$ shots to collect the classical shadow dataset. To ensure that the solution to LS system in Eq. (13) is spectrally ϵ_1 -close to its Trotterized version, we take $\epsilon_1 \rightarrow \mathcal{O}(\epsilon_1 I^{-1/2} k_{\text{max}}^{-1})$ and $\epsilon_{\text{noise}} \rightarrow \mathcal{O}(\epsilon_1 I^{-1/2} k_{\text{max}}^{-1})$, using the fact that the spectral norm of the data matrix deviation is at most $\sqrt{dI(K+1)}$ times the entry-wise error. Assigning $\epsilon_1 = \mathcal{O}(\epsilon)$ to be the algorithmic error then yields an *upper bound* on the total gate complexity for processing a multiobservable signal of length k_{max} [72],

$$N_{\text{gate,tot}}^{\text{MODMD}} = \tilde{\mathcal{O}} \left(\frac{\chi \log(I) I^{1+1/2p} k_{\text{max}}^{4+2/p} M(\|\vec{\kappa}_H\|_1)^{1+1/p}}{\epsilon^{2+1/p}} \right), \quad (56)$$

with a factor $\chi = \|\mathbf{X}^+\|_2 + \|\mathbf{X}^+\|_2^2 \|\mathbf{X}'\|_2$ arising from the matrix perturbation bound for pseudoinverse (recall that the LS solution is formally derived via pseudoinverse) [73,74]. In particular, χ captures the influence of k_{max} , I , and the observables $\{O_i\}_{i=1}^I$ through the ideal data matrices.

This scaling invites a heuristic comparison with that of QPE [19,23,24,75], where the time evolution is likewise Trotterized with a p th-order scheme. Here we generalize the derivation in [3] that only considers the second-order scheme. The Trotterized QPE implements time steps $t_k^{\text{QPE,Trotter}} = 2^{k-1} \Delta t^{\text{QPE,Trotter}}$. Note that the total energy error resulting from a Trotter step of $\Delta t^{\text{QPE,Trotter}}$ is roughly [3]

$$\epsilon \Delta t^{\text{QPE,Trotter}} \approx \pi \sqrt{2^{-2k_{\text{max}}^{\text{QPE}}-2} + \epsilon_1^2}, \quad (57)$$

which reflects the interplay between phase estimation and Trotterization. Setting both error sources to $2^{-k_{\text{max}}^{\text{QPE}}-1} = \epsilon_1 = \epsilon \Delta t^{\text{QPE,Trotter}} / (2\sqrt{2}\pi)$ locks the final error at ϵ . The phase error imposes $k_{\text{max}}^{\text{QPE}} \leq \log_2(\epsilon t / (2\sqrt{2} \Delta t^{\text{QPE,Trotter}}))$

while the Trotterization error imposes

$$\epsilon_1 = (\|\vec{k}_H\|_1 \Delta t^{\text{QPE, Trotter}})^{p+1} = \frac{\epsilon \Delta t^{\text{QPE, Trotter}}}{2\sqrt{2}\pi}, \quad (58)$$

which yields $\Delta t^{\text{Trotter, QPE}} = \|\vec{k}_H\|_1^{-(p+1)/p} (\epsilon/(2\sqrt{2}\pi))^{1/p}$ and therefore

$$k_{\max}^{\text{QPE}} \leq \frac{p+1}{p} \log_2 \left(\frac{2\sqrt{2}\pi \|\vec{k}_H\|_1}{\epsilon} \right). \quad (59)$$

Accordingly, the total gate complexity for QPE is

$$N_{\text{gate, tot}}^{\text{QPE}} = \tilde{O} \left(\sum_{n \in \mathcal{N}_{\text{eig}}} \frac{1}{|\alpha_n|^2} \frac{M(2\sqrt{2}\pi \|\vec{k}_H\|_1)^{1+1/p}}{\epsilon^{1+1/p}} \right), \quad (60)$$

where we recall that $\mathcal{N}_{\text{eig}}$ is the set of target eigenindices (in our case, those labeling the low-lying eigenstates) and $|\alpha_n|^2 = |\langle \psi_n | \phi_0 \rangle|^2$. The summation over $\mathcal{N}_{\text{eig}}$ accounts for the circuit repetitions necessary to land on each target eigenstate.

In regimes where simulation of time evolution achieves sufficient fidelity, as anticipated in the early fault-tolerant era, MODMD remains robust against constant shot noise $\epsilon_{\text{noise}} = \mathcal{O}(1)$, assuming access to well-prepared reference state with dominant overlap on the target eigenstates. In particular, a recent theoretical analysis [25] elucidates the notable noisy super-resolution property: MODMD under uniform time steps, $t_k = k\Delta t$, can achieve an algorithmic precision ϵ by evolving up to $t_{\max} = \tilde{O}(\epsilon^{-2/3})$ simulation time, enabling a shallower circuit than the sharp $\Omega(\epsilon^{-1})$ lower bound for QPE. This improved scaling arises when the reference state is sufficiently supported on the target eigenstates. However, the MODMD total runtime, $t_{\text{tot}} = \tilde{O}(\epsilon^{-4/3})$, gets asymptotically worse than the Heisenberg limit. Alternatively, another analysis [17] under adaptive time step, $t_k = \prod_{0 \leq k' < k} m_{k'}$ with $m_0 = 1$ and $m_{k' > 1} \geq 2$ for $k' > 1$, indicates that MODMD can similarly saturate the Heisenberg limit in total runtime.

In near-term regimes where Trotterization overhead is explicit, we underscore that the associated circuit depth, determined by a p -th order scheme in use, is an especially practical consideration beyond the total gate complexity. For MODMD, this depth is $\tilde{O}(\|\vec{k}_H\|_1^{1+1/p} t_{\max}^{1+1/p} \epsilon^{-1/p}) = \tilde{O}(\epsilon^{-2/3-5/3p})$ for the maximal runtime $t_{\max} = \tilde{O}(\epsilon^{-2/3})$, i.e., when the initial state has dominant overlap with the set of target eigenstates. For QPE, the Trotterized depth scales as $\Omega(\|\vec{k}_H\|_1^{1+1/p} \epsilon^{-1-1/p}) = \Omega(\epsilon^{-1-1/p})$ by Eq. (60). This suggests that, as circuit depth still poses limitations on current devices, higher-order Trotterization for $p \geq 3$ can yield a smaller relative depth overhead for MODMD. This, albeit, comes at the expense of larger total runtime due to more circuit runs.

3. Depolarizing noise channel

For pre-fault-tolerant implementation, hardware noise need to be considered since device errors can significantly degrade simulation coherence and, thus, the performance of MODMD. Perhaps the simplest and most pervasively modeled device error is the global depolarization channel, $\mathcal{D}_t : \rho(0) \mapsto \tilde{\rho}(t)$. This phenomenological model serves a coarse approximation to decoherence effects and captures uniform, unstructured noise.

Under this channel, the ancilla-system composite state evolves toward a maximally mixed state,

$$\tilde{\rho}(t) = e^{-r_D |t|} \rho(t) + \frac{1 - e^{-r_D |t|}}{2N} \text{Id}_a \otimes \text{Id}, \quad (61)$$

where $r_D > 0$ parametrizes the rate of depolarization. We note that the exponential decay degrades the intensity of time-evolved signal where $\text{Tr}[\tilde{\rho}(t)\Gamma_O] = e^{-r_D |t|} \text{Tr}[\rho(t)\Gamma_O]$ for any operator O . Therefore, the exponential damping is encoded in our multiobservable signal as the real part of the calculated MODMD phases $\tilde{\lambda}_n$ (in contrast to the imaginary part encoding eigenenergy estimates). That is, $r_D \approx -\text{Re} \log(\tilde{\lambda}_n)/\Delta t$. It is worth stating that Fourier-based algorithms [11, 41], under such damping in the time domain, are prone to corresponding spectral broadening in the energy domain. This can obscure spectral features of interest unless additional deconvolution techniques are applied.

We further argue that MODMD also performs robustly under specific Pauli noise channels, given a few additional assumptions. We illustrate with the Pauli channel $\mathcal{D}_t^P : \rho(t) \mapsto \tilde{\rho}(t)$,

$$\tilde{\rho}(t) = g_P \rho(t) + (1 - g_P) P_{\text{noise}} \rho(t) P_{\text{noise}}, \quad (62)$$

where $g_P \in [0, 1]$ controls the weight of the ideal evolution relative to the Pauli error, and P_{noise} acts as a single fixed Pauli string for simplicity. We assume, in particular, that the ancilla qubit does not undergo any bit-flip errors. We factor P_{noise} into ancilla and system components: $P_{\text{noise}}^a \otimes P_{\text{noise}}^{\text{sys}}$. Then for any Pauli string O ,

$$\begin{aligned} \text{Tr}[P_{\text{noise}} \rho(t) P_{\text{noise}} \Gamma_O] \\ = \pm \text{Re} \langle \phi_{\perp} | P_{\text{noise}}^{\text{sys}} | \phi_{\perp} \rangle \langle \phi_0(t) | O P_{\text{noise}}^a | \phi_0 \rangle, \end{aligned} \quad (63)$$

where a sign ambiguity arises here from both the ancillary contribution of $\langle 1 | P_{\text{noise}}^a | 1 \rangle$ and system contribution due to the (anti)commutation relation between $P_{\text{noise}}^{\text{sys}}$ and O . Observe that if $|\phi_{\perp}\rangle$ is a computational basis state, P_{noise} acts nontrivially only if it is composed entirely of Z gates. Moreover, if $|\phi_0\rangle$ is a computational basis state, then

$$\tilde{\gamma}_i(t) = \text{Tr}[\tilde{\rho}(t)\Gamma_O] = [g_P \pm (1 - g_P)] \gamma_i(t), \quad (64)$$

where the RHS is nonvanishing for $g_P \neq \frac{1}{2}$. In practice, as knowledge about the locality of the system error $P_{\text{noise}}^{\text{sys}}$

gets learned, one can prepare a more interesting reference state, such as sparse superposition of computational basis states or low-bond-dimension MPS, that is unaffected by the error with high probability. Quantifying robustness under physically realistic Pauli noise models, for example, channels induced by continuous-time Lindblad dynamics, is an exciting direction beyond the scope of this work.

Finally, we consider a general channel $\mathcal{D}_t^\omega : \rho(0) \mapsto \tilde{\rho}(t)$ subsuming the depolarizing and Pauli noise,

$$\tilde{\rho}(t) = g_t \rho(t) + (1 - g_t) \omega(t), \quad (65)$$

where g_t again defines the survival probability of the ideal time-evolved state and $\omega(t)$ is a noisy background state. Notice that $\text{Tr}[\omega(t)\Gamma_O] \equiv 0$ for a general operator O when $\omega(t)$ is block-diagonal in the ancilla basis. The deviation between exact observables, e.g., $\gamma_i(t)$ from Sec. IV C 2, and their counterparts, $\tilde{\gamma}_i(t)$, depends on specific models of the survival probability g_t . However, a universal upper bound on this deviation can be derived using the matrix Hölder inequality [76],

$$|\gamma_i(t) - \tilde{\gamma}_i(t)| \leq (1 - g_t) \|\vec{\kappa}_i\|_1 \|\rho(t) - \omega(t)\|_{*1} \quad (66)$$

where $\|\cdot\|_{*1}$ denotes the Schatten 1-norm and we also rely on the fact that $\|\Gamma_{O_i}\|_2 \leq \|O_i|\phi_0\rangle\|_2 \leq \|\vec{\kappa}_i\|_1$.

V. APPLICATIONS

In this section, we detail numerical studies conducted on representative many-body systems from condensed-matter physics and quantum chemistry to demonstrate the efficacy of the MODMD framework. Our numerical calculations precisely follow prescriptions in Sec. III.

A. Condensed-matter physics

We examine the convergence of the ground- and excited-state eigenenergies of the 1D transverse-field Ising model (TFIM) with a total of $L = 15$ spins and open boundary conditions. The system Hamiltonian is given by

$$H_{\text{TFIM}} = -J \sum_{i=1}^{L-1} Z_i Z_{i+1} - h \sum_{i=1}^L X_i, \quad (67)$$

for coupling constant J and external field strength h .

We first demonstrate the performance of MODMD for TFIM parameters fixed at $J = h = 1$, examining how the algorithm behaves with varying number of time steps or equivalently the maximal simulation time. As discussed in Secs. II and IV, increasing the dimensions d and K of the data matrix $\mathbf{X} \in \mathbb{R}^{d \times (K+1)}$ facilitates convergence of the MODMD eigenvalues $\tilde{\lambda}_n = |\tilde{\lambda}_n| e^{-i\tilde{E}_n \Delta t}$ (note that these eigenvalues are not necessarily confined to the unit circle in complex plane) to the eigenphases $\lambda_n = e^{-iE_n \Delta t}$ of the Hamiltonian. Figure 2 illustrates this convergence for

the first $N_{\text{eig}} = 4$ eigenenergy estimates. Specifically, we report the absolute error $|\delta E_n| = |\tilde{E}_n - E_n|$ as a function of the dimension K , with the ratio $K/d = \frac{5}{2}$ held constant throughout subsequent calculations as this ratio provides near-optimal performance for ODMD [9].

Figure 2(a) shows convergence in the multiobservable setting where $I = 7$. We assign $O_1 = \text{Id}$ as per the convention given in Sec. III C, and sample $\{O_i\}_{i=2}^I$ randomly from the set of 1-local Pauli strings. The initial state $|\phi_0\rangle$ is an equal superposition of six computational basis states. The use of simple random Pauli observables proves effective here since (1) the TFIM does not preserve the Hamming weight (local bit-flip X does not annihilate $\langle \phi_0 | O_i | \phi_0(t) \rangle$) and (2) our reference is a superposition (so local phase-flip Z does not add a trivial ± 1 overall phase). The reference $|\phi_0\rangle$ contains relatively small but sufficient overlap with the first few Hamiltonian eigenstates, where the squared overlap sums to $\sum_{n=0}^{N_{\text{eig}}-1} |\langle \psi_n | \phi_0 \rangle|^2 \approx 10^{-1}$. This allows for the algorithm to generate an adequate signal without requiring substantial similarity between the reference state and target eigenstates. To stimulate the shadow-induced errors, we additionally introduce Gaussian $\mathcal{N}(0, \varepsilon_{\text{noise}}^2)$ noise with $\varepsilon_{\text{noise}} = 10^{-3}$ to the multiobservable signal. The absolute energy errors shown are averaged over a total of 20 realizations of both the Gaussian noise and 1-local Pauli observables. For comparison, the right panel displays the single-observable convergence from the standard ODMD algorithm as our benchmark. We observe considerably faster convergence in the excited-state energies in MODMD for cases where ODMD nearly stagnates. We highlight that the quantum cost, or total number of shots required, is at most $2(K + d) \log(I) \varepsilon_{\text{noise}}^{-2}$, only a factor of $2 \log(I) < 4$ more than that of ODMD in this case.

The convergence of MODMD naturally divides into two regimes, each defined by a distinct error scaling. The noise level $\varepsilon_{\text{noise}}$ essentially determines the crossover between the two convergence regimes. When $|\delta E_n| > \varepsilon_{\text{noise}}$, we observe an exponentially decaying error typical of the classical subspace methods [60, 62, 64]. Conversely, in the regime where $|\delta E_n| < \varepsilon_{\text{noise}}$, increasing simulation time leads to slower, algebraic error decay with precision ultimately limited by Heisenberg scaling. This crossover between the exponential and algebraic error decay is shown explicitly within Appendix F 2. In practice, the onset of the algebraic error behavior can be considerably delayed, for example, via increased and noise-mitigated sampling of classical shadows. Similar analysis for the Heisenberg model, serving as another condensed-matter system, can be found in Appendix G.

Despite its simplicity, the TFIM is an instructive toy model as it undergoes a quantum phase transition, where the spectral gap between the first two eigenstates can be systematically tuned by varying the h/J ratio. In the thermodynamic limit $L \rightarrow \infty$, the gap closes at $h/J = 1$ and increases monotonically with h/J . To investigate

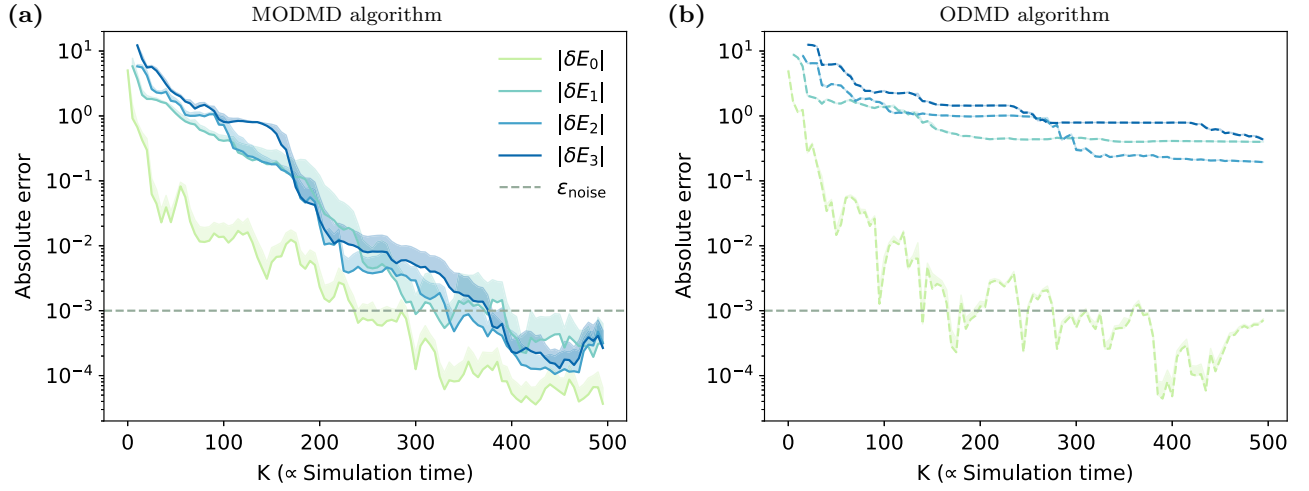


FIG. 2. Convergence of eigenenergies for the transverse-field Ising model. To obtain eigenenergy estimates $\tilde{E}_n$, we fix the multiobservable dynamic mode decomposition parameters $K/d = \frac{5}{2}$ and $\tilde{\delta} = 10^{-2}$ for constructing and thresholding the pair of data matrices $\mathbf{X}, \mathbf{X}' \in \mathbb{R}^{dl \times (K+1)}$. Gaussian $\mathcal{N}(0, \epsilon_{\text{noise}}^2)$ noise with $\epsilon_{\text{noise}} = 10^{-3}$ is added independently to the real or/and imaginary parts of the matrix elements. The absolute errors, $|\tilde{E}_n - E_n|$, in the first four eigenenergies of the transverse-field Ising model Hamiltonian are shown with respect to K proportional to the nondimensional maximal simulation time. The reference state $|\phi_0\rangle$ is an equal superposition of six computational basis states (see Appendix F 1) and we employ a time step of $\Delta t \approx 0.08$. Shading above the solid and dashed lines shows the standard deviation across trials for each quantity. For the transverse-field Ising model, the model Hamiltonian is of size $32\,768 \times 32\,768$ and the largest linear system in the corresponding multiobservable dynamic mode decomposition least-squares problem has size 1400×501 . (a) Absolute errors from the multiobservable (multiobservable dynamic mode decomposition) algorithm with $I = 7$ distinct observables, where the convergence results are averaged over 20 trials, each involving a Gaussian noise realization and a selection of $I - 1$ random observables. (b) Absolute errors from the single-observable (observable dynamic mode decomposition) algorithm where the convergence results are averaged over 20 trials, each involving a Gaussian noise realization.

the gap dependence near a phase transition, we demonstrate the difference in performance between the single-observable (ODMD) and multiobservable (MODMD) algorithms in the presence of near-degenerate target energies. In Fig. 3, we focus on comparing the error $|\delta E_1|$ in the first excited energy against the gap $E_1 - E_0$. The convergence results indicate that, in our multiobservable approach, the first excited-state energy E_1 can be accurately estimated when the noise level is slightly smaller than the spectral gap. In contrast, ODMD requires a visibly higher gap-to-noise ratio to achieve a comparable accuracy. This highlights a significant improvement of MODMD in distinguishing near-degenerate eigenstates.

B. Quantum chemistry

Electronic structure calculation is a fundamental problem in quantum chemistry. Here we evaluate performance of the MODMD algorithm for molecular Hamiltonians, whose second quantization can be efficiently implemented on the quantum computer. A minimal STO-3G basis set is used to construct the finite-dimensional Hamiltonian, which is then transformed into qubit form via the parity fermion-to-qubit mapping [77].

Moreover, we construct observables based on the Pauli representation of the Hamiltonian as per Eq. (23) and

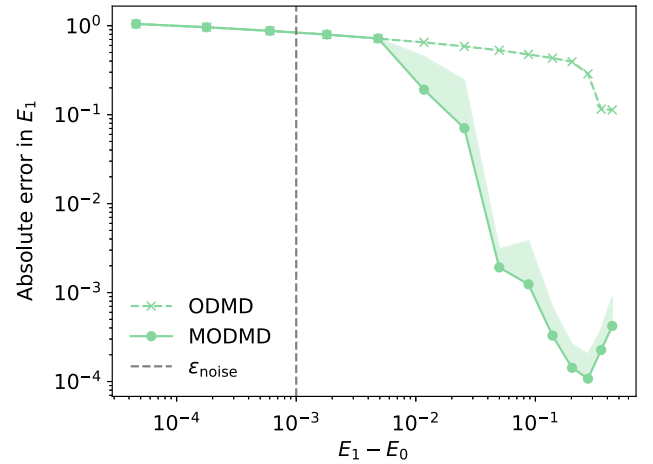


FIG. 3. Convergence of the first excited-state energy of the TFIM. Absolute error in the first excited-state energy is shown as a function of the spectral gap between the ground and the first excited state for fixed $K = 500$. The vertical dotted line marks the noise level $\epsilon_{\text{noise}} = 10^{-3}$. Convergence results are averaged over 20 trials each involving a Gaussian noise realization and, in the multiobservable dynamic mode decomposition case, also a selection of $I - 1$ random observables. The shading above the solid and dashed lines shows the standard deviation across trials for each quantity.

