Estimating Eigenenergies from Quantum Dynamics: A Unified Noise-Resilient Measurement-Driven Approach

Yizhi Shen[1], Daan Camps[2], Aaron Szasz[1], Siva Darbha[1,2], Katherine Klymko[2], David B. Williams–Young[1], Norm M. Tubman[3], and Roel Van Beeumen[1]

[1]Applied Mathematics and Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[2]National Energy Research Scientific Computing Center, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[3]NASA Ames Research Center, Moffett Field, CA 94035, USA

arXiv:2306.01858v5 [quant-ph] 13 Aug 2025

Ground state energy estimation in physical, chemical, and materials sciences is one of the most promising applications of quantum computing. In this work, we introduce a new hybrid approach that finds the eigenenergies by collecting real-time measurements and post-processing them using the machinery of dynamic mode decomposition (DMD). From the perspective of quantum dynamics, we establish that our approach can be formally understood as a stable variational method on the function space of observables available from a quantum many-body system. We also provide strong theoretical and numerical evidence that our method converges rapidly even in the presence of a large degree of perturbative noise, and show that the method bears an isomorphism to robust matrix factorization methods developed independently across various scientific communities. Our numerical benchmarks on spin and molecular systems demonstrate an accelerated convergence and a favorable resource reduction over state-of-the-art algorithms. The DMD-centric strategy can systematically mitigate noise and stands out as a leading hybrid quantum-classical eigensolver.

1 Introduction

In domains ranging from fundamental physics to practical chemistry, the characterization of many-body quantum systems requires the calculation of low-energy states, especially ground states. By identifying ground states of lattice models describing crystalline materials, we can look for novel phases of matter such as quantum spin liquids [1, 2] and topological insulators [3, 4], or even understand the mechanisms driving high-temperature superconductivity [5–7]. By finding the low-lying eigenstates of molecules, we can uncover reaction pathways and design optimized processes in various sectors from fertilizer production [8] to drug discovery [9].

Yizhi Shen: yizhis@lbl.gov

For strongly interacting many-body systems, quantum computers provide arguably the best tool for finding ground states. Naive classical approaches based on exact diagonalization of the many-body Hamiltonian scale exponentially in the number of quantum spins and/or electronic orbitals, while more sophisticated approaches such as tensor network simulations [10, 11] are inefficient when describing highly-entangled or high-dimensional systems. Simulation using a quantum computer can overcome these challenges, scaling polynomially in the number of spins or orbitals and faithfully representing states with any degree of entanglement.

With a fully fault-tolerant, error-corrected quantum computer, accurate ground state energies can be determined via the well-established quantum phase estimation (QPE) [12–14] algorithm. QPE is theoretically optimal: it saturates the Heisenberg limit given a single copy of the input reference state, taking only $\mathcal{O}(\epsilon^{-1})$ computational time to ascertain precision ϵ. However, QPE requires the use of $\mathcal{O}(\log(\epsilon^{-1}))$ ancilla qubits, making this algorithmic primitive infeasible without full fault tolerance.

In contrast, hybrid quantum-classical algorithms may bring a quantum advantage even on near-term devices. Hybrid strategies deploy a quantum computer for the tasks where it excels, for example evolving a wavefunction with unitary operators, and offload other parts of a simulation to a classical computer. Notably, the variational quantum eigensolver (VQE) [15–18] has emerged as a popular approach for current quantum hardware. However, the VQE method often relies on the heuristic selection of circuit ansatz, and can be prone to serious optimization challenges including vanishing gradients [19–21]. As such, the potential computational advantages of VQE are not clearly understood.

A promising alternative family of algorithms only requires access to measurements of state overlaps using a single ancilla qubit. These algorithms are known as early fault-tolerant because they allocate fewer qubits than full phase estimation but necessitate multiple copies of the input reference state. State-of-the-art examples within this category include variational

quantum phase estimation (VQPE) [22, 23], quantum complex exponential least-squares (QCELS) [24, 25], and quantum Lanczos [26], among others. Although these algorithms have promising potential to achieve Heisenberg scaling before full fault tolerance, they also come with their own limitations. For example, some algorithms require a substantial number of shots to reduce the statistical errors. Additionally, due to the prohibitive cost of state orthogonalization, recent attention has focused on algorithms with non-orthogonal bases [27, 28], resulting in ill-conditioned generalized eigenvalue problems. This difficulty is exacerbated by the presence of noise sources, even when accounting solely for the statistical noise from sampling on any practical hardware. Other algorithms build on the premise of a good input state that has significant overlap with the low energy eigenstates, which imposes serious extra resource overhead for initial state preparation for some problems.