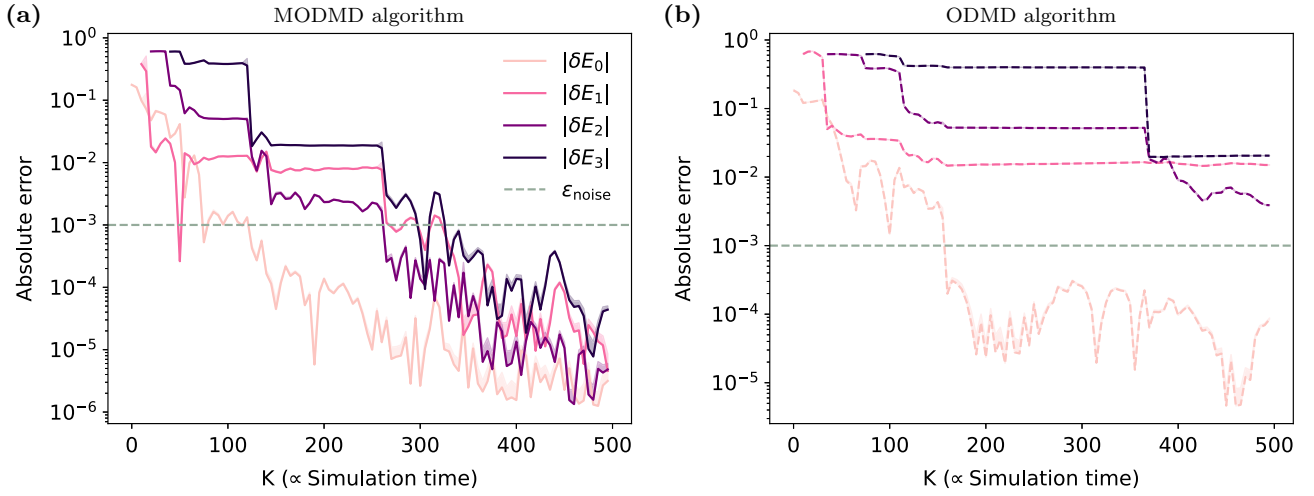


FIG. 4. Convergence of eigenenergies for the lithium hydride molecule. To obtain eigenenergy estimates $\tilde{E}_n$, we fix the multiobservable dynamic mode decomposition parameters $K/d = \frac{5}{2}$ and $\tilde{\delta} = 10^{-2}$ for constructing and thresholding the pair of data matrices $\mathbf{X}, \mathbf{X}' \in \mathbb{R}^{dl \times (K+1)}$. Gaussian $\mathcal{N}(0, \varepsilon_{\text{noise}}^2)$ noise with $\varepsilon_{\text{noise}} = 10^{-3}$ is added independently to the real or/and imaginary parts of the matrix elements. The reference state $|\phi_0\rangle$ is a superposition of six Slater determinants (see Appendix F 1) and we employ a time step of $\Delta t \approx 0.39$. The absolute errors, $|\tilde{E}_n - E_n|$, in the first four eigenenergies of the lithium hydride Hamiltonian are shown with respect to K proportional to the nondimensional maximal simulation time. Convergence results are averaged over 20 trials of Gaussian noise realizations with shading above the solid and dashed lines showing the standard deviation across trials for each quantity. For lithium hydride, the model Hamiltonian is of size 1024×1024 and the largest linear system in the corresponding multiobservable dynamic mode decomposition least-squares problem has size 1400×501 . (a) Absolute errors from the multiobservable (multiobservable dynamic mode decomposition) algorithm with $I = 7$ distinct observables. (b) Absolute errors from the single-observable (observable dynamic mode decomposition) algorithm.

deterministically select a subset of $I = 7$ Pauli operators P_v (where $O_1 = \text{Id}$) with medium-magnitude weights κ_v to maximize spectral independence among observables. The reference state $|\phi_0\rangle$ is prepared to be a superposition of six Slater determinants. These effective single-particle states map simply to computational basis states derived from a mean-field calculation. This choice ensures both a nonzero overlap with the target eigenstates and viability for experimental preparation.

We illustrate MODMD for the lithium hydride (LiH) molecule, where exact eigenenergies can be calculated by full diagonalization of the electronic Hamiltonian. Figure 4 shows the convergence of the first $N_{\text{eig}} = 4$ eigenenergy estimates for LiH at its equilibrium bond length of 1.59 Å. Similar to the TFIM results from Fig. 2, Fig. 4 exhibits a characteristic exponential-to-algebraic crossover in the error decay, with the transition occurring near noise level $\varepsilon_{\text{noise}}$ (see Appendix F 2). The plateaus observed for the higher energy levels at intermediate values of K can be attributed to the anomalous convergence to different but higher-lying eigenenergies, which we verified numerically. As we approach the interior of the Hamiltonian spectrum, MODMD can resolve the higher excited-state energies for sufficiently large K , whereas ODMD stagnates. Results of ODMD and MODMD on additional chemical systems, BeH_2 and N_2 , can be found in Appendix G.

In Fig. 5 we assess the noise robustness of the MODMD algorithm in our LiH example. Adopting the same hyperparameters as used in Fig. 4, we now fix

$K = 500$ and vary the noise level $\varepsilon_{\text{noise}}$. Importantly, we maintain the SVD threshold $\tilde{\delta} = 10\varepsilon_{\text{noise}}$ at a consistently larger value. We observe a power-law scaling of the absolute error with respect to the noise level, suggesting that a conservative truncation strategy can be used to protect the actual signal from noise across a wide range of noise magnitudes. By comparison, the ODMD algorithm yields larger errors for any given $\varepsilon_{\text{noise}}$. Furthermore, reducing noise only significantly improves the ODMD performance when $\varepsilon_{\text{noise}}$ reaches relatively small values. In other words, ODMD requires a more aggressive truncation $\tilde{\delta}$, which may cause a serious loss of the signal information and thereby slower convergence.

C. Comparison with alternative approaches

1. QMEGS

To evaluate the practical performance of MODMD, we compare its convergence with alternative state-of-the-art real-time methods. We first benchmark against QMEGS [11] as a leading signal processing algorithm. QMEGS estimates multiple Hamiltonian eigenenergies by reconstructing a Gaussian-filtered signal,

$$\mathcal{G}(E, O) = \left| \frac{1}{K} \sum_{k=0}^{K-1} \hat{s}_k e^{iEt_k} \right|, \quad (68)$$

$$\hat{s}_k \in \{\pm 1 \pm i\}, \quad \mathbb{E}(\hat{s}_k) = \langle \phi_0 | O e^{-iHt_k} | \phi_0 \rangle,$$

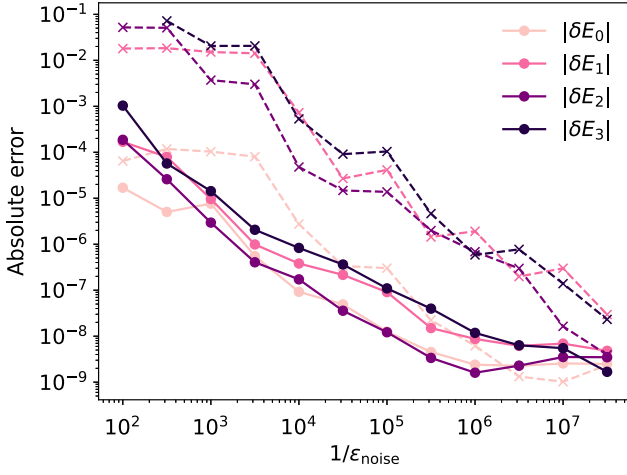


FIG. 5. Convergence of eigenenergies for lithium hydride under varying noise level. Absolute errors in the four lowest eigenenergies of the lithium hydride Hamiltonian are plotted against the noise level ϵ_{noise} for fixed $K = 500$ and $\tilde{\delta} = 10\epsilon_{\text{noise}}$. All convergence results reflect an average over 20 trials of Gaussian noise realizations. The solid lines correspond to the multiobservable (multiobservable dynamic mode decomposition) algorithm with $I = 7$ distinct observables and the dashed lines to the single-observable (observable dynamic mode decomposition) algorithm.

where the times $\{t_k\}_k$ are sampled i.i.d. from a truncated Gaussian distribution. Peaks of $\mathcal{G}(E, O)$ correspond to eigenvalues of H , reflecting a Gaussian-broadened version of the underlying eigenstate overlap $|\langle\psi_n|\phi_0\rangle|^2$. QMEGS retrieves the dominant energies upon locating the peaks using a classical search-and-block optimization strategy, with quantum data collected from single-shot Hadamard-test measurements for the observable O . Here we take the observable to be unitary (e.g., $O = \text{Id}$ or Pauli strings), consistent with the Hadamard-test setting.

Notably, the QMEGS algorithm differs from MODMD in its theoretical guarantee, which depends on the choice of a suitable reference state, although the algorithm itself can be initialized with arbitrary reference state. QMEGS reaches Heisenberg-limited scaling, under the assumption that the reference state satisfies the *sufficiently dominant condition*, namely that there is a set of eigenindices $\mathcal{N}_{\text{eig}}$ so that

$$\min_{n \in \mathcal{N}_{\text{eig}}} |\langle\psi_n|\phi_0\rangle|^2 > \sum_{n \notin \mathcal{N}_{\text{eig}}} |\langle\psi_n|\phi_0\rangle|^2. \quad (69)$$

Additionally, while MODMD can resolve the lowest-lying eigenenergies, QMEGS most naturally identifies energies with the largest eigenstate overlaps. To facilitate a direct comparison between the methods, we use an initial state $|\phi_0\rangle$ for which the first $N_{\text{eig}} = 4$ eigenstates are weighted preferentially but Eq. (69) is not fulfilled. Such an initial state is constructed in Appendix F 1.

Specifically, we perform similar simulations to those in Sec. V A, analyzing the error in $|\mathcal{N}_{\text{eig}}|$ eigenenergies of the TFIM Hamiltonian as a function of simulation time. To maintain consistency, the hyperparameters of the two algorithms are adjusted so that the total number of shots remains the same for each method. The size of QMEGS dataset K_{QMEGS} , given MODMD parameters K_{MODMD} , d_{MODMD} , I , and shot noise level ϵ_{noise} , is calibrated as $K_{\text{QMEGS}} = (K_{\text{MODMD}} + d_{\text{MODMD}}) \log(I) \epsilon_{\text{noise}}^{-2}$. For any maximal MODMD evolution time $t_{\text{max}} = (K + d)\Delta t$, we choose the Gaussian width and truncation in QMEGS to match t_{max} . As in Sec. V A, the observables O_i in each MODMD trial consist of $O_1 = \text{Id}$ and six additional candidates $O_{i \geq 2}$ drawn from the 1-local Pauli strings. We account for the fact that QMEGS uses a single observable by running the algorithm to reconstruct $\mathcal{G}(E, O_i)$ for each O_i in the MODMD observable set. Focusing on low-lying eigenenergies and noting that additional observables can only improve or leave unchanged their visibility, we take the minimum error $|\delta E_n^{\text{QMEGS}}| = \min_{1 \leq i \leq I} |\delta E_n^{\text{QMEGS}}(O_i)|$ across observables. For demonstration, we focus on a 12-qubit system with a simulated noise level of $\epsilon_{\text{noise}} = 10^{-5/2}$ and MODMD threshold $\tilde{\delta} = 10^{-3/2}$.

The relative enhancement of MODMD over QMEGS, $|\delta E_n^{\text{QMEGS}} / \delta E_n^{\text{MODMD}}|$, as a function of the nondimensional maximal simulation time is shown in Fig. 6. Beyond small simulation time, MODMD consistently outperforms QMEGS in the convergence of all four lowest-lying eigenenergies. The QMEGS hyperparameters α_{QMEGS} and q_{QMEGS} [11] are tuned to optimize performance, with details provided in Appendix H 2.

2. VQPE and UVQPE

As a further comparison, we examine the performance of MODMD relative to VQPE [3,4], a leading real-time subspace method. We benchmark both approaches using the TFIM Hamiltonian under a similar setting described in Sec. V A. Recall our previous initial state $|\phi_0\rangle$ was chosen so that $\sum_{n=0}^3 |\langle\psi_n|\phi_0\rangle|^2 \approx 10^{-1}$, thereby situating the system well beyond the *sufficiently dominant* regime, where signal processing methods are effective. To pursue our comparative analysis under general instance, here we consider a less structured (yet classically tractable) initial state $|\phi_0\rangle$ specified in Appendix F 1. This state does not preferentially overlap with low-lying eigenstates, allowing us to more rigorously assess the performance of MODMD and VQPE in the absence of spectral bias.

VQPE solves the subspace eigenvalue problem,

$$\mathbf{H}\Psi = \tilde{E}\mathbf{S}\Psi, \quad (70)$$

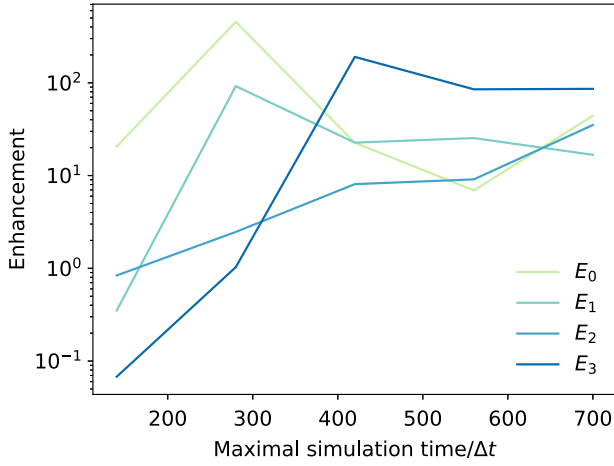


FIG. 6. Performance comparison of multiobservable dynamic mode decomposition and quantum multiple eigenvalue Gaussian filtered search on the transverse-field Ising model. Relative performance enhancement, $|\delta E_n^{\text{QMEGS}}/\delta E_n^{\text{MODMD}}|$, is computed from absolute errors in the first four eigenenergies (obtained by quantum multiple eigenvalue Gaussian filtered search and multiobservable dynamic mode decomposition with $I = 7$ distinct observables) and is shown with respect to K_{MODMD} proportional to the nondimensional maximal simulation time. Multiobservable dynamic mode decomposition and quantum multiple eigenvalue Gaussian filtered search errors are averaged over five trials: each employing a common observable set with $O_1 = \text{Id}$ and $I - 1$ observables randomly selected from a pool of Pauli strings. We search for 20 peaks in quantum multiple eigenvalue Gaussian filtered search, while the four lowest-lying eigenstates have the 1st, 7th, 10th, and 5th highest initial state overlap respectively. Note that the absolute errors for both algorithms do not reach the same order of magnitude as ϵ_{noise} until the larger values of maximal simulation time in this plot.

where the Hamiltonian and overlap matrix elements,

$$\mathbf{H}_{jk} = \langle \phi_j | H | \phi_k \rangle, \quad \mathbf{S}_{jk} = \langle \phi_j | \phi_k \rangle, \quad (71)$$

are obtained by projecting the full many-body eigenvalue problem onto the subspace spanned by the nonorthogonal real-time basis $\{|\phi_k\rangle = e^{-iHk\Delta t}|\phi_0\rangle\}_k$. Notice that these matrix elements correspond to two types of observables, $O = H$ and $O = \text{Id}$. Their real and imaginary parts are measured separately in VQPE using the Hadamard test, which requires sampling at least four distinct circuits per time step (even assuming access to block encoding of the problem Hamiltonian H). Importantly, VQPE remains effective despite imperfections in time evolution [78], such as those introduced by Trotterization, since the algorithm ultimately projects the exact Hamiltonian onto subspace spanned by some prepared states—regardless of whether they are exactly or approximately time-evolved.

Unitary VQPE (UVQPE), as a direct algorithmic twin, solves an isospectral subspace eigenvalue problem, $\mathbf{U}\Psi = e^{-i\tilde{E}\Delta t}\mathbf{S}\Psi$, where $\mathbf{U}_{jk} = \langle \phi_j | e^{-iH\Delta t} | \phi_k \rangle \equiv \mathbf{U}_{jk+1}$ define the

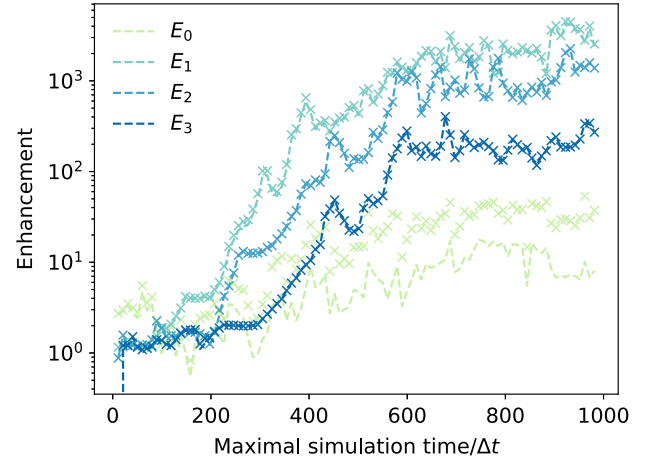


FIG. 7. Performance comparison of multiobservable dynamic mode decomposition and (U)VQPE on the transverse-field Ising model. Relative performance enhancement, $|\delta E_n^{(\text{U})\text{VQPE}}/\delta E_n^{\text{MODMD}}|$, is computed from absolute errors in the first four eigenenergies (obtained by unitary variational quantum phase estimation and multiobservable dynamic mode decomposition with $I = 7$ distinct observables) and is shown with respect to K proportional to the nondimensional maximal simulation time. Multiobservable dynamic mode decomposition errors are averaged over 20 trials, each involving a Gaussian noise realization and a set of $I - 1$ observables randomly selected from a pool of candidate operators. Cross markers represent comparison with unitary variational quantum phase estimation while dashed lines represent comparison with variational quantum phase estimation.

matrix elements of projected time-evolution operator. UVQPE mitigates resource overhead by circumventing the need to measure the Hamiltonian matrix elements; on the other hand, like MODMD, it still requires sufficiently accurate time evolution to extract spectral information. We view UVQPE as an effective variant of VQPE, bridging and encompassing state-of-the-art subspace approaches.

In Fig. 7, we compare (U)VQPE and MODMD for the first $N_{\text{eig}} = 4$ eigenenergy estimates. In particular, we show the relative enhancement, $|\delta E_n^{(\text{U})\text{VQPE}}/\delta E_n^{\text{MODMD}}|$, as a function of the maximal simulation time. It is worth noting that a $d \times (K + 1)$ linear system in MODMD corresponds to a generalized eigenvalue problem of size $(d + K + 1) \times (d + K + 1)$ in (U)VQPE.

When the simulation is stopped at $K = 500$, MODMD consistently achieves lower energy errors than (U)VQPE, often by one to three orders of magnitude. The advantage of MODMD becomes evident as the simulation extends, where (U)VQPE struggles to resolve eigenstates reliably. Leveraging a richer observable signal with essentially no increase in the measurement overhead per time step [note that $\log(I) \sim 4$ for $I \sim 10^2$, matching a minimal factor required by VQPE], MODMD delivers accurate energy estimates with significantly reduced quantum costs.

D. Prediction of longer-time dynamics

To demonstrate the utility of the MODMD framework beyond eigenenergy estimation, we present additional numerical simulations, where the system matrix A

is constructed from a set of observable measurements and then used to forecast future dynamics.

Following Eqs. (12) and (13), we measure the multi-observable expectations $\vec{s}_k = \vec{s}(k\Delta t)$ for $k = 0, 1, \dots, k^*$ and arrange these real-time snapshots to construct $\mathbf{X}$ and

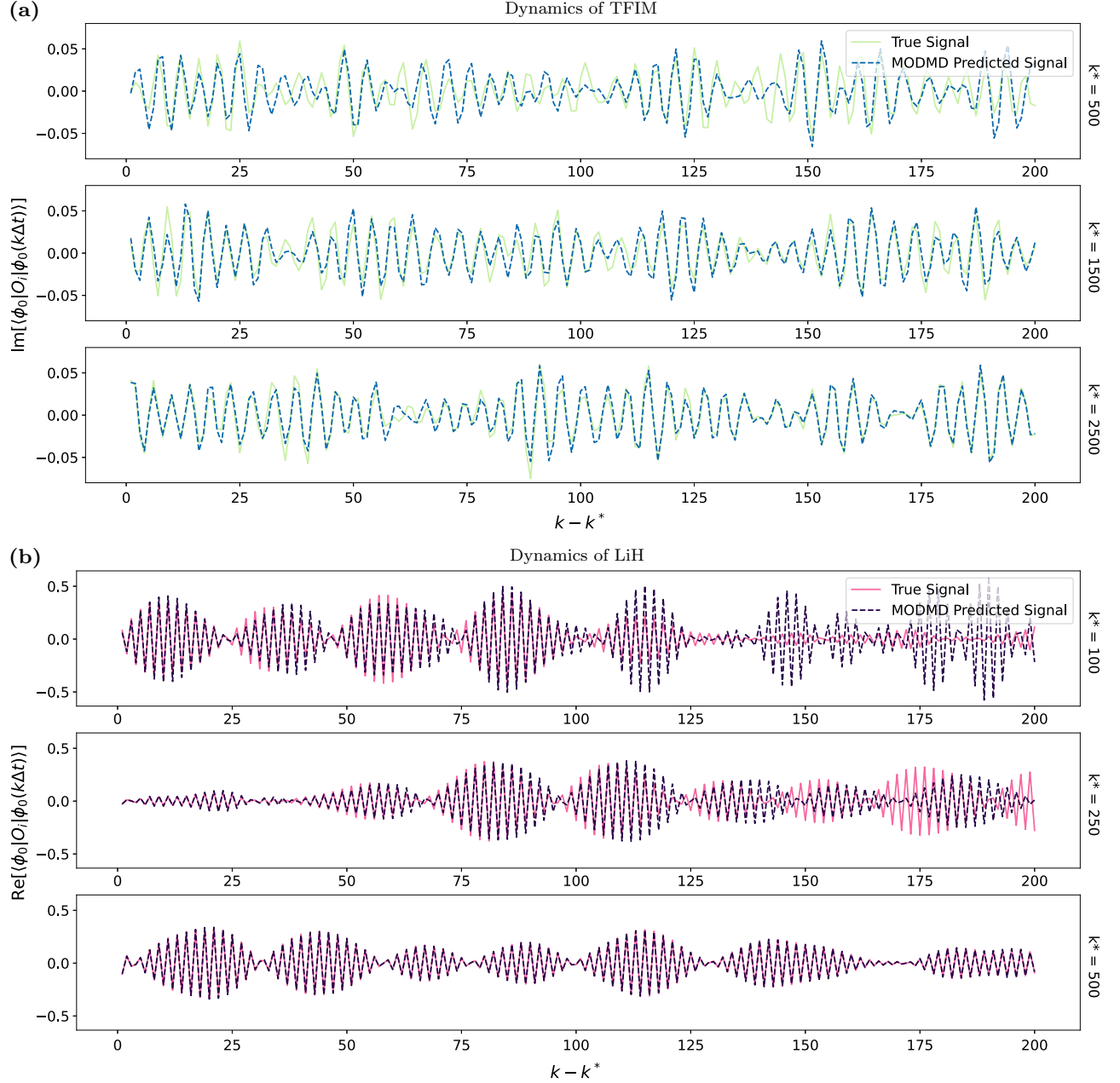


FIG. 8. Predicted observable dynamics from multiobservable dynamic mode decomposition. For different values of k^* (as specified on the right side of each panel), we construct the system matrix A from the multiobservable snapshots $\vec{s}_k = \vec{s}(k\Delta t)$ for $k = 0, 1, \dots, k^*$. We then predict $\vec{s}_k$ for $k = k^* + 1, k^* + 2, \dots, k^* + 200$ by computing $\vec{s}_k = A^{k-k^*} \mathbf{X}'$ and plot the i th element of predicted $\vec{s}_k$, i.e., $\text{Re}[\langle \phi_0 | O_i | \phi_0(k\Delta t) \rangle]$. The figure displays the results for $i = 5$, though similar convergence behavior is observed across all observables. The multiobservable dynamic mode decomposition hyperparameters $K/d = \frac{5}{2}$ and $\delta = 10^{-2}$ are fixed for constructing and thresholding the pair of data matrices $\mathbf{X}, \mathbf{X}' \in \mathbb{R}^{dI \times (K+1)}$. Gaussian $\mathcal{N}(0, \varepsilon_{\text{noise}}^2)$ noise with $\varepsilon_{\text{noise}} = 10^{-3}$ is added independently to the matrix elements. The reference states, time steps, and observable sets are the same as those used in Sec. V. The results are averaged over 20 trials, each involving a Gaussian noise realization. (a) Dynamics predictions for the transverse-field Ising model Hamiltonian. (b) Dynamics predictions for the lithium hydride Hamiltonian.

$\mathbf{X}'$. Here, k^* is a hyperparameter controlling how many snapshots are collected—and therefore the extent of past information exploited for prediction. After constructing the system matrix A , we propagate the dynamics through the MODMD prescription $\vec{s}_{k+1} \approx A\vec{s}_k$, as is outlined in Sec. III D. This allows us to predict $\vec{s}_k$ for future steps $k = k^* + 1, k^* + 2, \dots$ by computing $A^{k-k^*}\mathbf{X}'$.

In Fig. 8, we visualize real part of the i th observable component, i.e., $\text{Re}\langle\phi_0|O_i|\phi_0(k\Delta t)\rangle$, up to 200 time steps beyond k^* . For transverse-field Ising and LiH iltonians, we observe that the predictive accuracies improve with k^* in both cases. This aligns with our basic intuition that longer observation windows grant ability to predict dynamics further into the future.

Though we only predict the dynamics for the particular sets of observables used in Sec. V, one could perform extrapolation of other quantities of interest depending on the system and dynamics in question. In support of such broader applicability, we also investigate the approximate eigenstates extracted by MODMD in Appendix I, where our method exhibits reliable performance—suggesting its competence in capturing a wide range of low-energy properties.

VI. CONCLUSION

In this work, we developed a hybrid quantum-classical measurement-driven framework for effectively extracting information about the low-energy eigenspaces of quantum many-body systems. Our MODMD approach leverages real-time evolution on quantum hardware and classically unravels multidimensional signals, composed of real-time observables, from a limited number of randomized measurements. The simultaneous prediction of many observables leads to accurate estimates of eigenenergies and shallower circuits with shorter evolution time. We explored the theoretical underpinnings of MODMD, which exponentially suppresses spectral error in the low-noise regime. We numerically demonstrated its rapid convergence in the presence of perturbative noise using examples from condensed matter physics and quantum chemistry.

Compared to state-of-the-art real-time approaches, we highlight the unique strengths of our method in addition to its reliable convergence and noise resilience. To our best knowledge, MODMD is among the most resource-efficient for generating real-time signals. This is because (1) we evolve a single reference state for a duration shorter than required by single-observable approaches, where the reduction in the simulation time becomes more substantial as the number of observables included in the signal subspace increases, and (2) the reference state does not have to possess large overlaps with the low energy eigenstates of interest. Furthermore, our classical postprocessing consists of solving a simple LS problem followed by a standard eigenvalue problem, which is ansatz-free and thus

circumvents an exponential growth in optimization costs, whether quantum or classical, associated with the number of desired eigenenergies.

In fact, our MODMD framework is capable of retrieving ground- and excited-state properties beyond eigenenergy levels, demonstrating an extensive and timely application of the low-rank shadow. This versatility suggests that similar ideas may extend to dynamical properties as well. Nonequilibrium behavior is particularly intriguing in this context, and exploring such directions can further extend the scope of our multiobservable framework.

ACKNOWLEDGMENTS

This work was funded by the U.S. Department of Energy (DOE) under Contract No. DE-AC02-05CH11231, through the Office of Science, Office of Advanced Scientific Computing Research (ASCR) Exploratory Research for Extreme-Scale Science (Y.S., K.K., D.C., and R.V.B.). This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231. H.Y.H. would like to acknowledge the support from the Harvard Quantum Initiative. S.F.Y. and H.Y.H. would like to acknowledge funding from the DOE through the QUACQ center (No. DE-SC0025572).

DATA AVAILABILITY

The data that support the findings of this article are openly available [79].

Appendix A: Shadow under locally scrambling dynamics

Here we adopt a dynamical perspective on the efficient generation of the classical shadow, which introduces random scrambling via real-time evolution under a disordered local Hamiltonian $H_s(t)$. For concreteness, we focus on quantum spin glass, in particular the disordered TFIM with Hamiltonian,

$$H_s(\tau) = - \sum_{\langle ij \rangle} J_{ij}(\tau) X_i X_j - h(\tau) \sum_{i=1}^L Z_i, \quad (\text{A1})$$

where J_{ij} and h set the spin-spin coupling and external field strength respectively. We assume for simplicity that $J_{ij} = \sum_k J_{ij,k} \mathbb{1}_{[\tau_k, \tau_{k+1})}(\tau)$ and $h = \sum_k h_k \mathbb{1}_{[\tau_k, \tau_{k+1})}(\tau)$ remain piecewise constant in time, with $\mathbb{1}_{[\tau_k, \tau_{k+1})}$ being an indicator function of the k th-step interval. To incorporate scrambling disorder, we employ quenched random interactions, e.g., $J_{ij,k} \sim \mathcal{N}(0, \sigma_J^2)$ and $h_k \sim \mathcal{N}(0, \sigma_h^2)$. Thus under the time evolution $U_s = \prod_{k=K_s}^0 e^{-iH_s(\tau_k)(\tau_{k+1}-\tau_k)}$,

local quantum information gets scrambled and the resulting entanglement produced by the disorder dynamics facilitates the recovery of $\rho_\Phi(t)$. We remark that the number of scrambling time steps K_s serves the role analogous to the circuit depth when considering Clifford gates [80]. For any integer $T \geq 0$, $\lim_{K_s \rightarrow \infty} \mathcal{U}(\{\tau_k\}_{0 \leq k \leq K_s})$ is a T -design, whereby its first T statistical moments align with those of the Haar measure. To simulate long-time dynamics, we draw upon our recent tool of algebraic circuit compression [81,82], particularly suited to a disordered TFIM Hamiltonian $H_s(\tau)$ in Eq. (A1), to keep the effective depth of the time evolution U_s shallow. Remarkably, the depth post compression should exhibit no dependence on the maximal runtime τ_{K_s} , therefore allowing a significantly more efficient exploration of the Haar limit.

Appendix B: Observable dynamic mode decomposition

The standard DMD, originally developed in the field of numerical fluid dynamics, is a measurement-driven approximation for the temporal progression of a classical dynamical system [83–87]. Specifically, DMD samples the system snapshots at regular time intervals Δt and uses them to construct an efficient representation of the full dynamical trajectory. For simplicity, we consider a system whose N -dimensional state manifold is $\mathbb{C}^N$. The optimal linear approximation for the discretized time step $k \mapsto k+1$ is expressed as the LS relation,

$$\phi_{k+1} \stackrel{\text{LS}}{=} A_k \phi_k, \quad (\text{B1})$$

where $\phi_k \in \mathbb{C}^N$ specifies the system state at time $k\Delta t$ and $A_k \in \mathbb{C}^{N \times N}$ is the system matrix, i.e., the linear operator that minimizes the residual $\|\phi_{k+1} - A_k \phi_k\|_2$ to yield the LS relation above. Similarly, the optimal linear approximation for a sequence of successive snapshots $k = 0, 1, \dots, K+1$ can be determined by the solution,

$$\begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_{K+1} \end{bmatrix} \stackrel{\text{LS}}{=} \begin{bmatrix} \phi_0 & \phi_1 & \cdots & \phi_K \end{bmatrix} A, \quad (\text{B2})$$

where the system matrix A minimizes the sum of squared residuals over the length- K sequence. The linear flow described by Eq. (B2) naturally generates approximate dynamics $\phi_{\text{DMD}}(t) = A^{t/\Delta t} \phi_0$ governed by eigenmodes of A . DMD-based approaches can be remarkably effective despite their formal simplicity, since they are rooted in the general Koopman operator theory developed to describe the behavior of general (non)linear dynamical systems [53–55,57,58].

The standard DMD approach described above for classical dynamics cannot be immediately translated to quantum

dynamics. The DMD approximation of the system evolution would require complete knowledge of the system state, as specified by an N -dimensional complex vector at each time step. However, we do not have direct access to the full many-body quantum state. Instead, we can only access the state of a quantum system via measurement sampling of observables [88]. To address this challenge, we employ a technique motivated by Takens' embedding theorem [89–91] to obtain an effective state vector consisting of an operator measured at a sequence of successive times. We reformulate the linear model underpinning DMD in terms of these observable vectors to approximate the system dynamics.

Takens' embedding theorem [89,90] establishes a connection between the manifold of states, which an observer cannot directly access, and time-delayed measurements of an observable. In particular, the theorem asserts, under generous conditions, that a state on an N -dimensional (sub)manifold can be completely determined using a sequence of at most $d_* \leq 2N + 1$ time-delayed observables. This correspondence reads

$$\phi(t) \leftrightarrow \mathbf{o}_{t,d_*} = \begin{bmatrix} o(t) \\ o(t + \Delta\tau) \\ \vdots \\ o(t + (d_* - 1)\Delta\tau) \end{bmatrix}, \quad (\text{B3})$$

where $\Delta\tau > 0$ is the time delay, $\phi(t)$ is the system state, $o(t) = o[\phi(t)]$ is the measured observable, and $\mathbf{o}_t$ is the d_* -dimensional “observable trajectory” containing the dynamical information. The RHS of Eq. (B3) is known as a d_* -dimensional delayed embedding of the observable. Takens' theorem relates the evolution of microscopic degrees of freedom to the evolution history of macroscopic observables, providing a concrete probe into the dynamical properties of the system without direct access to the full states. Here we adopt the term Takens' embedding technique to refer to the method of applying time delays on the system observables, motivated by the rigorous results of Takens' embedding theorem.

In anticipation of efficiently leveraging near-term quantum resources, we choose the time delay in Takens' embedding technique to equal the DMD time interval, i.e., $\Delta\tau = \Delta t$. Given this choice, we then measure the system along time steps $\{t_k = k\Delta t\}_{k=0}^{K+1}$ and acquire the sequence of observable trajectories $\{\mathbf{o}_{t_k,d}\}_{k=0}^{K+1}$, each of some length $d \leq d_*$,

$$\mathbf{o}_{t_k,d} = \begin{bmatrix} o(t_k) \\ o(t_{k+1}) \\ \vdots \\ o(t_{k+d-1}) \end{bmatrix}, \quad 0 \leq k \leq K+1. \quad (\text{B4})$$

By construction, the first $(d-1)$ entries of $\mathbf{o}_{t_k,d}$ are identical to the last $(d-1)$ entries of $\mathbf{o}_{t_{k-1},d}$. Consequently, the

matrix assembled by arranging successive trajectories $\mathbf{o}_{t_k}$ as columns

$$\mathbf{X}_{k_1:k_2} = [\mathbf{o}_{t_{k_1},d} \quad \mathbf{o}_{t_{k_1+1},d} \quad \cdots \quad \mathbf{o}_{t_{k_2},d}] \quad (\text{B5})$$

has a Hankel form, i.e., the matrix elements on each anti-diagonal are equal. In the embedding space, we can identify the closest linear flow,

$$\mathbf{X}_{1:K+1} \stackrel{\text{LS}}{=} \mathbf{X}_{0:K} \implies A = \mathbf{X}_{1:K+1} (\mathbf{X}_{0:K})^+, \quad (\text{B6})$$

where $+$ denotes the Moore-Penrose pseudo-inverse. Here the system matrix A assumes a companion structure with just d free parameters. The approximation to the system dynamics is then stored in the d parameters inferred from measurements of $K + d + 1$ delayed observables. We hence name our LS embedding in the observable space the *observable dynamic mode decomposition* (ODMD).

Appendix C: Multiobservable dynamic mode decomposition

Recall that the multiobservable system matrix has a characteristic polynomial,

$$\mathcal{C}_A(z) = \det(z - A) = \det\left(\sum_{\ell=0}^d z^\ell A_\ell\right), \quad (\text{C1})$$

where $A_d \equiv \text{Id}_I$. We examine the simpler version of Eq. (13) in which we assume the d matrix blocks,

$$A_\ell = \text{Diag} \begin{bmatrix} A_{\ell,11} & & & \\ & A_{\ell,22} & & \\ & & \ddots & \\ & & & A_{\ell,II} \end{bmatrix}, \quad (\text{C2})$$

to be diagonal with $A_{\ell,ij} \equiv 0$ for $i \neq j$. In this case, the multivariate residual is

$$\vec{\text{Res}}_k(A) = \sum_{\ell=0}^d A_\ell \vec{s}_{\ell+k} = \sum_{\ell=0}^d A_\ell \sum_{n=0}^{N-1} \vec{c}_n \lambda_n^{\ell+k} = \sum_{n=0}^{N-1} \lambda_n^k \sum_{\ell=0}^d \lambda_n^\ell A_\ell \vec{c}_n = \sum_{n=0}^{N-1} \lambda_n^k \sum_{i=1}^I c_{n,i} \left(\sum_{\ell=0}^d \lambda_n^\ell A_{\ell,ii} \right) \vec{e}_i \quad (\text{C3})$$

$$= \sum_{n=0}^{N-1} \lambda_n^k \sum_{i=1}^I c_{n,i} \mathcal{C}_{A|i}(\lambda_n) \vec{e}_i = \sum_{i=1}^I \left(\sum_{n=0}^{N-1} c_{n,i} \lambda_n^k \mathcal{C}_{A|i}(\lambda_n) \right) \vec{e}_i, \quad (\text{C4})$$

where Eq. (C3) directly follows from the submatrices A_ℓ being diagonal and $\vec{e}_i$ denotes the canonical basis vector of $\mathbb{C}^I$. Moreover, we define in Eq. (C4) the I single-observable characteristic polynomials,

$$\mathcal{C}_{A|i}(z) = \sum_{\ell=0}^d z^\ell A_{\ell,ii} = \prod_{\ell=0}^{d-1} (z - \tilde{\lambda}_{\ell|i}), \quad (\text{C5})$$

each with d roots $\tilde{\lambda}_{\ell|i}(A_{0,ii}, A_{1,ii}, \dots, A_{d-1,ii})$. Clearly, Eq. (C4) indicates that the multivariate residual admits I decoupled components corresponding to the different observables. Minimizing the total residual, $\|\vec{\text{Res}}_k(A)\|_2$, is hence equivalent to minimizing each of the I observable residuals, $\left| \sum_{n=0}^{N-1} c_{n,i} \lambda_n^k \mathcal{C}_{A|i}(\lambda_n) \right|$. This is consistent with the fact that the matrix determinant $\det(z - A)$ factorizes into independent contributions,

$$\mathcal{C}_A(z) = \prod_{i=1}^I \left(\sum_{\ell=0}^d z^\ell A_{\ell,ii} \right) = \prod_{i=1}^I \mathcal{C}_{A|i}(z), \quad (\text{C6})$$

such that the eigenvalues of A also factorize into clusters based on single-observable residuals. The resulting MODMD estimates are bounded by the I single-observable ODMD estimates, e.g.,

$$\min_{1 \leq i \leq I} |\tilde{E}_{i,0} - E_0| \leq |\tilde{E}_0 - E_0| \leq \max_{1 \leq i \leq I} |\tilde{E}_{i,0} - E_0|, \quad (\text{C7})$$

where $\tilde{E}_0$ and $\tilde{E}_{i,0}$ designate, respectively, the ground-state energy estimate using the entire observable pool (the full system matrix A) and a single observable (one row of the system matrix $A_{\ell,ii}$).

More interesting convergence arises when the I observable residuals are coupled and give a total residual,

$$\vec{\text{Res}}_k(A) = \sum_{n=0}^{N-1} \lambda_n^k \sum_{i,j=1}^I c_{n,j} \left(\sum_{\ell=0}^d \lambda_n^\ell A_{\ell,ij} \right) \vec{e}_i, \quad (\text{C8})$$

$$= \sum_{i=1}^I \left(\sum_{n=0}^{N-1} c_{n,i} \lambda_n^k \mathcal{C}_{A|i}(\lambda_n) + \sum_{n=0}^{N-1} \lambda_n^k \sum_{\ell=0}^{d-1} \lambda_n^\ell \sum_{j \neq i}^I A_{\ell,ij} c_{n,j} \right) \vec{e}_i, \quad (\text{C9})$$

where we can lower the residual by utilizing the flexibility of off-diagonal elements in the submatrices A_ℓ . For example, the total residual may vanish completely when the multiobservable signal is $d(I-1)$ -sparse in the eigenfrequency basis, where, for $1 \leq i \leq I$, the coefficients $c_{n,i}$ are supported on at most $d(I-1)$ eigenindices $\mathcal{N}_i = \{0 \leq n \leq N-1 : c_{n,i} \neq 0\}$. That is, $\max_{1 \leq i \leq I} |\mathcal{N}_i| \leq d(I-1)$. In this case, a vanishing residual is possible, provided that (a) $|\cup_{i=1}^I \mathcal{N}_i| \leq d(I-1)$ and (b) the matrix elements $\{A_{\ell,ij}\}_{\ell,j \neq i}$ can be set appropriately such that $\forall 1 \leq i \leq I$,

$$\begin{bmatrix} \beta_{0,i1}(1) & \dots & \beta_{\ell,ij \neq i}(1) & \dots & \beta_{d-1,iI}(1) \\ \beta_{0,i1}(2) & \dots & \beta_{\ell,ij \neq i}(2) & \dots & \beta_{d-1,iI}(2) \\ \vdots & & \ddots & & \vdots \\ \beta_{0,i1}(|\mathcal{N}_i|) & \dots & \beta_{\ell,ij \neq i}(|\mathcal{N}_i|) & \dots & \beta_{d-1,iI}(|\mathcal{N}_i|) \\ \beta_{0,i1}(1) & \dots & \beta_{\ell,1j \neq i}(1) & \dots & \beta_{d-1,iI}(1) \\ \vdots & & \ddots & & \vdots \\ \beta_{0,i1}(|\mathcal{N}_I|) & \dots & \beta_{\ell,Ij \neq i}(|\mathcal{N}_I|) & \dots & \beta_{d-1,iI}(|\mathcal{N}_I|) \end{bmatrix} \begin{bmatrix} A_{0,i1} \\ A_{0,i2} \\ \vdots \\ A_{\ell,ij} \\ \vdots \\ A_{d-1,iI} \end{bmatrix} = - \begin{bmatrix} \beta_{0,ii}(1) \mathcal{C}_{A|i}(\lambda_{n_1[i]}) \\ \beta_{0,ii}(2) \mathcal{C}_{A|i}(\lambda_{n_2[i]}) \\ \vdots \\ \beta_{0,ii}(|\mathcal{N}_i|) \mathcal{C}_{A|i}(\lambda_{n_{|\mathcal{N}_i|}[i]}) \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (\text{C10})$$

where $n_\xi[i]$ numerates support eigenindices in the set $\mathcal{N}_i$ for observable O_i and $\beta_{\ell,ij}(\xi) = \lambda_{n_\xi[i]}^\ell c_{n_\xi[i],j}$. Given $\{A_{\ell,ii}\}_\ell$, a formal solution $\{A_{\ell,ij}\}_{\ell,j \neq i}$ to Eq. (C10) exists if the matrix on the LHS has a full rank. This directly implies that the total residual may vanish completely for a dI -sparse multiobservable signal in the eigenfrequency basis (as expected), because $\{A_{\ell,ii}\}_\ell$ can always be chosen so that the single-observable characteristic polynomial $\mathcal{C}_{A|i}(z)$ encompasses the roots $\{\lambda_{n_\xi[i]}\}_\xi$ corresponding to at most d members of $\mathcal{N}_i$. Without additional degrees of freedom from the off-diagonal matrix elements of A , the residual only vanishes for a d -sparse signal.

We now relax the sparsity assumption on the signal to explore the general case. Observe that

$$\min_A \sum_{k=0}^K \left\| \vec{\text{Res}}_k(A) \right\|_2^2 = \min_A \sum_{k=0}^K \sum_{i=1}^I \left| \sum_{n=0}^{N-1} c_{n,i} \lambda_n^k \mathcal{C}_{A|i}(\lambda_n) + \sum_{j \neq i} \sum_{n=0}^{N-1} c_{n,j} \lambda_n^k \mathcal{C}_{A|ij}(\lambda_n) \right|^2, \quad (\text{C11})$$

$$\leq \min_{\{A_{\ell,ii}: 0 \leq \ell \leq d-1, 1 \leq i \leq I\}} \sum_{k=0}^K \sum_{i=1}^I \left| \sum_{n \notin \mathcal{N}_i^*} c_{n,i} \lambda_n^k \mathcal{C}_{A|i}(\lambda_n) + \sum_{j \neq i} \sum_{n \notin \mathcal{N}_i^*} c_{n,j} \lambda_n^k \bar{\mathcal{C}}_{A|ij}(\lambda_n) \right|^2, \quad (\text{C12})$$

where $\mathcal{C}_{A|ij}(z) = \sum_{\ell=0}^d z^\ell A_{\ell,ij} \equiv \sum_{\ell=0}^{d-1} z^\ell A_{\ell,ij}$ for $j \neq i$ defines a degree- $(d-1)$ polynomial via the off-diagonal matrix elements $\{A_{\ell,ij}\}_\ell$, $\mathcal{N}_i^* \subset [N]$ is an eigenindex subset of size $|\mathcal{N}_i^*| = d(I-1)$, and $\bar{\mathcal{C}}_{A|ij}(z) = \sum_{\ell=0}^{d-1} z^\ell A_{\ell,ij}(\mathcal{N}_i^*; \{A_{\ell,ii}\}_\ell)$ fixes $\mathcal{C}_{A|ij}(z)$ by evaluating specific coefficients $\{A_{\ell,ij}\}_\ell$ such that

$$\sum_{j \neq i} c_{n,j} \bar{\mathcal{C}}_{A|ij}(\lambda_n) = -c_{n,i} \mathcal{C}_{A|i}(\lambda_n), \quad \forall n \in \mathcal{N}_i^*. \quad (\text{C13})$$