In this paper, we introduce a new hybrid algorithm for finding ground state energies and other ground state properties based on the machinery of dynamic mode decomposition (DMD) [29–33]. DMD processes snapshots of a dynamical system to identify the principal oscillating modes and constructs an efficient representation of the full dynamical trajectory. To apply DMD to a quantum system, where we do not have access to the full state vector, we use a vector of time-delayed measured observables as a proxy. In particular, the oscillating modes of a quantum system are its energy eigenstates: if we use expectation values of the time evolution operator as the observables, we expect DMD to return the ground state energy of the system. We name this method observable dynamic mode decomposition (ODMD). By fully exploiting real-time data from quantum hardware, our approach introduces a unified eigensolver framework that preserves the salient features of different classes of approaches, including variational convergence, compactness, and noise resilience.

Similar to state-of-the-art approaches such as VQPE and QCELS, ODMD only requires matrix elements of the time evolution operator, computed via state overlaps, and at most one ancilla qubit to find the ground state energy. However, ODMD can overcome some limitations of those approaches and other hybrid algorithms. On the one hand, in contrast to subspace projection techniques such as VQPE, the classical portion of our algorithm involves a least-squares (LS) problem rather than a generalized eigenvalue problem, improving conditioning and stability against perturbative noise. More importantly, ODMD can be formulated as a functional subspace method applied to the space of observables in our setting. This connection can be established via Koopman operator analysis [34–38], which allows us to further generalize ODMD to a multi-observable framework for far more efficient excited state computation.

On the other hand, while ODMD shares the frequency retrieval perspective underlying QCELS and its variants, ODMD distinguishes from these recent signal processing algorithms in the following crucial aspects. First, ODMD does not hinge on the requirement that the reference state must have a dominant overlap with specific eigenstates of interest. For other signal processing algorithms targeting ground state energy, this dominance condition translates into a restricted ground state overlap greater than $\frac{1}{2}$ [24, 25, 39], which imposes an extra quantum cost for state preparation. Second, the classical cost for ODMD scales polynomially with the number of target eigenenergies due to its matrix-based post-processing. This is exponentially more favorable than that of the nonlinear approaches [24, 25]. We note that although the sparse Fourier approach of [39] also exponentially and moreover optimally reduces such cost, it fails to provide accurate energy estimates when the reference state evolves inexactly, for example in the presence of a common depolarizing noise channel. This is because damping or any deviation from a purely sinusoidal time signal can severely limit the resolution of frequencies achievable through standard Fourier recovery. Third, ODMD reveals information about the eigenstates, whereas alternative signal processing approaches recover only the eigenenergies.

Other benefits of ODMD include the ability to use only the real parts of the overlap matrix elements, therefore reducing the number of shots by half. Furthermore, our method progressively updates the ground state estimate (or the target eigenstate estimates in general) with the addition of each new matrix element, so resources can be utilized efficiently to achieve the desired level of precision.

Our manuscript is organized as follows: In Section 2, we give an overview of existing real-time evolution-based approaches to finding ground state energy, and we discuss their relative strengths and weaknesses. We introduce in Section 3 a new measurement-driven framework, which we name the observable dynamic mode decomposition (ODMD), starting from the machinery of classic dynamic mode decomposition (DMD). A rigorous examination of the ODMD convergence is found in Section 4. For readers interested in the theoretical underpinnings of ODMD, in Section 5 we show how the method can be reinterpreted in terms of subspace approaches in function space and in terms of matrix factorization methods from the applied mathematics literature; readers who are primarily interested in ODMD as a practical tool could skip this technical section on a first read. In Section 6, we numerically demonstrate the efficacy of our approach for ground and excited state problems using representative spin and molecular many-body systems. Section 7 provides a discussion and concluding remarks.

2 Background and landscape of quantum methods

There has been a growing interest in leveraging real-time methods for finding ground state energies [22, 40–43]. The primary appeal of these methods comes from the unitarity of the time evolution operator e^{-iHt}. Unitarity makes real-time based approaches naturally compatible with quantum hardware, which accordingly translates to shorter-depth circuits compared to other, naturally non-unitary methods such as the power method [44], Lanczos method [26], or imaginary-time implementations [45].

2.1 Real-time evolution-based algorithms for ground state estimation

Recent studies have proposed subspace methods [22, 23, 40–42], which project the many-body Schrödinger equation onto a subspace of the Hilbert space constructed through real-time evolution. These approaches yield generalized eigenvalue equations built from non-orthogonal real-time basis states. The subspace Hamiltonian and overlap matrix elements,

$$\mathbf{H}_{ij} = \langle \boldsymbol{\phi}_i | H | \boldsymbol{\phi}_j \rangle = \langle \boldsymbol{\phi}_0 | e^{iHt_i} H e^{-iHt_j} | \boldsymbol{\phi}_0 \rangle \,, \tag{1}$$

$$\mathbf{S}_{ij} = \langle \boldsymbol{\phi}_i | \boldsymbol{\phi}_j \rangle = \langle \boldsymbol{\phi}_0 | e^{iH(t_i - t_j)} | \boldsymbol{\phi}_0 \rangle \,, \tag{2}$$

are obtained via measurements performed on quantum hardware and the resulting projected eigenvalue problem, $\mathbf{H}\boldsymbol{\Psi} = \tilde{E}\mathbf{S}\boldsymbol{\Psi}$, is solved classically for approximate energies $\tilde{E}$. For example, quantum filter diagonalization (QFD) and the VQPE [22] method have this form with a uniform timestep grid $t_j = j\Delta t$. The multireference selected quantum Krylov (MRSQK) algorithm [43] is a variant where time evolution is applied to a set of initial states in order to reduce the condition number of the resulting generalized eigenvalue equation. As suggested by recent work [22, 23, 41], a relatively small number of real-time states $\{|\boldsymbol{\phi}_i\rangle\}_i$ suffice to achieve accurate estimates for the low-lying energies in many physically relevant cases, making these methods relevant for near-term calculations on quantum hardware.