Resembling Eq. (C10), we reserve the notation $n_\xi^*[i]$ for the support eigenindices in $\mathcal{N}_i^*$. Equation (C13) then holds only if the following coefficient matrix has full rank, i.e.,

$$\text{rank}[\beta_{\mathcal{N}_i^*}] = \text{rank} \begin{bmatrix} \beta_{0,i1}^*(1) & \dots & \beta_{\ell,ij \neq i}^*(1) & \dots & \beta_{d-1,iI}^*(1) \\ \beta_{0,i1}^*(2) & \dots & \beta_{\ell,ij \neq i}^*(2) & \dots & \beta_{d-1,iI}^*(2) \\ \vdots & & \ddots & & \vdots \\ \beta_{0,i1}^*(|\mathcal{N}_i^*|) & \dots & \beta_{\ell,ij \neq i}^*(|\mathcal{N}_i^*|) & \dots & \beta_{d-1,iI}^*(|\mathcal{N}_i^*|) \end{bmatrix} = d(I-1), \quad (\text{C14})$$

for $\beta_{\ell,ij}^*(\xi) = \lambda_{n_{\xi[i]}^*}^{\ell} c_{n_{\xi[i]}^*}^{\ell} c_{n_{\xi[i]}^*}^{\ell}$. An arbitrary index set $\mathcal{N}_i^*$ can be assigned as long as the full-rank condition is met; ideally, we aim for $\mathcal{N}_i^* = \mathcal{N}_i^*(\{A_{\ell,ii}\}_{\ell})$ to satisfy the optimal property,

$$\left| \sum_{j=1}^I \sum_{n \notin \mathcal{N}_i^*} c_{n,j} \lambda_n^k \bar{C}_{A|ij}(\lambda_n) \right| = \min_{\mathcal{N} \subset [N]: |\mathcal{N}|=d(I-1)} \left| \sum_{j=1}^I \sum_{n \notin \mathcal{N}} c_{n,j} \lambda_n^k \bar{C}_{A|ij}(\lambda_n) \right|, \quad (\text{C15})$$

which minimizes the effective residual parametrized by the d diagonal matrix elements $\{A_{\ell,ii}\}_{\ell}$, with $\bar{C}_{A|ii}(z) = C_{A|i}(z)$. By the triangle inequality, we can bound the RHS of Eq. (C12) as

$$\left| \sum_{n \notin \mathcal{N}_i^*} c_{n,i} \lambda_n^k C_{A|i}(\lambda_n) + \sum_{j \neq i} \sum_{n \notin \mathcal{N}_i^*} c_{n,j} \lambda_n^k \bar{C}_{A|ij}(\lambda_n) \right| \leq \sum_{j=1}^I \sum_{n \notin \mathcal{N}_i^*} |c_{n,j} \bar{C}_{A|ij}(\lambda_n)| \leq \bar{c}_i \sum_{j=1}^I \bar{C}_{ij}, \quad (\text{C16})$$

where $\bar{c}_i = \max_{1 \leq j \leq I} \sum_{n \notin \mathcal{N}_i^*} |c_{n,j}|$ and $\bar{C}_{ij} = \max_{n \notin \mathcal{N}_i^*} |\bar{C}_{A|ij}(\lambda_n)|$. With suitable choices of the reference state $|\phi_0\rangle$ and operator pool $\{O_i\}_{i=1}^I$, we can strategically adjust the coefficients $\{c_{n,i}\}_{n,i}$, therefore controlling the constant $\bar{c}_i$ and the conditioning of the matrix $\beta_{\mathcal{N}_i^*}$ (or equivalently $\|\beta_{\mathcal{N}_i^*}^{-1}\|_2$). Accordingly, for given index sets $\{\mathcal{N}_i^*\}_i$, we have

$$\begin{aligned} & \min_{\{A_{\ell,ii}: 0 \leq \ell \leq d-1, 1 \leq i \leq I\}} \sum_{k=0}^K \sum_{i=1}^I \left(\bar{c}_i \sum_{j=1}^I \bar{C}_{ij} \right)^2 \\ & \leq (K+1) \sum_{i=1}^I \bar{c}_i^2 \min_{\{A_{\ell,ii}: 0 \leq \ell \leq d-1\}} \left(\max_{n \notin \mathcal{N}_i^*} |C_{A|i}(\lambda_n)| + \sum_{\ell=0}^{d-1} \sum_{j \neq i} |A_{\ell,ij}(\mathcal{N}_i^*; \{A_{\ell,ii}\}_{\ell})| \right)^2, \end{aligned} \quad (\text{C17})$$

where the residual from the off-diagonal matrix elements can be bounded by

$$\sum_{\ell=0}^{d-1} \sum_{j \neq i} |A_{\ell,ij}(\mathcal{N}_i^*; \{A_{\ell,ii}\}_{\ell})| \leq d(I-1) \max_{n \in \mathcal{N}_i^*} |c_{n,i}| \|\beta_{\mathcal{N}_i^*}^{-1}\|_2 \max_{n \in \mathcal{N}_i^*} |C_{A|i}(\lambda_n)|, \quad (\text{C18})$$

using an identity analogous to Eq. (C10). To further bound the RHS of Eq. (C17), we attempt to tightly approximate the minimizer over the d diagonal matrix elements $\{A_{\ell,ii}\}_{\ell}$ or alternatively over the set of degree- d monic polynomials defined along the circular arc $\mathcal{A}_{\theta} = \{z \in \mathbb{C} : |z| = 1, -\theta \leq \arg(z) \leq \theta\}$. In particular let us consider the complex-valued Chebyshev polynomials,

$$\hat{C}(z) = \arg \min_{C \text{ monic: deg}(C)=d} \sup_{z \in \mathcal{A}_{\theta}} |C(z)|, \quad (\text{C19})$$

whose parametric representations are explicitly constructible by tools such as Jacobi's elliptic and theta functions [92]. For notational convenience, we denote $\|\hat{C}\|_{\theta} = \sup_{z \in \mathcal{A}_{\theta}} |\hat{C}(z)|$. Similar to the real-valued Chebyshev polynomials over the interval $[-1, 1]$, $\hat{C}(z)$ retains the minimal-norm property on $\mathcal{A}_{\theta}$, where $\|\hat{C}\|_{\theta} \sim 2 \sin^d(\theta/2) \cos^2(\theta/4)$ asymptotically for large d [93]. Setting $\theta_H(\Delta t) = (E_{N-1} - E_0) \Delta t / 2$, the total residual decays at least exponentially,

$$\min_A \sum_{k=0}^K \left\| \vec{\text{Res}}_k(A) \right\|_2^2 \leq (K+1) \|\hat{C}\|_{\theta_H(\Delta t)}^2 \sum_{i=1}^I \bar{c}_i^2 \left[1 + d(I-1) \max_{n \in \mathcal{N}_i^*} |c_{n,i}| \|\beta_{\mathcal{N}_i^*}^{-1}\|_2 \right]^2, \quad (\text{C20})$$

where the prefactor improves significantly, compared to the single-observable case derived in [9], when

$$\max_{1 \leq j \leq I} \sum_{n \notin \mathcal{N}_i^*} |c_{n,j}| \ll \frac{1}{1 + d(I-1) \max_{n \in \mathcal{N}_i^*} |c_{n,i}| \|\beta_{\mathcal{N}_i^*}^{-1}\|_2}, \quad \forall 1 \leq i \leq I. \quad (\text{C21})$$

Appendix D: Proof of theorems

1. Theorem D1

Theorem D1. Let $\tilde{E}_0(d)$ be the approximate ground-state energy extracted from the d -dimensional function subspace spanned by $\{g_1, \mathcal{K}_{\Delta t}[g_1], \dots, \mathcal{K}_{\Delta t}^{d-1}[g_1]\}$. For $d \geq 1$, there exists time step Δt such that the error $\delta E_g(d) = \tilde{E}_0(d) - E_0$ is bounded by

$$\delta E_0(d) \leq \frac{\text{cond}(\mathcal{B})^{1/2} |\sin[(E_{N-1} - E_0)\Delta t]|}{\tilde{\epsilon}_{0 \rightarrow 1}^{2(d-1)} \Delta t} \tan^2 \Xi, \quad (\text{D1})$$

where $\text{cond}(\mathcal{B}) = \|\mathcal{B}^{-1}\| \|\mathcal{B}\|$ denotes the condition number of the $N \times N$ eigenfunction Gram matrix $\mathcal{B}$ with entries $\mathcal{B}_{mn} = \langle f_m, f_n \rangle_\mu / (\|f_m\|_{L^2(\mu)} \|f_n\|_{L^2(\mu)})$, $\cos^2 \Xi = |\langle f_0, g_1 \rangle_\mu|^2 / (\|f_0\|_{L^2(\mu)}^2 \|g_1\|_{L^2(\mu)}^2)$ is the squared overlap between the reference function g_1 and the true ground-state $\mathcal{K}_{\Delta t}$ eigenfunction, while $\tilde{\epsilon}_{0 \rightarrow 1} = 1 + 3(E_1 - E_0)\Delta t/2\pi \in [1, 2]$ characterizes the normalized spectral gap of the Hamiltonian H .

Proof. Let us define

$$\theta(f) = \arg \frac{\langle f, \mathcal{K}_{\Delta t} f \rangle_\mu}{\langle f, f \rangle_\mu} \in [-\pi, \pi], \quad (\text{D2})$$

which returns the expected dynamical phase factor associated with a function f under Koopman evolution, where f belongs to the invariance subspace $\text{span}\{f_n\}_{n=0}^{N-1}$. Given a suitable symmetrizing spectral shift such that $\|H\| \Delta t \leq \pi$, we note that $\theta(f_0) = \max_f \theta(f) = -E_0 \Delta t$. This variational principle implies

$$\tilde{E}_0(d) = -\frac{1}{\Delta t} \max_{f \in \text{Kr}_d} \theta(f) = -\frac{1}{\Delta t} \max_{p \in \mathcal{P}_{d-1}} \theta(p(\mathcal{K}_{\Delta t})g_1), \quad (\text{D3})$$

for which we use Kr_d to denote our d -dimensional function Krylov subspace $\text{span}\{g_1, \dots, \mathcal{K}_{\Delta t}^{d-1}[g_1]\}$ and $\mathcal{P}_{d-1}$ the set of degree- $(d-1)$ polynomials over $\mathbb{C}$.

Hereafter, we assume that the eigenfunctions $\{f_n\}_{n=0}^{N-1}$ are orthogonal, i.e., Eq. (44) simplifies to a diagonal relation where $\mathcal{B}_{mn} = 0$ if $m \neq n$. Drawing on the convergence analysis for nonnormal operators in [66], we establish that the error bound in the cases of nonorthogonal eigenfunctions is amplified by at most a multiplicative factor $\text{cond}(\mathcal{B})^{1/2} \geq 1$. This is because $\mathcal{B}$ admits the representation $\mathcal{B} = \mathcal{T}^\dagger \mathcal{T}$ where $\mathcal{T}$ collects the eigenfunctions written in standard function basis, so that $\text{cond}(\mathcal{B}) = \text{cond}(\mathcal{T})^2$. In the orthogonal case, we can expand g_1 in the eigenfunctions,

$$\frac{g_1}{\|g_1\|_{L^2(\mu)}} = \sum_{n=0}^{N-1} \frac{z_n f_n}{\|f_n\|_{L^2(\mu)}}, \quad (\text{D4})$$

with $\sum_{n=0}^{N-1} |z_n|^2 = 1$. Thus Eq. (D3) in the eigenbasis reads

$$\tilde{E}_0(d) = -\frac{1}{\Delta t} \max_{p \in \mathcal{P}_{d-1}} \arg \frac{\langle p(\mathcal{K}_{\Delta t})g_1, \mathcal{K}_{\Delta t} p(\mathcal{K}_{\Delta t})g_1 \rangle_\mu}{\langle p(\mathcal{K}_{\Delta t})g_1, p(\mathcal{K}_{\Delta t})g_1 \rangle_\mu}, \quad (\text{D5})$$

$$= -\frac{1}{\Delta t} \max_{p \in \mathcal{P}_{d-1}} \arg \frac{\sum_{n=0}^{N-1} \lambda_n |z_n p(\lambda_n)|^2}{\sum_{n=0}^{N-1} |z_n p(\lambda_n)|^2}, \quad (\text{D6})$$

where $\lambda_n = e^{-iE_n \Delta t}$ label the eigenvalues of the Koopman operator. We note that

$$\max_{p \in \mathcal{P}_{d-1}} \arg \frac{\sum_{n=0}^{N-1} \lambda_n |z_n p(\lambda_n)|^2}{\sum_{n=0}^{N-1} |z_n p(\lambda_n)|^2} \geq \max_{p \in \mathcal{P}_{d-1}} \arg \frac{\lambda_0 |z_0 p(\lambda_0)|^2 + \lambda_{N-1} \sum_{n=1}^{N-1} |z_n p(\lambda_n)|^2}{\sum_{n=0}^{N-1} |z_n p(\lambda_n)|^2}, \quad (\text{D7})$$

by convexity of the circular sector with angle $[-E_{N-1} \Delta t, -E_0 \Delta t]$ in the complex plane. Then

$$\tilde{E}_0(d) \leq -\frac{1}{\Delta t} \max_{p \in \mathcal{P}_{d-1}} \arg \left[\lambda_0 + (\lambda_{N-1} - \lambda_0) \frac{\sum_{n=1}^{N-1} |z_n p(\lambda_n)|^2}{\sum_{n=0}^{N-1} |z_n p(\lambda_n)|^2} \right], \quad (\text{D8})$$

$$= -\frac{1}{\Delta t} \arg \left[\lambda_0 + (\lambda_{N-1} - \lambda_0) \min_{p \in \mathcal{P}_d} \frac{\sum_{n=1}^{N-1} |z_n p(\lambda_n)|^2}{\sum_{n=0}^{N-1} |z_n p(\lambda_n)|^2} \right], \quad (\text{D9})$$

where we observe $\arg(\lambda_{N-1} - \lambda_0) = \frac{1}{2}[\arg \lambda_{N-1} + \arg(-\lambda_0)] \leq \arg \lambda_0$ assuming $E_0 \leq 0$ WLOG. To proceed, we seek a family of polynomials defined over the unit circle $\mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\}$ to bound the fraction on the RHS of Eq. (D9). Now let us fix some $q \in (0, 1]$ and consider the handy choice of complex-valued Rogers-Szegő polynomials [94,95],

$$W_d(z | q) = \begin{cases} 1 & d = 0, \\ z + 1 & d = 1, \\ (1+z)W_{d-1}(z | q) & \\ -(1-q^{j-1})zW_{d-2}(z | q) & d \geq 2, \end{cases} \quad (\text{D10})$$

over the circle $\mathbb{S}^{1,q} = \{z : |z| = q^{-1/2}\}$. For simplicity, we rewrite $z = -q^{-1/2} \exp(-i\vartheta)$ where $\vartheta \in [-\pi, \pi]$ denotes an angular phase. Here a prefactor of -1 is included to periodically translate the polynomials so that $W_d(\vartheta | q)$ adapts the symmetry $W_d(-\vartheta) = W_d(\vartheta)^*$ (we also omit a conditional dependence of W_d on q for notational clarity). Such family of polynomials shares the key properties that (i) $|W_d(\vartheta)|$ remains bounded below unity over some angular window $\mathcal{W} = [-\Omega, \Omega] \subset [-\pi, \pi]$ and (ii) $|W_d(\vartheta)|$ grows rapidly outside $\mathcal{W}$.

Note that the constant q controls the width of our truncated angular window $\mathcal{W}$. In the limit of $q \rightarrow 1$, one can verify that these polynomials converge to

$$W_d(\vartheta) \rightarrow \sum_{k=0}^d \binom{d}{k} \exp[-ik(\vartheta + \pi)], \quad (\text{D11})$$

which simply gives the sum of evenly spaced points along $\mathbb{S}^1$ weighted by binomial coefficients. As a consequence, $\sup_{\vartheta} |W_d(\vartheta)| \approx 2^d$ for $q \approx 1$. To bound the fraction in Eq. (D9) tightly, we look for a suitable linear transformation $\mathcal{L} : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ acting on the N individual eigenphases, $\lambda_n = e^{-i\vartheta_n(\Delta t)}$, such that $\mathcal{L}$ nudges excited-state angles $\vartheta_{n \geq 1}$ all inside the truncated window $\mathcal{W}$ while keeping the ground-state angle ϑ_0 outside. It is safe to assume $\vartheta_{N-1} - \vartheta_1 \leq 2\Omega$ and $2\Omega \leq \vartheta_{N-1} - \vartheta_0 \leq \pi + \Omega$ by adapting a suitable time step size Δt , e.g.,

$$\Delta t = \sup_{\tau} \left\{ \tau \in \mathbb{R}^+ : \vartheta_{N-1}(\tau) - \vartheta_1(\tau) \leq 2\Omega, \vartheta_1(\tau) - \vartheta_0(\tau) \leq \Omega^c \right\}, \quad (\text{D12})$$

with $\Omega^c = \pi - \Omega$. Therefore a natural choice of $\mathcal{L}$ is the phase multiplication or angular translation, $\mathcal{L} : \vartheta \mapsto \vartheta - \Omega + \vartheta_1$, which circularly moves $\{\lambda_n\}_{n=0}^{N-1}$ so that $\mathcal{L}(\vartheta_0) \leq -\Omega = \mathcal{L}(\vartheta_1) \leq -\mathcal{L}(\vartheta_{N-1})$ as desired. With $\mathcal{L}$ chosen above, we establish a variational upper bound on the ground-state energy error by substituting the trial polynomials $p = W_{d-1} \circ \mathcal{L}$ into Eq. (D9),

$$\tilde{E}_0(d) \leq -\frac{1}{\Delta t} \arg \left[\lambda_0 + (\lambda_{N-1} - \lambda_0) \frac{\sum_{n=1}^{Q-1} |z_n W_{d-1}(\mathcal{L}(\lambda_n))|^2}{\sum_{n=0}^{N-1} |z_n W_{d-1}(\mathcal{L}(\lambda_n))|^2} \right], \quad (\text{D13})$$

$$\leq -\frac{1}{\Delta t} \arg \left[\lambda_0 + (\lambda_{N-1} - \lambda_0) \frac{\sum_{n=1}^{Q-1} |z_n W_{d-1}(\Omega)|^2}{|z_0 W_{d-1}(\mathcal{L}(\lambda_0))|^2} \right], \quad (\text{D14})$$

$$= -\frac{1}{\Delta t} \arg \left[\lambda_0 + (\lambda_{N-1} - \lambda_0) \frac{\sin^2 \Xi}{|W_{d-1}(\mathcal{L}(\lambda_0))|^2 \cos^2 \Xi} \right], \quad (\text{D15})$$

where in arriving at Eq. (D15) we have utilized the property (i) of $W_{d-1}(\vartheta)$ and defined the angle Ξ by $\cos^2 \Xi = |z_0|^2 = |\langle f_0, g_1 \rangle_{\mu}|^2 / (\|f_0\|_{L^2(\mu)}^2 \|g_1\|_{L^2(\mu)}^2)$ which specifies the L^2 -projection of our reference function g_1 onto the ground-state eigenfunction. For the limiting case $q = 1$, it is rather straightforward to show that $\Omega = \pi/3$ and

$$|W_d(\mathcal{L}(\lambda_0))|^{1/d} = \sqrt{2 - 2 \cos(\vartheta_1 - \vartheta_0 + \Omega)} \geq 1 + C\epsilon_{0 \rightarrow 1}, \quad (\text{D16})$$

where $\epsilon_{0 \rightarrow 1} = (\vartheta_1 - \vartheta_0)/\Omega^c = (E_1 - E_0)\Delta t/\Omega^c$ gives the normalized spectral gap times the time step and C is a constant for which Eq. (D16) holds with $\epsilon_{0 \rightarrow 1} \in [0, 1]$. For example, $C = 1$ can be justified by concavity of the LHS of the inequality above with respect to $(\vartheta_1 - \vartheta_0)$. Hence

we can further bound Eq. (D15) using Eq. (D16),

$$\tilde{E}_0(d) \leq -\frac{1}{\Delta t} \arg \left[\lambda_0 \left(1 - \tilde{\epsilon}_{0 \rightarrow 1}^{-2(d-1)} \tan^2 \Xi \right) + \lambda_{N-1} \tilde{\epsilon}_{0 \rightarrow 1}^{-2(d-1)} \tan^2 \Xi \right], \quad (\text{D17})$$

$$\leq \frac{1}{\Delta t} \arctan \left[\frac{(1 - \zeta) \sin \vartheta_0 + \zeta \sin \vartheta_{N-1}}{(1 - \zeta) \cos \vartheta_0 + \zeta \cos \vartheta_{N-1}} \right], \quad (\text{D18})$$

$$= E_0 + \frac{\sin(\vartheta_{N-1} - \vartheta_0)}{\Delta t} \zeta + \mathcal{O}(\zeta^2), \quad (\text{D19})$$

where we have defined $\tilde{\epsilon}_{0 \rightarrow 1} = 1 + C\epsilon_{0 \rightarrow 1} > 1$ and $\zeta = \tilde{\epsilon}_{0 \rightarrow 1}^{-2(d-1)} \tan^2 \Xi$. The last equality can be derived from a Taylor expansion up to leading order in ζ . Accounting

for the eigenfunction conditioning $\text{cond}^{1/2}(\mathcal{B})$, we have thereby proved the statement as claimed. ■

2. Aside: canonical angles

To set up the proof of Theorem D2, we first introduce the notion of subspace overlap for subsequent discussions [67]. Suppose that we have two subspaces $\mathcal{X}$ and $\mathcal{Y}$ of $\mathbb{C}^N$ with $\dim \mathcal{X} \leq \dim \mathcal{Y}$. Let $B_{\mathcal{X}} \in \mathbb{C}^{N \times \dim \mathcal{X}}$ and $B_{\mathcal{Y}} \in \mathbb{C}^{N \times \dim \mathcal{Y}}$ be the orthonormal basis matrices of $\mathcal{X}$ and $\mathcal{Y}$ ($B_{\mathcal{X}}$ and $B_{\mathcal{Y}}$ are thus not unique). The canonical angles between the subspaces are defined as

$$\theta_\ell[\mathcal{X}, \mathcal{Y}] = \arccos \sigma_\ell \in \left[0, \frac{\pi}{2}\right], \quad 1 \leq \ell \leq \dim \mathcal{X}, \quad (\text{D20})$$

where $\sigma_1 \leq \sigma_2 \leq \dots \leq \sigma_{\dim \mathcal{X}}$ denote the singular values of $B_{\mathcal{X}}^\dagger B_{\mathcal{Y}}$. For convenience, we use a diagonal matrix

$$\Theta[\mathcal{X}, \mathcal{Y}] = \begin{bmatrix} \theta_1 & & & \\ & \theta_2 & & \\ & & \ddots & \\ & & & \theta_{\dim \mathcal{X}} \end{bmatrix} \quad (\text{D21})$$

to record the set of $\dim \mathcal{X}$ canonical angles. Notice when $\dim \mathcal{X} = 1$, the canonical angle is determined by the familiar ℓ^2 -inner product on the Hilbert space. Here we present elementary and known results [74] that are helpful for deriving the block convergence bound.

Result 1. For $\mathcal{Y}_0 \subseteq \mathcal{Y}$ with $\dim \mathcal{Y}_0 = \dim \mathcal{X}$, $\theta_\ell[\mathcal{X}, \mathcal{Y}] \leq \theta_\ell[\mathcal{X}, \mathcal{Y}_0]$ and the equality can be saturated.

Result 2. Suppose $\dim \mathcal{X} = \dim \mathcal{Y}$ and $\theta_\ell[\mathcal{X}, \mathcal{Y}] \neq \pi/2$.

- (a) For $\mathcal{Y}_0 \subseteq \mathcal{Y}$, there exists unique $\mathcal{X}_0 \subseteq \mathcal{X}$ with $\dim \mathcal{X}_0 = \dim \mathcal{Y}_0$ such that $\Pi_{\mathcal{Y}} \mathcal{X}_0 = \mathcal{Y}_0$ for orthogonal projection $\Pi_{\mathcal{Y}}$ onto $\mathcal{Y}$. Moreover,

$$\theta_\ell[\mathcal{X}_0, \mathcal{Y}_0] \leq \theta_\ell[\mathcal{X}, \mathcal{Y}], \quad (\text{D22})$$

for $1 \leq \ell \leq \dim \mathcal{Y}_0$ and hence

$$\|\sin \Theta[\mathcal{X}_0, \mathcal{Y}_0]\| \leq \|\sin \Theta[\mathcal{X}, \mathcal{Y}]\| \quad (\text{D23})$$

for any unitarily invariant norm $\|\cdot\|$.

- (b) For any orthonormal vectors $[y_1, \dots, y_Q]$ in $\mathcal{Y}$, there exists linearly independent vectors $[x_1, \dots, x_Q]$ in $\mathcal{X}$ so that $\Pi_{\mathcal{Y}} x_\ell = y_\ell$ and (a) holds for $\mathcal{X}_0 = \text{span}\{x_1, \dots, x_Q\}$ and $\mathcal{Y}_0 = \text{span}\{y_1, \dots, y_Q\}$ with $\dim \mathcal{X}_0 = \dim \mathcal{Y}_0 = Q$.

3. Theorem D2

Theorem D2. Let $\tilde{E}_n(d)$ be the approximate n th eigenenergy extracted from the dI -dimensional function subspace $\text{span}\{\tilde{g}, \mathcal{K}_{\Delta t}[\tilde{g}], \dots, \mathcal{K}_{\Delta t}^{d-1}[\tilde{g}]\}$, and $\delta E_n(d) = \tilde{E}_n(d) - E_n$ be the approximation error. Consider the error matrix,

$$\Delta_I(d) = \text{diag} \begin{bmatrix} \delta E_0(d) & & \\ & \ddots & \\ & & \delta E_{I-1}(d) \end{bmatrix}, \quad (\text{D24})$$

which contains approximations to the lowest I energies. For $d \geq 1$, there exists time step Δt such that the spectral approximation $\Delta_I(d)$ is bounded by

$$\|\Delta_I(d)\|_2 \leq \frac{\text{cond}(\mathcal{B})^{1/2} |\sin[(E_{N-1} - E_0)\Delta t]|}{\tilde{\epsilon}_{I-1 \rightarrow I}^{2(d-1)} \Delta t} \|\tan^2 \Theta\|_2, \quad (\text{D25})$$

for the operator norm $\|\cdot\|_2$. Here Θ denotes the canonical angle between the two subspaces $\text{span}\{f_n : 0 \leq n \leq I-1\}$ and $\text{span}\{g_i : 1 \leq i \leq I\}$, which generalizes the squared overlap in Theorem D1 (see Appendixes D2 and D3). In the denominator, $\tilde{\epsilon}_{I-1 \rightarrow I} \in [1, 2]$ depends on the I th spectral gap $(E_I - E_{I-1})$ of the Hamiltonian H .

Proof. We begin by treating the case of mutually orthogonal eigenfunctions.

Here, it suffices to establish the more general result for the error matrix $\Delta = \text{diag}[\delta E_{n_1}(d), \delta E_{n_1+1}(d), \dots, \delta E_{n_2}(d)]$ for $n_2 - n_1 + 1 = I$. From here on, we will use the operator norm as a specific example of a unitarily invariant norm by setting $\|\cdot\| = \|\cdot\|_2$. Our key contribution lies in extending the proof of the theorem, building on existing results for block Krylov methods [67], to the real-time setting. Applying Result 2(b), we know there exists $B_I = [x_1, \dots, x_I]$ with $\text{span} B_I = \text{span}\{g_1, \dots, g_I\}$ so that $\Pi_{\text{invariant subspace}} B_I = [f_{n_1}, \dots, f_{n_2}]$. WLOG we assume that $B_I \mapsto B_I(B_I^\dagger B_I)^{-1/2}$ has orthonormal columns, i.e., $\langle x_i, x_j \rangle_\mu = \delta_{ij}$ for $1 \leq i, j \leq I$. For a degree- $(d-1)$ polynomial p , we define

$$R_I = p(\mathcal{K}_{\Delta t}) B_I = \sum_{n=0}^{N-1} f_n p(\lambda_n) f_n^\dagger B_I, \quad (\text{D26})$$

where the dual $f_n^\dagger$ are defined with respect to the functional inner product $\langle \cdot, \cdot \rangle_\mu$. We will further write the equality above as

$$\begin{aligned} R_I &= F p(\Lambda) F^\dagger B_I \\ &= \left[F_a p(\Lambda_a) F_a^\dagger + F_b p(\Lambda_b) F_b^\dagger + F_c p(\Lambda_c) F_c^\dagger \right] B_I, \end{aligned} \quad (\text{D27})$$

where $F = [F_a \ F_b \ F_c]$ and $\Lambda = \Lambda_a \oplus \Lambda_b \oplus \Lambda_c$ contain the N Koopman eigenfunctions and eigenvalues respectively in a block form. Here we use three subscripts a , b , and c to label the partition of Hamiltonian spectrum into three disjoint blocks with energies below E_{n_1} , from E_{n_1} to E_{n_2} , and above E_{n_2} . In particular, the middle block b is represented by $F_b = [f_{n_1}, \dots, f_{n_2}]$ and $\Lambda_b = \text{diag}[\lambda_{n_1}, \dots, \lambda_{n_2}]$.

By our construction, $F_b^\dagger B_I$ is nonsingular so

$$R_I(F_b^\dagger R_I)^{-1} = R_I \left[p(\Lambda_b) F_b^\dagger B_I \right]^{-1} = F_a p(\Lambda_a) F_a^\dagger B_I \left[p(\Lambda_b) F_b^\dagger B_I \right]^{-1} + F_b + F_c p(\Lambda_c) F_c^\dagger B_I \left[p(\Lambda_b) F_b^\dagger B_I \right]^{-1}, \quad (\text{D28})$$

whenever $p(\Lambda_b)$ is nonsingular by suitable choice of the polynomial p . Since $\text{span} R_I$ is clearly a subspace of the Krylov subspace $\text{Kr}_{d,I} := \text{span}\{\vec{g}, \dots, \mathcal{K}_{\Delta t}^{d-1}[\vec{g}]\}$, Result 1 implies

$$\|\tan \Theta[\text{span} F_b, \text{Kr}_{d,I}]\| \leq \|\tan \Theta[\text{span} F_b, \text{span} R_I]\|, \quad (\text{D29})$$

$$= \|\sin \Theta[\text{span} F_b, \text{span} R_I] (\cos \Theta[\text{span} F_b, \text{span} R_I])^{-1}\|, \quad (\text{D30})$$

$$= \left\| [F_a \ F_c]^\dagger R_I (F_b^\dagger R_I)^{-1} \right\| = \left\| \begin{bmatrix} p(\Lambda_a) & 0 \\ 0 & p(\Lambda_c) \end{bmatrix} \begin{bmatrix} F_a^\dagger B_I (F_b^\dagger B_I)^{-1} \\ F_c^\dagger B_I (F_b^\dagger B_I)^{-1} \end{bmatrix} p(\Lambda_b)^{-1} \right\|, \quad (\text{D31})$$

$$\leq \left\| \begin{bmatrix} p(\Lambda_a) & 0 \\ 0 & p(\Lambda_c) \end{bmatrix} \right\| \|p(\Lambda_b)^{-1}\| \left\| \begin{bmatrix} F_a^\dagger B_I (F_b^\dagger B_I)^{-1} \\ F_c^\dagger B_I (F_b^\dagger B_I)^{-1} \end{bmatrix} \right\|, \quad (\text{D32})$$

$$\leq \frac{\max_{n < n_1 \vee n > n_2} |p(\lambda_n)|}{\min_{n_1 \leq n \leq n_2} |p(\lambda_n)|} \|\tan \Theta[\text{span} F_b, \text{span} B_I]\|, \quad (\text{D33})$$

where Eq. (D31) results from the substitution of Eq. (D28).

The inequality above characterizes the convergence of the Krylov subspace (toward the target eigenstates), and our task is to choose a polynomial p that can tightly bound the fraction in Eq. (D33). As in Appendix D 1, we exploit the Rogers-Szegő polynomials $W_{d-1}(z | q)$ and work in the limit $q = 1$ to derive relevant results. For $n_2 = N - 1$, let us consider the polynomial

$$p(z) = \frac{W_{d-1}(e^{-i\varphi} z)}{W_{d-1}(e^{-i\varphi} \lambda_{n_1})}, \quad (\text{D34})$$

where $e^{-i\varphi}$ is a constant phase offset shifting the angles $\vartheta_{n < n_1}$ (recall that $\vartheta_n = E_n \Delta t$) inside the window $\mathcal{W} = [-\Omega, \Omega]$ while keeping the angles $\vartheta_{n_1 \leq n \leq n_2}$ outside $\mathcal{W}$. In this case, observe that

$$\min_{n_1 \leq n \leq n_2} |p(\lambda_n)| = |p(\lambda_{n_1})| = 1, \quad (\text{D35})$$

$$\max_{n < n_1 \vee n > n_2} |p(\lambda_n)| = \max_{n < n_1} |p(\lambda_n)| \leq \frac{1}{|W_{d-1}(e^{-i\varphi} \lambda_{n_1})|}. \quad (\text{D36})$$

By employing suitable time step size Δt , e.g., that from Eq. (D12), we assume $\vartheta_{n_1-1} - \vartheta_0 \leq 2\Omega$ and $2\Omega \leq \vartheta_{N-1} - \vartheta_0 \leq \pi + \Omega$. Specifically, $\varphi = \Omega - \vartheta_{n_1-1}$ circularly shifts the eigenphases $\{\lambda_n\}_{n=0}^{N-1}$. This ensures that $|\varphi + \vartheta_0| \leq \varphi + \vartheta_{n_1-1} = \Omega \leq \varphi + \vartheta_{n_1}$ as intended. From

Theorem D1, we recall $\Omega = \pi/3$ and $|W_{d-1}(e^{-i\varphi} \lambda_{n_1})|^{1/(d-1)} \geq 1 + C\epsilon_{n_1-1 \rightarrow n_1}$ where $\epsilon_{n_1-1 \rightarrow n_1} = 3(E_{n_1} - E_{n_1-1})\Delta t/2\pi$ encodes the normalized n_1 th energy gap and C is some appropriate constant that can be set to 1 [cf. Eq. (D16)]. This immediately implies an exponential convergence to the eigenspaces,

$$\begin{aligned} \|\tan \Theta[\text{span} F_b, \text{Kr}_{d,I}]\| \\ \leq \frac{1}{\epsilon_{n_1-1 \rightarrow n_1}^{d-1}} \|\tan \Theta[\text{span} F_b, \text{span}\{g_1, \dots, g_I\}]\|, \end{aligned} \quad (\text{D37})$$

for the base case $n_2 = N - 1$. When $n_2 < N - 1$, we consider the polynomial $p = p_1 \cdot p_2$ with

$$p_1(z) = \frac{W_{d-N+n_2}(e^{-i\varphi} z)}{W_{d-N+n_2}(e^{-i\varphi} \lambda_{n_1})}, \quad (\text{D38})$$

and

$$p_2(z) = \prod_{n=n_2+1}^{N-1} \frac{z - \lambda_n}{\lambda_{n_2} - \lambda_n}, \quad (\text{D39})$$

where p is factorized into lower order polynomials p_1 of degree $(d - N + n_2)$ and p_2 of degree $(N - n_2 - 1)$. Observe that by design $p(\lambda_n) = 0$ for $n \geq n_2 + 1$ and thus

$$\min_{n_1 \leq n \leq n_2} |p(\lambda_n)| \geq \min_{n_1 \leq n \leq n_2} |p_1(\lambda_n)| \min_{n_1 \leq n \leq n_2} |p_2(\lambda_n)| = \min_{n_1 \leq n \leq n_2} \prod_{m=n_2+1}^{N-1} \left| \frac{\lambda_n - \lambda_m}{\lambda_{n_2} - \lambda_m} \right|, \quad (\text{D40})$$

$$\max_{n < n_1 \vee n > n_2} |p(\lambda_n)| \leq \max_{n < n_1} |p_1(\lambda_n)| \max_{n < n_1} |p_2(\lambda_n)| \leq \frac{1}{|W_{d-N+n_2}(e^{-i\varphi} \lambda_{n_1})|} \max_{n < n_1} \prod_{m=n_2+1}^{N-1} \left| \frac{\lambda_n - \lambda_m}{\lambda_{n_2} - \lambda_m} \right|, \quad (\text{D41})$$

which directly follows from applying our base case result to the eigensector $\{\lambda_n\}_{n=0}^{n_2}$. Note that our argument requires $d \geq N - n_2$: a symmetric bound can be readily established using the exact same argument with the roles of n_1 and n_2 exchanged, i.e., for $d \geq n_1 + 1$, upon a spectral flip $H \mapsto -H$.

Now we proceed to analyze the convergence of the Krylov energy approximation. For simplicity, we focus again on approximating from below Hamiltonian eigenvalues $E_n \lesssim E_{N-1}$ near the right edge of the spectrum. The argument for approximating from above eigenvalues near the left edge of the spectrum can be easily adapted with a spectral flip. We observe that up to the trivial spectral rotation $\mathcal{K}_{\Delta t} \mapsto \lambda_{n_2}^{-1} \mathcal{K}_{\Delta t}$ leaving the functional Krylov subspace invariant, we can assume WLOG $E_{n_2-1} \leq 0 < E_{n_2+1}$ for the remainder of the proof. For $V_0 := R_I(F_b^\dagger R_I)^{-1}$, we first construct $V_- = V_0(V_0^\dagger V_0)^{-1/2}$, which contains orthonormal columns. This allows us to introduce eigenvalues, $\hat{\lambda}_n = |\hat{\lambda}_n| e^{-i\hat{E}_n \Delta t}$ for $n_1 \leq n \leq n_2$, of the $\text{span} V_0$ -projected Koopman operator $V_-^\dagger \mathcal{K}_{\Delta t} V_-$, with the ordering $\arg(\hat{\lambda}_{n_1}) \geq \dots \geq \arg(\hat{\lambda}_{n_2})$. These eigenvalues depend explicitly on choices of the degree- $(d-1)$ polynomial p (as R_I does). Since $\text{span} V_0 \subseteq \text{Kr}_{d,I}$ irrespective of p , the variational principle gives

$$\begin{aligned} \hat{E}_n(d) \leq \tilde{E}_n(d) \leq E_n &\implies 0 \leq \delta E_n(d) \\ &\leq E_n - \hat{E}_n(d), \quad n_1 \leq n \leq n_2. \end{aligned} \quad (\text{D42})$$

The variational characterization, as we shall see, leads to an upper bound on the spectral error

$$\Delta_{I,n_1,n_2} = \text{diag} \begin{bmatrix} \delta E_{n_1} & & \\ & \ddots & \\ & & \delta E_{n_2} \end{bmatrix}, \quad (\text{D43})$$

which generalizes Δ_I in Theorem D2 ($\Delta_{I,n_1,n_2} = \Delta_I$ when $n_1 = N - I$ and $n_2 = N - 1$). With our previous assumption $E_{n_2} = 0$, we have $\hat{E}_n \leq 0$ for $n_1 \leq n \leq n_2$. Consequently, $\forall x \in \mathbb{C}^I$

$$-\frac{1}{\Delta t} \theta(V_0 x) = -\frac{1}{\Delta t} \arg \frac{\langle V_0 x, \mathcal{K}_{\Delta t} V_0 x \rangle_\mu}{\langle V_0 x, V_0 x \rangle_\mu} \leq 0, \quad (\text{D44})$$

where we adopt the same notation as in Theorem D1 [cf. Eq. (D2)]. By unfolding V_0 via Eq. (D28), we arrive at

$$V_0^\dagger \mathcal{K}_{\Delta t} V_0 = G_a^\dagger \Lambda_a G_a + \Lambda_b + G_c^\dagger \Lambda_c G_c \quad (\text{D45})$$

and

$$V_0^\dagger V_0 = G_a^\dagger G_a + \text{Id}_I + G_c^\dagger G_c, \quad (\text{D46})$$

for $G_a = p(\Lambda_a) F_a^\dagger B_I [p(\Lambda_b) F_b^\dagger B_I]^{-1}$ and $G_c = p(\Lambda_c) F_c^\dagger B_I [p(\Lambda_b) F_b^\dagger B_I]^{-1}$. Substituting Eqs. (D45) and (D46) into the variational inequality, we can derive

$$-\frac{1}{\Delta t} \arg \frac{x^\dagger (G_a^\dagger \Lambda_a G_a + \eta_b \lambda_{n_2} \text{Id}_I) x}{x^\dagger x} - \frac{1}{\Delta t} \arg \frac{x^\dagger \Lambda_b x}{x^\dagger x} + \frac{1}{\Delta t} \arg \lambda_{n_2} \leq -\frac{1}{\Delta t} \arg \frac{x^\dagger (G_a^\dagger \Lambda_a G_a + \Lambda_b) x}{x^\dagger x}, \quad (\text{D47})$$

$$\leq -\frac{1}{\Delta t} \arg \frac{x^\dagger (G_a^\dagger \Lambda_a G_a + \Lambda_b + G_c^\dagger \Lambda_c G_c) x}{x^\dagger (G_a^\dagger G_a + \text{Id}_I + G_c^\dagger G_c) x}, \quad (\text{D48})$$

$$= -\frac{1}{\Delta t} \arg \frac{\langle V_0 x, \mathcal{K}_{\Delta t} V_0 x \rangle_\mu}{\langle V_0 x, V_0 x \rangle_\mu} \leq 0, \quad (\text{D49})$$

where $\eta_b = \inf_{x \in \mathbb{C}^I: x^\dagger x = 1} \|x^\dagger \Lambda_b x\|$ is a constant of $\mathcal{O}(1)$ when $\vartheta_{n_2} - \vartheta_{n_1} < \pi/2$, and we recall that $E_{n_2} = 0 \implies \lambda_{n_2} = 1$ in our current setting. In particular, the inequality Eq. (D47) above holds since for any unit vector $x \mapsto x/\sqrt{x^\dagger x}$,

$$\arg \frac{\lambda_{n_2} x^\dagger (G_a^\dagger \Lambda_a G_a + \Lambda_b) x}{x^\dagger (G_a^\dagger \Lambda_a G_a + \eta_b \lambda_{n_2} \text{Id}_I) x \cdot x^\dagger \Lambda_b x} = \arg \frac{\frac{\lambda_{n_2}}{x^\dagger \Lambda_b x} + \frac{\lambda_{n_2}}{x^\dagger G_a^\dagger \Lambda_a G_a x}}{1 + \frac{\eta_b \lambda_{n_2}}{x^\dagger G_a^\dagger \Lambda_a G_a x}} \leq \arg \frac{\frac{\lambda_{n_2}}{x^\dagger \Lambda_b x} + \frac{\lambda_{n_2}}{x^\dagger G_a^\dagger \Lambda_a G_a x}}{1 + \frac{\|x^\dagger \Lambda_b x\| \lambda_{n_2}}{x^\dagger G_a^\dagger \Lambda_a G_a x}} \leq 0, \quad (\text{D50})$$

where the numerator has a smaller phase than the denominator. The inequality Eq. (D49), on the other hand, follows from our basic assumption that $E_n \leq 0$ for $n \leq n_2$. The inequalities indicate that the $G_a^\dagger \Lambda_a G_a$ eigenvalues, denoted as $\bar{\lambda}_n$ for $n_1 \leq n \leq n_2$ (with the eigenvalues ordered by decreasing arguments), are related to the $V_-^\dagger \mathcal{K}_{\Delta t} V_-$ eigenvalues through

$$-\frac{1}{\Delta t} \arg(\bar{\lambda}_n + \eta_b \lambda_{n_2}) + E_n - E_{n_2} \leq \hat{E}_n, \quad n_1 \leq n \leq n_2. \quad (\text{D51})$$

Consequently, we may relax the error bound in Eq. (D42),

$$\delta E_n \leq E_n - \hat{E}_n \leq \frac{1}{\Delta t} \arg(\bar{\lambda}_n + \eta_b \lambda_{n_2}) + E_{n_2}, \quad (\text{D52})$$

$$\leq \frac{1}{\Delta t} \arg \left[\lambda_0 \frac{e_n^\dagger G_a^\dagger G_a e_n}{\eta_b e_n^\dagger e_n} + \lambda_{n_2} \right] + E_{n_2}, \quad (\text{D53})$$

$$\leq \frac{1}{\Delta t} \arg \left[\lambda_0 \frac{y_n^\dagger G_a^\dagger G_a y_n}{\eta_b y_n^\dagger y_n} + \lambda_{n_2} \right] + E_{n_2}, \quad (\text{D54})$$

where e_n denotes the eigenvector of $G_a^\dagger \Lambda_a G_a$ corresponding to $\bar{\lambda}_n$ and y_n the eigenvector of $G_a^\dagger G_a$ corresponding to the $(n - n_1 + 1)$ th largest eigenvalue. Importantly, we remark that Eqs. (D47) and (D54) can be established without phase ambiguity of $2\pi\mathbb{Z}$ by choosing suitable time step size Δt .

We therefore seek a tight bound on the eigenvalues of $G_a^\dagger G_a$ (or equivalently the singular values of G_a). Notice that the singular values of $p(\Lambda_a)^{-1} \cdot G_a \cdot p(\Lambda_b) = p(\Lambda_a)^{-1} \cdot p(\Lambda_a) F_a^\dagger B_I [p(\Lambda_b) F_b^\dagger B_I]^{-1} \cdot p(\Lambda_b) = F_a^\dagger B_I [F_b^\dagger B_I]^{-1}$ are bounded above by the singular values of the canonical tangents [cf. Eq. (D33)]

$$\tan \Theta[\text{span} F_b, \text{span} B_I] = \begin{bmatrix} F_a^\dagger B_I (F_b^\dagger B_I)^{-1} \\ F_c^\dagger B_I (F_b^\dagger B_I)^{-1} \end{bmatrix}. \quad (\text{D55})$$

This allows us to effectively control the spectral radius of $G_a^\dagger G_a$,

$$\begin{aligned} \|G_a^\dagger G_a\| &\leq \|p(\Lambda_a)\|^2 \|p(\Lambda_b)^{-1}\|^2 \|\tan^2 \Theta[\text{span} F_b, \text{span} B_I]\|, \\ &= \frac{\max_{n < n_1} |p(\lambda_n)|^2}{\min_{n_1 \leq n \leq n_2} |p(\lambda_n)|^2} \|\tan^2 \Theta[\text{span} F_b, \text{span} B_I]\|, \end{aligned} \quad (\text{D56})$$

where it suffices to identify a degree- $(d-1)$ polynomial p that tightly bounds the fraction on the RHS of the

expression above. In addition, here p must satisfy the key orthogonality constraints,

$$\tilde{f}_n^\dagger R_I = [0 \ 0 \ \dots \ 0], \quad n_2 + 1 \leq n \leq N - 1, \quad (\text{D58})$$

where $\tilde{f}_n \in \text{Kr}_{d,I}$ denotes the approximate eigenfunction corresponding to the Krylov eigenvalue $\tilde{\lambda}_n = |\tilde{\lambda}_n| e^{-i\tilde{E}_n \Delta t}$. The set of $(N - n_2)$ constraints ensure that $\text{span} R_I = p(\mathcal{K}_{\Delta t}) \text{span} B_I = p(\mathcal{K}_{\Delta t}) \text{span}\{g_1, \dots, g_I\}$ belongs to the correct eigensector of the Krylov subspace. Let us consider the familiar polynomial $p = p_1 \cdot p_2$ with p_1 of degree- $(d - N + n_2)$ and p_2 of degree $(N - n_2 - 1)$ given by

$$p_1(z) = \frac{W_{d-N+n_2}(e^{-i\varphi} z)}{W_{d-N+n_2}(e^{-i\varphi} \lambda_{n_1})} \quad (\text{D59})$$

and

$$p_2(z) = \prod_{n=n_2+1}^{N-1} \frac{z - \tilde{\lambda}_n}{\lambda_{n_2} - \tilde{\lambda}_n}, \quad (\text{D60})$$

as introduced in Eqs. (D38) and (D39) respectively. Accordingly, we can generalize Eq. (D17) from Theorem D1,

$$\delta E_n \leq -\frac{1}{\Delta t} \arctan \left[\frac{\zeta_{n_1, n_2} \sin \vartheta_0 + \sin \vartheta_{n_2}}{\zeta_{n_1, n_2} \cos \vartheta_0 + \cos \vartheta_{n_2}} \right] + E_{n_2}, \quad (\text{D61})$$

$$= \frac{\sin(\vartheta_{n_2} - \vartheta_0)}{\Delta t} \zeta_{n_1, n_2} + \mathcal{O}(\zeta_{n_1, n_2}^2), \quad (\text{D62})$$

where $\zeta_{n_1, n_2} = \eta_b^{-1} \iota_{n_1, n_2}^2 \tilde{\epsilon}_{n_1-1 \rightarrow n_1}^{-2(d-N+n_2)} \|\tan^2 \Theta\|$ for

$$\iota_{n_1, n_2} = \begin{cases} 1 & \text{when } n_2 = N - 1, \\ \frac{\max_{n < n_1} \prod_{m=n_2+1}^{N-1} \left| \frac{\lambda_n - \tilde{\lambda}_m}{\lambda_{n_2} - \tilde{\lambda}_m} \right|}{\min_{n_1 \leq n \leq n_2} \prod_{m=n_2+1}^{N-1} \left| \frac{\lambda_n - \tilde{\lambda}_m}{\lambda_{n_2} - \tilde{\lambda}_m} \right|} & \text{otherwise.} \end{cases} \quad (\text{D63})$$

This immediately implies

$$\|\Delta_{I, n_1, n_2}\| \leq \frac{\iota_{n_1, n_2}^2 |\sin[(E_{n_2} - E_0) \Delta t]| \|\tan^2 \Theta\|}{\eta_b \tilde{\epsilon}_{n_1-1 \rightarrow n_1}^{-2(d-N+n_2)} \Delta t}, \quad (\text{D64})$$

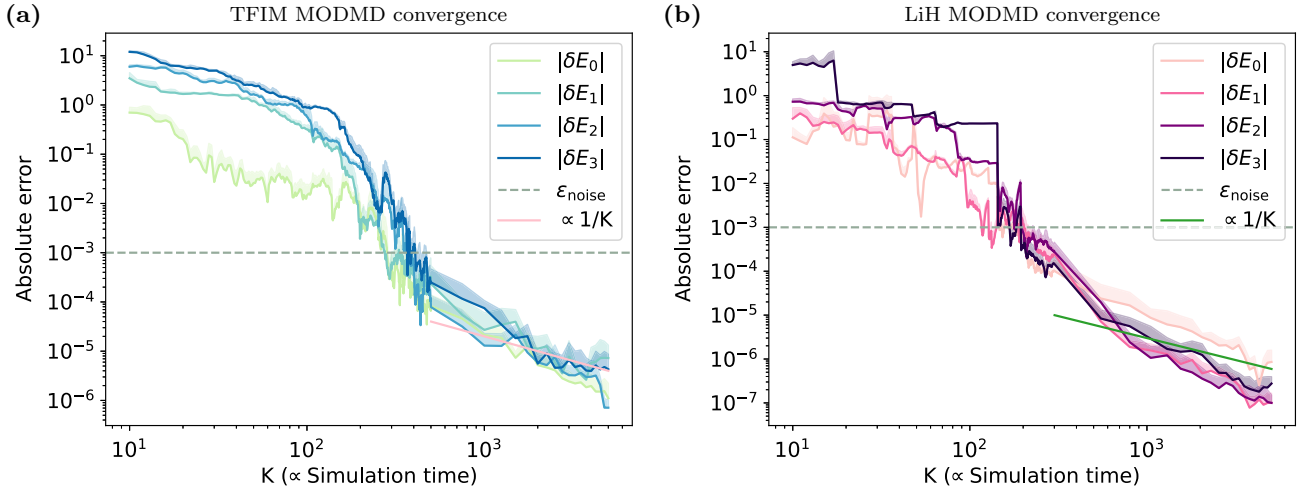


FIG. 9. Convergence of eigenenergies for transverse-field Ising model and lithium hydride. To obtain eigenenergy estimates $\tilde{E}_n$, we fix the multiobservable dynamic mode decomposition parameters $K/d = \frac{5}{2}$ and $\tilde{\delta} = 10^{-2}$ for constructing and thresholding the data matrix pair $\mathbf{X}, \mathbf{X}' \in \mathbb{R}^{dl \times (K+1)}$. Gaussian $\mathcal{N}(0, \varepsilon_{\text{noise}}^2)$ noise with $\varepsilon_{\text{noise}} = 10^{-3}$ is added independently to distinct matrix elements. The absolute errors, $|\tilde{E}_n - E_n|$, in the first four lowest target eigenenergies of the transverse-field Ising model and lithium hydride Hamiltonians are shown with respect to K proportional to the nondimensional maximal simulation time. The reference states $|\phi_0\rangle$ are sparse in the computational basis as prescribed in Appendix F 1. The shading above the solid lines indicates standard deviation across repeated trials for each quantity. (a) Transverse-field Ising model absolute errors from the multiobservable dynamic mode decomposition algorithm with $\Delta t \approx 0.08$ and $I = 7$ (plotted on a log-log scale), where the convergence results are averaged over five trials, each involving a Gaussian noise realization and a random selection of $I - 1$ observables. (b) Lithium hydride absolute errors from the multiobservable dynamic mode decomposition algorithm with $\Delta t \approx 0.39$ and $I = 7$ (on a log-log scale), where the results are averaged over 20 trials, each involving a Gaussian noise realization.

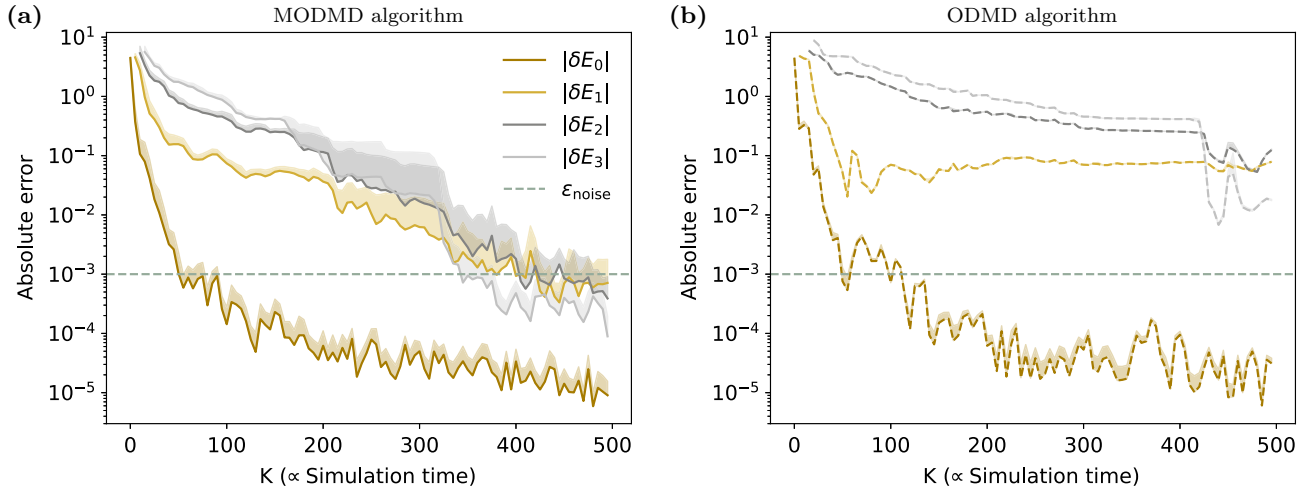


FIG. 10. Convergence of eigenenergies for the isotropic Heisenberg model. To obtain eigenenergy estimates $\tilde{E}_n$, we fix the multiobservable dynamic mode decomposition parameters $K/d = \frac{5}{2}$ and $\tilde{\delta} = 10^{-2}$ for constructing and thresholding the pair of data matrices $\mathbf{X}, \mathbf{X}' \in \mathbb{R}^{dl \times (K+1)}$. Gaussian $\mathcal{N}(0, \varepsilon_{\text{noise}}^2)$ noise with $\varepsilon_{\text{noise}} = 10^{-3}$ is added independently to the real or/and imaginary parts of the matrix elements. Absolute errors, $|\tilde{E}_n - E_n|$, in the first four eigenenergies of the Heisenberg Hamiltonian are shown with respect to K proportional to the nondimensional maximal simulation time. The reference state $|\phi_0\rangle$ is an equal superposition of four computational basis states (see Appendix F 1) and we employ a time step of $\Delta t \approx 0.09$. The shading above the solid and dashed lines shows the standard deviation across trials for each quantity. For the Heisenberg model, the Hamiltonian is of size $32\,768 \times 32\,768$ and the largest linear system in the corresponding multiobservable dynamic mode decomposition least-squares problem has size 1400×501 . (a) Absolute errors from the multiobservable (multiobservable dynamic mode decomposition) algorithm with $I = 7$ distinct observables, where the convergence results are averaged over 20 trials, each involving a Gaussian noise realization and a selection of $I - 1$ random observables. (b) Absolute errors from the single-observable (observable dynamic mode decomposition) algorithm where the convergence results are averaged over 20 trials, each involving a Gaussian noise realization.

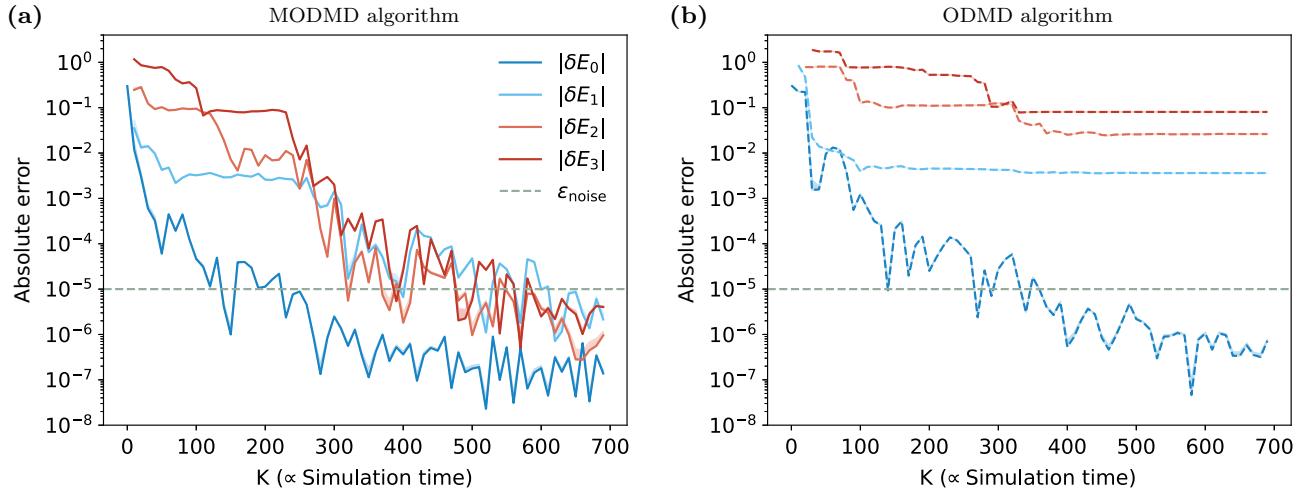


FIG. 11. Convergence of eigenenergies for the beryllium hydride molecule. To obtain eigenenergy estimates $\tilde{E}_n$, we fix the multiobservable dynamic mode decomposition parameters $K/d = \frac{5}{2}$ and $\tilde{\delta} = 10^{-4}$ for constructing and thresholding the pair of data matrices $\mathbf{X}, \mathbf{X}' \in \mathbb{R}^{dl \times (K+1)}$. Gaussian $\mathcal{N}(0, \epsilon_{\text{noise}}^2)$ noise with $\epsilon_{\text{noise}} = 10^{-5}$ is added independently to the real or/and imaginary parts of the matrix elements. The reference state $|\phi_0\rangle$ is a superposition of six Slater determinant states (see Appendix F 1) and we employ a time step of $\Delta t = 0.20$. The absolute errors, $|\tilde{E}_n - E_n|$, in the first four eigenenergies of the beryllium hydride Hamiltonian are shown with respect to K proportional to the nondimensional maximal simulation time. Convergence results are averaged over 20 trials of Gaussian noise realizations with shading above the solid and dashed lines showing the standard deviation across trials for each quantity. For beryllium hydride, the model Hamiltonian is of size 4096×4096 and the largest linear system in the corresponding multiobservable dynamic mode decomposition least-squares problem has size 1960×701 . (a) Absolute errors from the multiobservable (multiobservable dynamic mode decomposition) algorithm with $I = 7$ observable types. (b) Absolute errors from the single-observable (observable dynamic mode decomposition) algorithm.

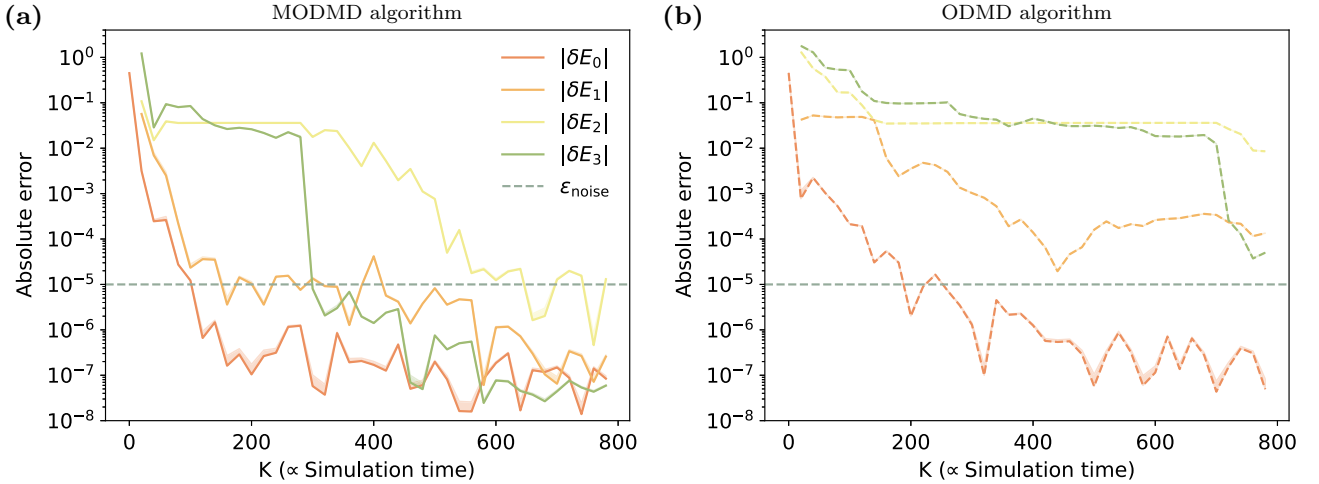


FIG. 12. Convergence of eigenenergies for molecular nitrogen (N_2). To obtain eigenenergy estimates $\tilde{E}_n$, we fix the multiobservable dynamic mode decomposition parameters $K/d = \frac{5}{2}$ and $\tilde{\delta} = 10^{-4}$ for constructing and thresholding the pair of data matrices $\mathbf{X}, \mathbf{X}' \in \mathbb{R}^{dl \times (K+1)}$. Gaussian $\mathcal{N}(0, \epsilon_{\text{noise}}^2)$ noise with $\epsilon_{\text{noise}} = 10^{-5}$ is added independently to the real or/and imaginary parts of the matrix elements. The reference state $|\phi_0\rangle$ is a superposition of six Slater determinant states (see Appendix F 1) and we use a time step of $\Delta t = 0.37$. The absolute errors, $|\tilde{E}_n - E_n|$, in the first four eigenenergies of the N_2 Hamiltonian are shown with respect to K proportional to the nondimensional maximal simulation time. Convergence results are averaged over 20 trials of Gaussian noise realizations with shading above the solid and dashed lines showing the standard deviation across trials for each quantity. For N_2 , the model Hamiltonian is of size 16384×16384 and the largest linear system in the corresponding multiobservable dynamic mode decomposition least-squares problem has size 2240×801 . (a) Absolute errors from the multiobservable (multiobservable dynamic mode decomposition) algorithm with $I = 7$ observable types. (b) Absolute errors from the single-observable (observable dynamic mode decomposition) algorithm.

which simplifies to the desired result when $n_2 = N - 1$ as claimed. Finally, the dependence on conditioning, $\text{cond}(\mathcal{B})^{1/2}$, can be reinstated to accommodate the nonorthogonal eigenfunction cases in analogy with Theorem D1. ■

Appendix E: Conditioning of Koopman eigenfunctions in degenerate energy subspaces

1. Single-observable case

Let us first consider the case when the ground-state energy is degenerate. All Koopman eigenfunctions corresponding to E_0 acquire the same phase factor at each time step. Consequently, the single-observable signal collapses a degenerate eigenspace into a single spectral mode: while the energy E_0 remains identifiable, the individual Hamiltonian eigenstates do not. The mode $f_0(|\phi\rangle) = \langle\psi_0 | \phi\rangle$ is replaced by an effective eigenfunction,

$$\tilde{f}_0(|\phi\rangle) = \sum_{n:E_n=E_0} w_n \langle\psi_n | \phi\rangle, \quad \sum_{n:E_n=E_0} |w_n|^2 = 1, \quad (\text{E1})$$

which represents the ground-state subspace. In the presence of degeneracies, the condition number $\text{cond}(\mathcal{B})$ in Theorem D1 should be interpreted as $\text{cond}(\tilde{\mathcal{B}})$, where $\tilde{\mathcal{B}}$ is the smaller $L^2(\mu)$ Gram matrix formed from the effective eigenfunctions associated with distinct energies. In other words, degenerate ground states (in fact any degenerate eigenstates) become linearly dependent under our $L^2(\mu)$ inner product and can never be resolved by a single-observable signal.

2. Multiobservable case

In comparison, degeneracies can be resolved by multiobservable signals. To see this, we revisit the example above of the degenerate ground-state energy E_0 with some multiplicity r_0 . Our signals $\tilde{g}(|\phi_k\rangle)$ contain, up to the common phase $e^{-iE_0 k \Delta t}$, a collection of amplitudes $\mathcal{A}_{i\alpha} = \langle\phi_0 | O_i | \psi_{0,\alpha}\rangle$ across observables O_i and degenerate states $\{|\psi_{0,\alpha}\rangle\}_{\alpha=1}^{r_0}$. The amplitudes span a data-identifiable subspace, whose dimension is the rank of the $I \times r_0$ amplitude matrix $\mathcal{A}$. The ground-state mode thereby decomposes into $\text{rank}(\mathcal{A})$ effective eigenchannels,

$$\tilde{f}_{E_0,b}(|\phi\rangle) = \sum_{\alpha=1}^{r_0} w_{b,(0,\alpha)} \langle\psi_{0,\alpha} | \phi\rangle, \quad b = 1, \dots, \text{rank}(\mathcal{A}), \quad (\text{E2})$$

with the weight vectors $w_b \in \mathbb{C}^{r_0}$ forming an orthonormal basis of the row space $\text{row}(\mathcal{A})$.

More generally, the conditioning factor in Theorem D2 should be interpreted as the condition number of a *block Gram matrix* $\mathcal{B}_{\text{block}}$, constructed from Koopman eigenchannels $\tilde{f}_{E,b}$ that are distinguishable under the $L^2(\mu)$

inner product. With this replacement, $\text{rank}(\mathcal{A})$ matches the ground-state multiplicity whenever the chosen observables resolve ground-state subspace and $\text{cond}(\mathcal{B}_{\text{block}})$ becomes $\text{cond}(\tilde{\mathcal{B}})$ in the single-observable limit. Therefore, the block Krylov structure underlying MODMD conceptually points to its ability to extract degenerate energies, and, with suitable observables, the associated eigenspaces.

Appendix F: Additional simulation details

1. Reference states

Below we provide the exact reference states $|\phi_0\rangle$ employed for the TFIM and LiH calculations. The TFIM reference is an equal superposition of six computational basis states,

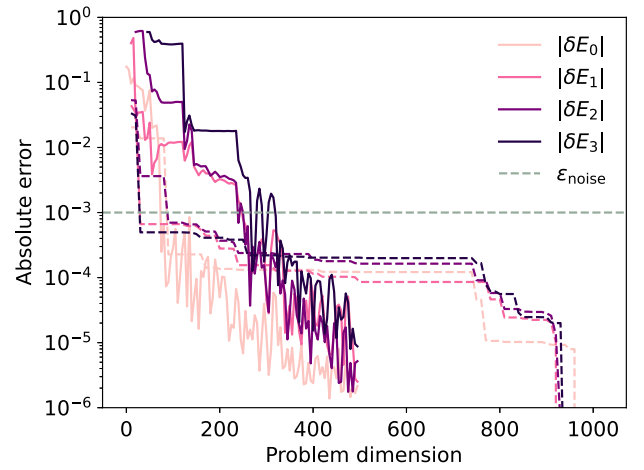


FIG. 13. Performance comparison of multiobservable dynamic mode decomposition and truncated configuration interaction for the lithium hydride molecule. To obtain eigenenergy estimates $\tilde{E}_n$, we fix the multiobservable dynamic mode decomposition parameters $K/d = \frac{5}{2}$ and $\tilde{\delta} = 10^{-2}$ for constructing and thresholding data matrices $\mathbf{X}, \mathbf{X}' \in \mathbb{R}^{df \times (K+1)}$. Gaussian $\mathcal{N}(0, \varepsilon_{\text{noise}}^2)$ noise with $\varepsilon_{\text{noise}} = 10^{-3}$ is added independently to the real or/and imaginary parts of the matrix elements. The reference state $|\phi_0\rangle$ is a superposition of six Slater determinants (see Appendix F 1) and we employ a time step of $\Delta t \approx 0.39$. As a baseline, eigenenergy estimates are also computed through exact (and therefore classical) diagonalization of truncated configuration interaction Hamiltonian, where we project the full problem onto a subspace spanned by lowest-energy Slater determinants. The absolute errors, $|\tilde{E}_n - E_n|$, in the first four eigenenergies of the lithium hydride Hamiltonian are shown with respect to K proportional to the nondimensional maximal simulation time. Multiobservable dynamic mode decomposition convergence results are averaged over 20 trials of Gaussian noise realizations. Solid and dashed lines indicate the multiobservable dynamic mode decomposition and truncated configuration interaction estimates. For lithium hydride, the full Hamiltonian is of size 1024×1024 whereas the largest linear system in the corresponding multiobservable dynamic mode decomposition least-squares problem has size 1400×501 .

$$|\phi_0\rangle_{\text{TFIM}} = \frac{1}{\sqrt{6}} (|0000000000000000\rangle + |1111111111111111\rangle + |1000000000000000\rangle \\ + |0000000011111111\rangle + |0000000111111111\rangle + |0000001111111111\rangle),$$

while the LiH reference is a superposition of six Slater determinant states,

$$|\phi_0\rangle_{\text{LiH}} = \frac{1}{3} (|0000100001\rangle + 2|0001100001\rangle + |0000100011\rangle + |0111100001\rangle + |0000101111\rangle + |0000100111\rangle),$$

where we work with the STO-3G basis set (two core and two valence electrons) to represent the second-quantized molecular Hamiltonian. Note that both reference states are superpositions of a small number of computational basis states and thus have efficient classical representation.

The reference states for the additional simulations in Appendix G are

$$|\phi_0\rangle_{\text{Heisenberg}} = \frac{1}{2} (|0000000000000000\rangle + |1000000000000000\rangle + |1100000000000000\rangle + |1110000000000000\rangle), \\ |\phi_0\rangle_{\text{BeH}_2} = \frac{1}{3} (|000010111101\rangle + |001110111101\rangle + |000110111101\rangle \\ + |000010110001\rangle + 2|000010111001\rangle + |001100111101\rangle) \\ |\phi_0\rangle_{\text{N}_2} = \frac{1}{3} (|00010101110101\rangle + |00010101000101\rangle + 2|00010101100101\rangle \\ + |01110101110101\rangle + |00110101110101\rangle + |00010101001001\rangle).$$

The unequal superposition for all three molecular Hamiltonians (LiH, BeH₂, and N₂) was necessary to ensure nonzero overlap with the first four eigenstates of each system.

For our comparative study between MODMD and QMEGS in Sec. VC 1, we employ the reference state,

$$|\phi_0\rangle_{\text{TFIM}} = \frac{1}{\sqrt{6}} (|00000000000000\rangle + |00000000000001\rangle + |00011111111111\rangle \\ + |10000000000000\rangle + |11111111111110\rangle + |11111111111111\rangle). \quad (\text{F1})$$

For our comparative study between MODMD and (U)VQPE in Sec. VC 2, we employ the reference state,

$$|\phi_0\rangle_{\text{TFIM}} = \frac{1}{\sqrt{18}} (|0000000000000000\rangle + |1111111111111111\rangle + |1000000000000000\rangle + |0000000011111111\rangle \\ + |0000000111111111\rangle + |0000001111111111\rangle + |0000011111111111\rangle + |0000111111111111\rangle \\ + |0001111111111111\rangle + |0011111111111111\rangle + |1111111111111110\rangle + |0000000000000001\rangle \\ + |1100000000000000\rangle + |1110000000000000\rangle + |1111000000000000\rangle + |1111100000000000\rangle \\ + |1111110000000000\rangle + |1111111000000000\rangle), \quad (\text{F2})$$

which comprises a superposition of 18 computational basis states and is less structured than $|\phi_0\rangle_{\text{TFIM}}$ in Sec. VA.

2. Convergence of eigenenergies

In order to better illustrate the regime division in error convergence, as explained in Sec. V, we provide alternative visualizations of Figs. 2(a) and 4(a). Figure 9 shows the absolute energy error as a function of time steps on

the log-log scale for a larger range of K values, highlighting the transition from the exponential to algebraic convergence. As the number of time steps increases, the log-log plot demonstrates the linear behavior expected for algebraic error decay.

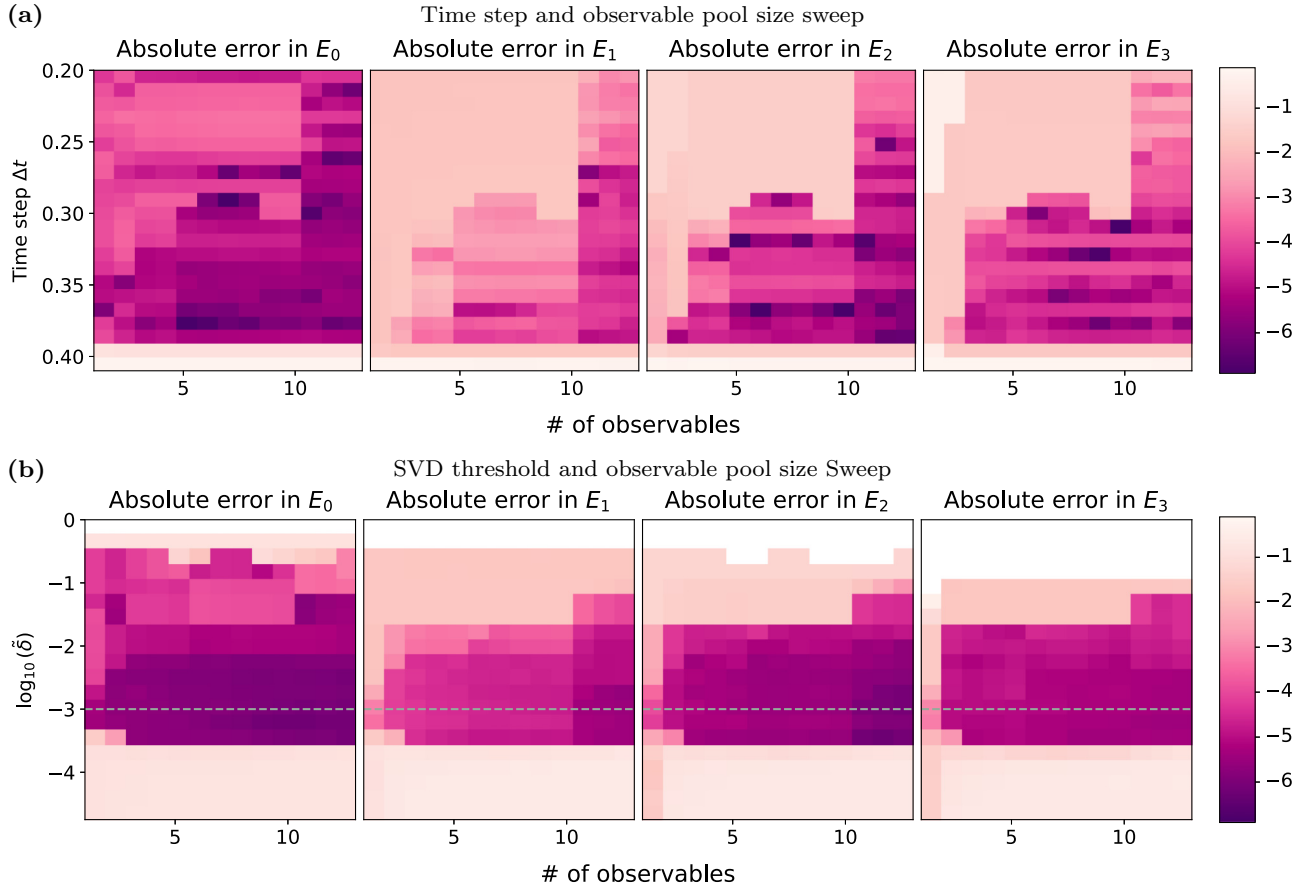


FIG. 14. Parameter sweeps for convergence of eigenenergies for the lithium hydride molecule. To obtain eigenenergy estimates $\tilde{E}_n$, we fix the multiobservable dynamic mode decomposition parameters $K/d = \frac{5}{2}$ and $K = 500$ for constructing and thresholding the pair of data matrices $\mathbf{X}, \mathbf{X}' \in \mathbb{R}^{dI \times (K+1)}$. Gaussian $\mathcal{N}(0, \varepsilon_{\text{noise}}^2)$ noise with $\varepsilon_{\text{noise}} = 10^{-3}$ is added independently to the real or/and imaginary parts of the matrix elements. The reference state $|\phi_0\rangle$ is a superposition of six Slater determinants (see Appendix F 1). The absolute errors, $|\tilde{E}_n - E_n|$, in the first four eigenenergies of the lithium hydride Hamiltonian are shown where convergence results are averaged over 20 trials of Gaussian noise realizations. (a) Absolute errors from the multiobservable dynamic mode decomposition algorithm for varying Δt and I values. The singular value threshold is fixed at the empirically near-optimal value of $\tilde{\delta} = 10^{-2}$. (b) Absolute errors from the multiobservable dynamic mode decomposition algorithm for varying $\tilde{\delta}$ and I values. The time step is fixed at the empirically near-optimal value of $\Delta t \approx 0.39$.

APPENDIX G: ADDITIONAL BENCHMARKS ON QUANTUM SYSTEMS

In this appendix, we present additional benchmark results similar to those presented in Sec. V, but for different quantum systems. The behaviors illustrated in Figs. 10–12 are consistent with those highlighted in Sec. V. While the ground-state energy converges at a somewhat comparable rate for both the single- and multi-observable algorithms, MODMD achieves high accuracies for the excited-state energies significantly more rapidly, reaching errors below the noise level for relatively small K .

1. Lithium Hydride Molecule

While increasing the number of observables enhances richness of the signal subspace and thus improves accuracy

at a favorable quantum cost, it unavoidably raises the classical cost of solving the relevant matrix LS problem in MODMD. To more quantitatively assess the trade-off between quantum and classical resources, we compare the performance of MODMD with that of the classical truncated configuration interaction (CI) method on the LiH molecule.

In particular we can construct the truncated CI Hamiltonian [96] in a natural way: for a given subspace dimension, we select the Slater determinants with the lowest diagonal energies and project the full CI Hamiltonian onto the corresponding low-energy subspace. For each linear system solved in the MODMD LS problem, we solve a truncated CI problem of the same subspace dimension, therefore enabling a direct comparison between the two approaches in terms of accuracy and computational cost.

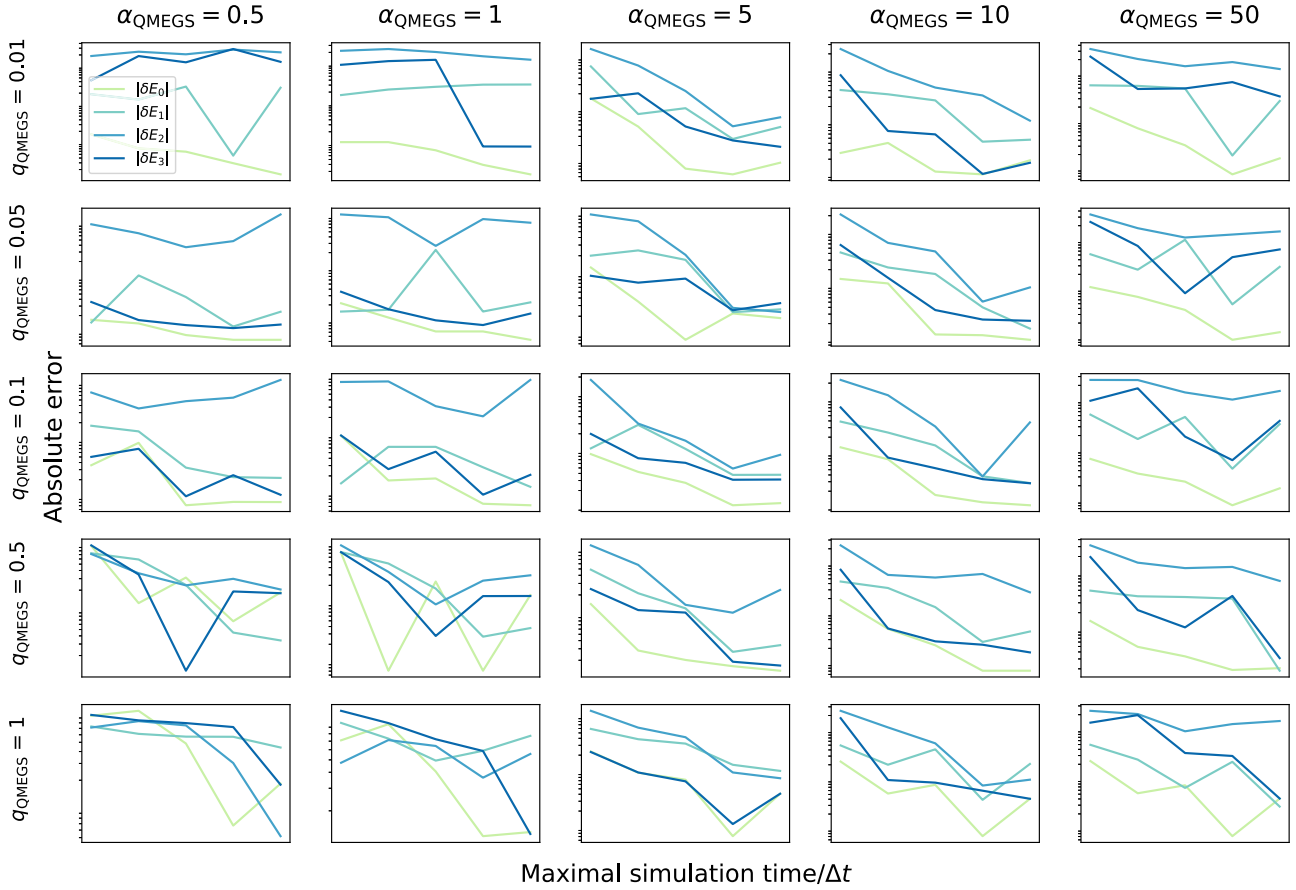


FIG. 15. Convergence of quantum multiple eigenvalue Gaussian filtered search eigenenergy approximations on the 12-qubit transverse-field Ising model for varying values of the quantum multiple eigenvalue Gaussian filtered search hyperparameters α_{QMEGS} and q_{QMEGS} .

In Fig. 13, we observe that the classical truncated CI method outperforms MODMD when the number of time steps is either relatively small—due to the greedy selection of initial Slater determinant(s)—or relatively large, where the CI subspace approaches exact diagonalization. For intermediate values of time step number, however, MODMD achieves measurably higher accuracies, outperforming truncated CI by one to two orders of magnitude. Such a nonmonotonic dependence points to a regime in which MODMD offers a favorable balance of accuracy and resource efficiency.

2. Quantum Heisenberg Model

As another example from condensed matter physics, we examine the convergence of the ground and excited-state eigenenergies of the 15-qubit isotropic (also known as XXX) Heisenberg model subject to open boundary conditions. The system Hamiltonian is given by

$$H_{\text{Heisenberg}} = -J \sum_{i=1}^{L-1} X_i X_{i+1} + Y_i Y_{i+1} + Z_i Z_{i+1} - h \sum_{i=1}^L Z_i. \quad (\text{G1})$$

The strategy used to select the observable pool is the same as in the TFIM calculations, and the reference state is an equal superposition of four computational basis states.

3. Beryllium Hydride Molecule

The second chemical Hamiltonian for which we evaluate the performance of MODMD is that of beryllium hydride (BeH_2). The 12-qubit Hamiltonian is generated in an equilibrium linear geometry with a bond length of $\approx 2.67 \text{ \AA}$ using the STO-3G basis set and parity fermion-to-qubit mapping. The observable pool was constructed in a similar fashion to the LiH calculations and the reference state is a superposition of six Slater determinants, each defining a Hartree-Fock-like configuration with single-particle excitations. Here, the simulated ϵ_{noise} and defined δ values were set smaller than those for LiH in the main text since there is a near-degeneracy between two of the lowest eigenstates, requiring less severe SVD truncation in order to accurately resolve the spectrum.

4. Molecular Nitrogen

As a third chemical example, we evaluate the performance of MODMD on the molecular nitrogen (N_2)

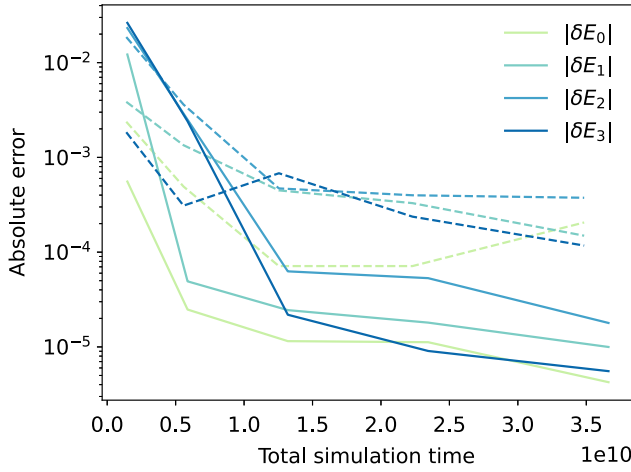


FIG. 16. Comparison of multiobservable dynamic mode decomposition and quantum multiple eigenvalue Gaussian filtered search on the transverse-field Ising model. Absolute errors in the first four eigenenergies (obtained via quantum multiple eigenvalue Gaussian filtered search and multiobservable dynamic mode decomposition with $I = 7$ distinct observables) are shown with respect to total simulation time for each algorithm. Multiobservable dynamic mode decomposition errors (solid) are averaged over a total of 20 trials, each involving a Gaussian noise realization and a set of $I - 1$ observables randomly selected within a pool of candidate operators. quantum multiple eigenvalue Gaussian filtered search errors (dashed) are calculated using the hyperparameter $(\alpha_{\text{QMEGS}}, q_{\text{QMEGS}}) = (5, 0.05)$. We search for a total of 20 peaks in quantum multiple eigenvalue Gaussian filtered search, where the four lowest-lying eigenstates have the 1st, 7th, 10th, and 5th highest initial state overlap respectively.

Hamiltonian. The 14-qubit Hamiltonian is generated in an equilibrium geometry with a bond length of $\approx 1.10 \text{ \AA}$ using the STO-3G basis set and parity fermion-to-qubit mapping. The observables were chosen following a procedure similar to that used in the LiH calculations and the reference state is a superposition of six Slater determinants, each defining a Hartree-Fock-like configuration with single-particle excitations. Here, the simulated $\varepsilon_{\text{noise}}$ and defined $\tilde{\delta}$ values were set smaller than those for LiH in Sec. V, due to a near-degeneracy between two of the lowest eigenstates.

APPENDIX H: PARAMETER SWEEPS

1. MODMD: Lithium Hydride Molecule

Here we present additional results for the eigenenergy convergence of the LiH molecule. We examine

the performance of MODMD under different parameter regimes upon sweeping the time step Δt against the size of the observable pool I in one set of calculations [Fig. 14(a)] and the singular value threshold $\tilde{\delta}$ against I in another [Fig. 14(b)].

As discussed in Sec. III C, we expect that an excessively large time step will cause an undesired aliasing ambiguity in eigenphases, independent of the number of observables, as is evident in the edge behavior of Fig. 14(a). Conversely, an overly small time step also slows down the convergence. We find that using a richer subspace, constructed from our multiobservable signal, can overcome this limitation. For all four eigenenergies, increasing the number of observables generally leads to reliable convergence in a larger range of Δt values, particularly for the excited-state energies. This robustness can be advantageous in an experimental setting where the achieved evolution time is relatively imprecise. Additionally, a smaller time step reduces the overall Trotter error [97].

As mentioned in Sec. III C, using a $\tilde{\delta}$ value that is too small (for instance, below the noise level) retains excessive noise, preventing proper eigenenergy convergence regardless of how many observables we employ. Conversely, a $\tilde{\delta}$ value that is too large truncates too much of the information recovered from MODMD measurements. This, however, can be overcome to some extent by increasing the number of observables in the observable pool. We observe that increasing I generally increases the range of $\tilde{\delta}$ values at which the lowest four eigenenergies achieve good convergence. Measuring several observables thus allows for a harsher truncation (larger $\tilde{\delta}$) in the postprocessing stage, better ensuring that the algorithm (i) does not fall into excessive noise retention regime and (ii) is less sensitive to hyperparameter tuning.

2. QMEGS: Transverse-Field Ising Model

QMEGS features two hyperparameters in its classical search-and-block postprocessing: the block size α_{QMEGS} and the search resolution q_{QMEGS} . Following [11], we adopt $q_{\text{QMEGS}} < \alpha_{\text{QMEGS}} \ll t_{\text{max}}$. To maximize performance, we also conduct a systematic hyperparameter sweep around the values reported in the original QMEGS work, $\alpha_{\text{QMEGS}} = 5$ and $q_{\text{QMEGS}} = 0.05$. The optimal pair is selected by minimizing the total QMEGS error defined as the sum of errors over the four lowest energy levels and simulations of increasing maximal runtime,

$$(\alpha_{\text{QMEGS}}^*, q_{\text{QMEGS}}^*) = \arg \min_{(\alpha_{\text{QMEGS}}, q_{\text{QMEGS}})} \sum_{t_{\text{max}}} \sum_{n=0}^{N_{\text{eig}}-1} |\delta E_n(t_{\text{max}}, \alpha_{\text{QMEGS}}, q_{\text{QMEGS}})|, \quad (\text{H1})$$

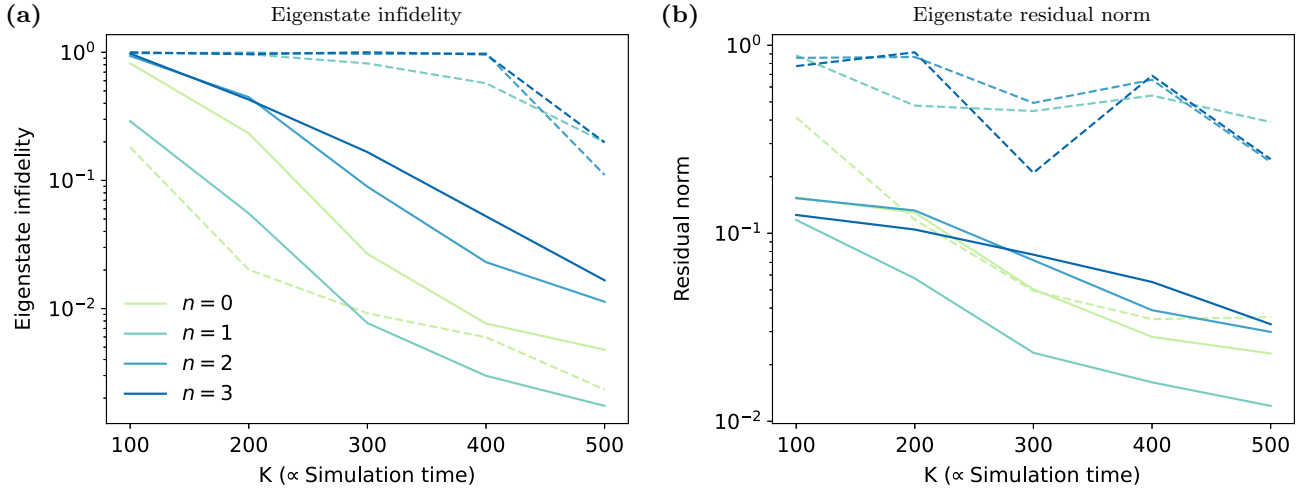


FIG. 17. Convergence of eigenstates for the transverse-field Ising model. To obtain eigenstate estimates $\tilde{\psi}_n$, we fix the multiobservable dynamic mode decomposition parameters $K/d = \frac{5}{2}$ and $\tilde{\delta} = 10^{-3}$ for constructing and thresholding the pair of data matrices $\mathbf{X}, \mathbf{X}' \in \mathbb{R}^{d \times (K+1)}$. Gaussian $\mathcal{N}(0, \varepsilon_{\text{noise}}^2)$ noise with $\varepsilon_{\text{noise}} = 10^{-4}$ is added independently to the real or/and imaginary parts of the matrix elements. The reference state $|\phi_0\rangle$ is an equal superposition of six computational basis states and we use a time step of $\Delta t \approx 0.08$. The state infidelity $1 - |\langle \tilde{\psi}_n | \psi_n \rangle|^2$ and normalized residual norm $\|H |\tilde{\psi}_n\rangle - \tilde{E}_n |\tilde{\psi}_n\rangle\|_2 / \|H\|_2$ are shown with respect to K proportional to the nondimensional maximal simulation time. Convergence results are averaged over 20 trials of Gaussian noise realizations and (for multiobservable dynamic mode decomposition) random 1-local Pauli observables. The solid lines depict the results for the multiobservable (multiobservable dynamic mode decomposition) algorithm with $I = 7$ and the dashed lines depict the results for the single-observable (observable dynamic mode decomposition) algorithm. (a) Approximate eigenstate fidelity $1 - |\langle \tilde{\psi}_n | \psi_n \rangle|^2$. (b) Normalized residual norm $\|H |\tilde{\psi}_n\rangle - \tilde{E}_n |\tilde{\psi}_n\rangle\|_2 / \|H\|_2$.

where Fig. 15 displays the results of this hyperparameter sweep, including relative performance plots similar to Fig. 6 for varying values of α_{QMEGS} and q_{QMEGS} . As

explained in Sec. V C 1, the error δE_n given in Eq. (H1) is optimized over the set of observables $\{O_i\}_{i=1}^I$ such that $|\delta E_n^{\text{QMEGS}}| = \min_{1 \leq i \leq I} |\delta E_n^{\text{QMEGS}}(O_i)|$.

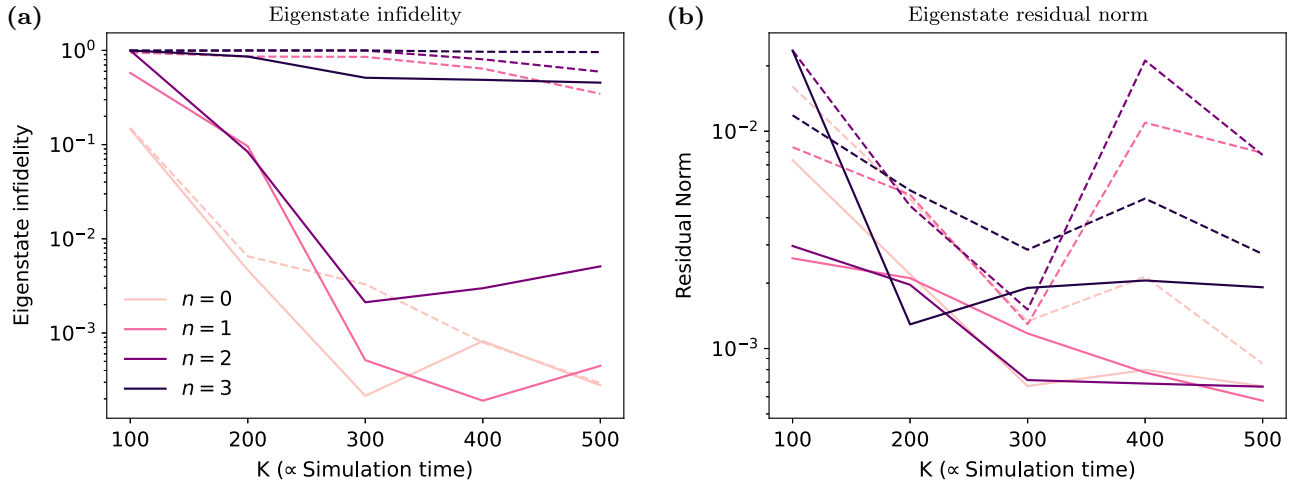


FIG. 18. Convergence of eigenstates for the lithium hydride molecule. To obtain eigenstate estimates $\tilde{\psi}_n$, we fix the multiobservable dynamic mode decomposition parameters $K/d = \frac{5}{2}$ and $\tilde{\delta} = 10^{-2}$ for constructing and thresholding the pair of data matrices $\mathbf{X}, \mathbf{X}' \in \mathbb{R}^{d \times (K+1)}$. Gaussian $\mathcal{N}(0, \varepsilon_{\text{noise}}^2)$ noise with $\varepsilon_{\text{noise}} = 10^{-3}$ is added independently to the real or/and imaginary parts of the matrix elements. The reference state $|\phi_0\rangle$ is a superposition of six computational basis states and we employ a time step of $\Delta t \approx 0.39$. The state infidelity, $1 - |\langle \tilde{\psi}_n | \psi_n \rangle|^2$, and normalized residual norm, $\|H |\tilde{\psi}_n\rangle - \tilde{E}_n |\tilde{\psi}_n\rangle\|_2 / \|H\|_2$, are shown with respect to K proportional to the nondimensional maximal simulation time. Convergence results are averaged over 20 trials of Gaussian noise realizations. The solid lines depict the results for the multiobservable (multiobservable dynamic mode decomposition) algorithm with $I = 7$ and the dashed lines depict the results for the single-observable (observable dynamic mode decomposition) algorithm. (a) Approximate eigenstate fidelity $1 - |\langle \tilde{\psi}_n | \psi_n \rangle|^2$. (b) Normalized residual norm $\|H |\tilde{\psi}_n\rangle - \tilde{E}_n |\tilde{\psi}_n\rangle\|_2 / \|H\|_2$.

Fixing the pair $(\alpha_{\text{QMEGS}}^*, q_{\text{QMEGS}}^*)$, we also examine the energy error with respect to the total runtime defined as

$$t_{\text{tot}}^{\text{QMEGS}} = \sum_{k=0}^{K_{\text{QMEGS}}-1} |t_k^{\text{QMEGS}}|, \quad t_{\text{tot}}^{\text{MODMD}} = \sum_{k=0}^{K+d} |t_k^{\text{MODMD}}| = \sum_{k=0}^{K+d} \frac{k\Delta t}{\epsilon^2} = \frac{(K+d)(K+d+1)}{2\epsilon^2}, \quad (\text{H2})$$

as shown in Fig. 16.

APPENDIX I: EIGENSTATE CONVERGENCE

As discussed in Sec. III B, the MODMD algorithm can approximate the low-lying eigenstates of the Hamiltonian in addition to the eigenenergies via Eq. (18). Here we demonstrate the convergence of the TFIM and LiH eigenstates with respect to the number of time steps K . We note that both the noise and SVD threshold in the TFIM simulation were set differently (smaller by an order of magnitude) compared to those in Sec. V (Figs. 17 and 18).

Here, we use two metrics to evaluate the quality of our eigenstate estimates $|\tilde{\psi}_n\rangle$, which are obtained via MODMD as linear combinations of time-evolved states acted on by selected operators. First, we examine the infidelity $1 - |\langle\tilde{\psi}_n|\psi_n\rangle|^2$ determined by the squared overlap of the approximate and exact eigenstates. Clearly, the (in)fidelity provides a direct measure of how well the approximate eigenstate represents the true wavefunction. As a second alternative, we consider the normalized residual norm, $\|H|\tilde{\psi}_n\rangle - \tilde{E}_n|\tilde{\psi}_n\rangle\|_2/\|H\|_2$, which is indicative of the energy variance of the approximate eigenstate and is commonly monitored in numerical linear algebra tasks to track eigenspace convergence.

In both metrics, MODMD captures most of the first three excited states with accuracies close to that of the ground state—a remarkable behavior not observed in OMDM. This uniformity in performance across the low-lying eigenstates underscores the effectiveness of our multiobservable approach in stabilizing recovery of essential spectral information.

-
- [1] R. M. Parrish, and P. L. McMahon, Quantum filter diagonalization: Quantum eigendecomposition without full quantum phase estimation, [arXiv:1909.08925](#).
 - [2] N. H. Stair, R. Huang, and F. A. Evangelista, A multireference quantum Krylov algorithm for strongly correlated electrons, *J. Chem. Theory Comput.* **16**, 2236 (2020).
 - [3] K. Klymko, C. Mejuto-Zaera, S. J. Cotton, F. Wudarski, M. Urbanek, D. Hait, M. Head-Gordon, K. B. Whaley, J. Moussa, N. Wiebe, *et al.*, Real-time evolution for ultracompact Hamiltonian eigenstates on quantum hardware, *PRX Quantum* **3**, 020323 (2022).
 - [4] Y. Shen, K. Klymko, J. Sud, D. B. Williams-Young, W. A. de Jong, and N. M. Tubman, Real-time Krylov theory for quantum computing algorithms, *Quantum* **7**, 1066 (2023).

- [5] C. L. Cortes, and S. K. Gray, Quantum Krylov subspace algorithms for ground- and excited-state energy estimation, *Phys. Rev. A* **105**, 022417 (2022).
- [6] N. H. Stair, C. L. Cortes, R. M. Parrish, J. Cohn, and M. Motta, Stochastic quantum Krylov protocol with double-factorized Hamiltonians, *Phys. Rev. A* **107**, 032414 (2023).
- [7] Z. Ding, and L. Lin, Even shorter quantum circuit for phase estimation on early fault-tolerant quantum computers with applications to ground-state energy estimation, *PRX Quantum* **4**, 020331 (2023).
- [8] Z. Ding, and L. Lin, Simultaneous estimation of multiple eigenvalues with short-depth quantum circuit on early fault-tolerant quantum computers, *Quantum* **7**, 1136 (2023).
- [9] Y. Shen, D. Camps, A. Szasz, S. Darbha, K. Klymko, D. B. Williams-Young, N. M. Tubman, and R. V. Beeumen, Estimating eigenenergies from quantum dynamics: A unified noise-resilient measurement-driven approach, *Quantum* **9**, 1836 (2025).
- [10] G. Wang, D. S. França, R. Zhang, S. Zhu, and P. D. Johnson, Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision, *Quantum* **7**, 1167 (2023).
- [11] Z. Ding, H. Li, L. Lin, H. Ni, L. Ying, and R. Zhang, Quantum multiple eigenvalue Gaussian filtered search: An efficient and versatile quantum phase estimation method, *Quantum* **8**, 1487 (2024).
- [12] Z. Zhang, A. Wang, X. Xu, and Y. Li, Measurement-efficient quantum Krylov subspace diagonalisation, *Quantum* **8**, 1438 (2024).
- [13] M. Motta, W. Kirby, I. Liepuoniute, K. J. Sung, J. Cohn, A. Mezzacapo, K. Klymko, N. Nguyen, N. Yoshioka, and J. E. Rice, Subspace methods for electronic structure simulations on quantum computers, *Electron. Struct.* **6**, 013001 (2024).
- [14] N. Maskara, S. Ostermann, J. Shee, M. Kalinowski, A. McClain Gomez, R. Araiza Bravo, D. S. Wang, A. I. Krylov, N. Y. Yao, M. Head-Gordon, M. D. Lukin, and S. F. Yelin, Programmable simulations of molecules and materials with reconfigurable quantum processors, *Nat. Phys.* **21**, 289 (2025).
- [15] M. R. Wall, and D. Neuhauser, Extraction, through filter-diagonalization, of general quantum eigenvalues or classical normal mode frequencies from small number of residues or a short-time segment of a signal. I. Theory and application to a quantum-dynamics model, *J. Chem. Phys.* **102**, 8011 (1995).
- [16] E. N. Epperly, L. Lin, and Y. Nakatsukasa, A theory of quantum subspace diagonalization, *SIAM J. Matrix Anal. Appl.* **43**, 1263 (2022).

- [17] H. Li, H. Ni, and L. Ying, Adaptive low-depth quantum algorithms for robust multiple-phase estimation, *Phys. Rev. A* **108**, 062408 (2023).
- [18] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, in *Graduate Studies in Mathematics* (American Mathematical Society (AMS), Boston, MA, USA, 2002).
- [19] M. A. Nielsen, and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, New York, NY, USA, 2010).
- [20] R. B. Griffiths, and C.-S. Niu, Semiclassical Fourier transform for quantum computation, *Phys. Rev. Lett.* **76**, 3228 (1996).
- [21] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Entanglement-free Heisenberg-limited phase estimation, *Nature* **450**, 393 (2007).
- [22] D. W. Berry, B. L. Higgins, S. D. Bartlett, M. W. Mitchell, G. J. Pryde, and H. M. Wiseman, How to perform the most accurate possible phase measurements, *Phys. Rev. A* **80**, 052114 (2009).
- [23] M. Zwiernik, C. A. Pérez-Delgado, and P. Kok, General optimality of the Heisenberg limit for quantum metrology, *Phys. Rev. Lett.* **105**, 180402 (2010).
- [24] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, *Nat. Photonics* **5**, 222 (2011).
- [25] Z. Ding, E. N. Epperly, L. Lin, and R. Zhang, in *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE Computer Society, Los Alamitos, CA, USA, 2024), pp. 2344–2366.
- [26] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, *Nat. Commun.* **9**, 4812 (2018).
- [27] A. Arrasmith, Z. Holmes, M. Cerezo, and P. J. Coles, Equivalence of quantum barren plateaus to cost concentration and narrow gorges, *Quantum Sci. Technol.* **7**, 045015 (2022).
- [28] H. Qi, L. Wang, H. Zhu, A. Gani, and C. Gong, The barren plateaus of quantum neural networks: Review, taxonomy and trends, *Quantum Inf. Process.* **22**, 435 (2023).
- [29] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, *Nat. Phys.* **16**, 1050 (2020).
- [30] H.-Y. Hu, and Y.-Z. You, Hamiltonian-driven shadow tomography of quantum states, *Phys. Rev. Res.* **4**, 013054 (2022).
- [31] A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, and P. Zoller, The randomized measurement toolbox, *Nat. Rev. Phys.* **5**, 9 (2023).
- [32] C. Bertoni, J. Haferkamp, M. Hinsche, M. Ioannou, J. Eisert, and H. Pashayan, Shallow shadows: Expectation estimation using low-depth random Clifford circuits, *Phys. Rev. Lett.* **133**, 020602 (2024).
- [33] K. Van Kirk, J. Cotler, H.-Y. Huang, and M. D. Lukin, Hardware-efficient learning of quantum many-body states, *arXiv:2212.06084*.
- [34] A. Zhao, N. C. Rubin, and A. Miyake, Fermionic partial tomography via classical shadows, *Phys. Rev. Lett.* **127**, 110504 (2021).
- [35] Z. Liu, Z. Hao, and H.-Y. Hu, Predicting arbitrary state properties from single Hamiltonian quench dynamics, *Phys. Rev. Res.* **6**, 043118 (2024).
- [36] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Quantum algorithms revisited, *Proc. R. Soc. Lond. A* **454**, 339 (1998).
- [37] 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* **567**, 209 (2019).
- [38] H.-Y. Hu, S. Choi, and Y.-Z. You, Classical shadow tomography with locally scrambled quantum dynamics, *Phys. Rev. Res.* **5**, 023027 (2023).
- [39] A. A. Akhtar, H.-Y. Hu, and Y.-Z. You, Scalable and flexible classical shadow tomography with tensor networks, *Quantum* **7**, 1026 (2023).
- [40] H.-Y. Hu, A. Gu, S. Majumder, H. Ren, Y. Zhang, D. S. Wang, Y.-Z. You, Z. Mineev, S. F. Yelin, and A. Seif, Demonstration of robust and efficient quantum property learning with shallow shadows, *Nat. Commun.* **16**, 2943 (2025).
- [41] H. H. S. Chan, R. Meister, M. L. Goh, and B. Koczor, Algorithmic shadow spectroscopy, *PRX Quantum* **6**, 010352 (2025).
- [42] A. Moitra, in *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*, STOC '15 (Association for Computing Machinery, New York, NY, USA, 2015), pp. 821–830.
- [43] W. Li, W. Liao, and A. Fannjiang, Super-resolution limit of the ESPRIT algorithm, *IEEE Trans. Inform. Theory* **66**, 4593 (2020).
- [44] M. Ippoliti, Y. Li, T. Rakovszky, and V. Khemani, Operator relaxation and the optimal depth of classical shadows, *Phys. Rev. Lett.* **130**, 230403 (2023).
- [45] T. Schuster, J. Haferkamp, and H.-Y. Huang, Random unitaries in extremely low depth, *Science* **389**, 92 (2025).
- [46] Y. Dong, L. Lin, and Y. Tong, Ground-state preparation and energy estimation on early fault-tolerant quantum Computers via quantum eigenvalue transformation of unitary matrices, *PRX Quantum* **3**, 040305 (2022).
- [47] D. Potts, M. Tasche, and T. Volkmer, Efficient spectral estimation by MUSIC and ESPRIT with application to sparse FFT, *Front. Appl. Math. Stat.* **2**, 1 (2016).
- [48] N. Gomes, J. Yin, S. Niu, C. Yang, and W. A. de Jong, A hybrid method for quantum dynamics simulation, *arXiv:2307.15231*.
- [49] A. F. Kemper, C. Yang, and E. Gull, Denoising and extension of response functions in the time domain, *Phys. Rev. Lett.* **132**, 160403 (2024).
- [50] R. Kaneko, M. Imada, Y. Kabashima, and T. Ohtsuki, Forecasting long-time dynamics in quantum many-body systems by dynamic mode decomposition, *Phys. Rev. Res.* **7**, 013085 (2025).
- [51] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, NY, USA, 1987).
- [52] D. Limmer, *Statistical Mechanics and Stochastic Thermodynamics: A Textbook on Modern Approaches in and Out of Equilibrium*, Oxford Graduate Texts (Oxford University Press, 2024).
- [53] B. O. Koopman, Hamiltonian systems and transformation in Hilbert space, *Proc. Natl. Acad. Sci.* **17**, 315 (1931).
- [54] G. D. Birkhoff, and B. O. Koopman, Recent contributions to the ergodic theory, *Proc. Natl. Acad. Sci.* **18**, 279 (1932).

- [55] S. L. Brunton, B. W. Brunton, J. L. Proctor, and J. N. Kutz, Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control, *PloS One* **11**, e0150171 (2016).
- [56] 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).
- [57] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, *Proc. Natl. Acad. Sci.* **113**, 3932 (2016).
- [58] H. Arbabi, and I. Mezić, Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the Koopman operator, *SIAM J. Appl. Dyn. Syst.* **16**, 2096 (2017).
- [59] G. R. de Prony, Essai expérimental et analytique sur les lois de la dilatabilité des fluides élastiques, et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alcool, à différentes températures, *Journal de l'École polytechnique* **1**, 24 (1795).
- [60] C. C. Paige, The computation of eigenvalues and eigenvectors of very large sparse matrices (Doctoral dissertation, University of London, 1971).
- [61] Y. Saad, On the rates of convergence of the Lanczos and the Block-Lanczos methods, *SIAM J. Numer. Anal.* **17**, 687 (1980).
- [62] R. A. Horn, and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, New York, NY, USA, 1985).
- [63] $\lim_{K \rightarrow \infty} \frac{1}{K+1} \sum_{k=0}^K f(e^{-iHk\Delta t} |\phi_0\rangle)$ converges uniformly to $\int_{\mathcal{H}} d\bar{\mu} f$ for all $|\phi_0\rangle \in \mathcal{H}$, provided that the limit measure $\bar{\mu}$ is uniquely invariant. For quantum systems, the action of dynamical evolution can be regarded as a collection of independent rotations up to a change of basis. If the Hamiltonian spectrum and evolution time step remain general, $\bar{\mu}$ is then the uniform measure on $\mathcal{H}$ as the unique invariant measure.
- [64] Y. Saad, *Numerical Methods for Large Eigenvalue Problems* (Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2011).
- [65] Y. Saad, and M. H. Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM J. Sci. Stat. Comput.* **7**, 856 (1986).
- [66] J. Liesen, and P. Tichý, Convergence analysis of Krylov subspace methods, *GAMM-Mitt.* **27**, 153 (2004).
- [67] R.-C. Li, and L.-H. Zhang, Convergence of the block Lanczos method for eigenvalue clusters, *Numer. Math.* **131**, 83 (2015).
- [68] M. Suzuki, General theory of fractal path integrals with applications to many-body theories and statistical physics, *J. Math. Phys.* **32**, 400 (1991).
- [69] N. Hatano, and M. Suzuki, in *Quantum Annealing and Other Optimization Methods*, edited by A. Das and B. K. Chakrabarti (Springer, Berlin, 2005), pp. 37–68.
- [70] A. M. Childs, and N. Wiebe, Hamiltonian simulation using linear combinations of unitary operations, *Quantum Info. Comput.* **12**, 901 (2012).
- [71] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders, Efficient quantum algorithms for simulating sparse Hamiltonians, *Commun. Math. Phys.* **270**, 359 (2007).
- [72] For a given problem, the scrambling unitaries required in classical shadow tomography have finite depth, as low as logarithmic in the system size [40], for each independent circuit run. So the gate count associated with the classical shadow component is $\mathcal{O}(t_{\text{tot}}Q) = \mathcal{O}(\log(I)k_{\text{max}}^4\epsilon_1^{-2})$ and is therefore subleading.
- [73] P.-Å. Wedin, Perturbation theory for pseudo-inverses, *BIT Numer. Math.* **13**, 217 (1973).
- [74] G. W. Stewart, and J.-g. Sun, *Matrix Perturbation Theory*, Computer science and scientific computing (Academic Press, Boston, 1990).
- [75] A. Y. Kitaev, Quantum computations: Algorithms and error correction, *Russ. Math. Surv.* **52**, 1191 (1997).
- [76] R. Bhatia, *Matrix Analysis* (Springer, New York, NY, USA, 1997), Vol. 169.
- [77] J. T. Seeley, M. J. Richard, and P. J. Love, The Bravyi-Kitaev transformation for quantum computation of electronic structure, *J. Chem. Phys.* **137**, 224109 (2012).
- [78] W. Kirby, Analysis of quantum Krylov algorithms with errors, *Quantum* **8**, 1457 (2024).
- [79] <https://github.com/alexbuzalicse/modmd> (2025).
- [80] Y. Nakata, C. Hirche, M. Koashi, and A. Winter, Efficient quantum pseudorandomness with nearly time-independent Hamiltonian dynamics, *Phys. Rev. X* **7**, 021006 (2017).
- [81] E. Kökcü, D. Camps, L. Bassman Oftelie, J. K. Freericks, W. A. de Jong, R. Van Beeumen, and A. F. Kemper, Algebraic compression of quantum circuits for Hamiltonian evolution, *Phys. Rev. A* **105**, 032420 (2022).
- [82] D. Camps, E. Kökcü, L. Bassman Oftelie, W. A. de Jong, A. F. Kemper, and R. Van Beeumen, An algebraic quantum circuit compression algorithm for Hamiltonian simulation, *SIAM J. Matrix Anal. Appl.* **43**, 1084 (2022).
- [83] I. Mezić, and A. Banaszuk, Comparison of systems with complex behavior, *Physica D: Nonlinear Phenom.* **197**, 101 (2004).
- [84] I. Mezić, Spectral properties of dynamical systems, model reduction and decompositions, *Nonlinear Dyn.* **41**, 309 (2005).
- [85] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, Spectral analysis of nonlinear flows, *J. Fluid Mech.* **641**, 115 (2009).
- [86] P. J. Schmid, Dynamic mode decomposition of numerical and experimental data, *J. Fluid Mech.* **656**, 5 (2010).
- [87] I. Mezić, Analysis of fluid flows via spectral properties of the Koopman operator, *Annu. Rev. Fluid Mech.* **45**, 357 (2013).
- [88] While “observable” is typically used in quantum mechanics to refer specifically to Hermitian operators (with real expectation value), here we use a broader definition, encompassing also complex scalar quantities that can be computed from measurements on a quantum computer.
- [89] D. Ruelle, and F. Takens, On the nature of turbulence, *Commun. Math. Phys.* **20**, 167 (1971).
- [90] F. Takens, in *Lecture Notes in Mathematics* (Springer, Berlin, 1981), pp. 366–381.
- [91] I. L. Gutiérrez, F. Dietrich, and C. B. Mendl, Quantum process tomography of unitary maps from time-delayed measurements, *Quant. Inf. Proc.* **22**, 251 (2023).
- [92] N. I. Akhiezer, Über einige funktionen die in gegebenen Intervallen am wenigsten von null abweichen, *Nachr. Phys-Math. Univ. Kazan* **3**, 1 (1928).

- [93] K. Schiefermayr, Chebyshev polynomials on circular arcs, [Acta Sci. Math.](#) **85**, 629 (2019).
- [94] C. Schmid, and K. J. DeMars, Angular correlation using Rogers-Szegő-Chaos, [Mathematics](#) **8**, 171 (2020).
- [95] M. Foupouagnigni, and W. Koepf, in *Orthogonal Polynomials: 2nd AIMS-Volkswagen Stiftung Workshop, Douala, Cameroon, 5–12 October, 2018* (Birkhäuser, Cham, 2020).
- [96] D. Zgid, E. Gull, and G. K.-L. Chan, Truncated configuration interaction expansions as solvers for correlated quantum impurity models and dynamical mean-field theory, [Phys. Rev. B](#) **86**, 165128 (2012).
- [97] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu, Theory of Trotter error with commutator scaling, [Phys. Rev. X](#) **11**, 011020 (2021).