A principal challenge in subspace methods is the classical solution of the generalized eigenvalue equation. Given that the matrix elements are measured on quantum hardware, they are inevitably subject to noise, which can become quite significant in practice. Due to the non-orthogonality of the real-time basis states, the noise leads to a highly ill-conditioned problem. One common approach to address ill-conditioned eigenvalue equations is to regularize or precondition the subspace $\mathrm{span}\{|\boldsymbol{\phi}_i\rangle\}_i$. For example, this can be done by pruning small singular values of the overlap matrix $\mathbf{S}$ beneath a noise-consistent threshold. However, such techniques are only viable up to a specific limit of the subspace dimension. Beyond this limit, noise accumulation in the matrix elements adversely affects the accuracy of the eigenvalue estimate as more real-time states are introduced [46].

In addition to explicit subspace diagonalization, an emerging class of approaches tackles the eigenvalue problem via a sparse Fourier perspective [47]. For example, QCELS [24] derives its inspiration from real-time signal processing techniques by singling out the dominant eigenphases through the optimization,

$$\tilde{E} = \arg\min_{\theta \in \mathbb{R}} \left\{ \min_{r \in \mathbb{C}} \sum_i \left| \langle \boldsymbol{\phi}_0 | e^{-iHt_i} | \boldsymbol{\phi}_0 \rangle - r e^{-i\theta t_i} \right|^2 \right\}, \tag{3}$$

where a classical minimization routine is called to search for the best amplitude and phase parameters (r, θ) fitting the quantum data. The fitted energy $\tilde{E}$ converges rapidly to the ground state energy if the chosen initial state $|\boldsymbol{\phi}_0\rangle$ is well-aligned with the exact many-body ground state.

Indeed, the effectiveness of many real-time approaches thus far has been mostly constrained to situations where an optimal starting or reference state $|\boldsymbol{\phi}_0\rangle$, which has a considerable overlap with the desired eigenstates, is achievable. In such scenarios, short time evolution suffices for convergence, and a basic truncation strategy can effectively remove perturbative noise and mitigate issues of ill-conditioning. However, creating starting states with substantial eigenstate overlap is difficult in general, especially for increasingly large and strongly correlated many-body systems.

2.2 Computing overlaps

The real-time methods described in Section 2.1 all involve calculating expectation values of the form $\langle \boldsymbol{\phi} | e^{-iHt} | \boldsymbol{\phi} \rangle$, where $|\boldsymbol{\phi}\rangle$ is some input state of interest and the expectation specifies the overlap between the input and time-evolved states. In particular, the methods do not require direct extraction of time-evolved quantum states in the high-dimensional Hilbert space, which is intractable for large systems owing to an exponential cost. In contrast, since the time evolution operator is often expressible on quantum hardware, the expectation value can be efficiently evaluated.

The time-evolved overlaps have the more general form $\langle \boldsymbol{\phi} | U | \boldsymbol{\phi} \rangle$, where U is a unitary operator. Such expectation values can be accessed using projective quantum measurements. In particular, they can be measured in the Pauli basis through the Hadamard test [48] or mirror-type fidelity estimation [42, 49]. The Hadamard test, illustrated in Fig. 1, evaluates the expectation using a single ancilla and controlled time evolution, $|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{-iHt}$. We measure the ancilla in the Pauli basis and record the measurement result as either $+1$ or -1. We perform repeated measurements and average over the $\{\pm 1\}$ results to

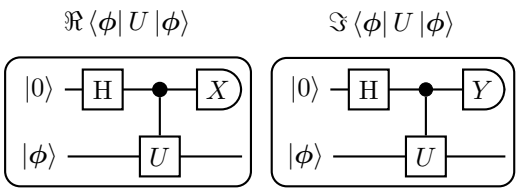

Figure 1: Hadamard test circuits for measuring the real part (**left**) and imaginary part (**right**) of the state overlap $\langle\phi|U|\phi\rangle \in \mathbb{C}$. A Hadamard gate is applied to an ancilla qubit initialized in the $|0\rangle$ state. The ancilla then acts as the control qubit of the unitary U operating on the register qubits. Ancilla measurements are performed in the Pauli basis to evaluate the expectation.

converge the expectation. When the ancilla measurements are performed in the Pauli X and Y bases, we acquire the real and imaginary parts of the complex expectation, respectively. The mirror circuits circumvent the need for controlled evolution, which can sometimes be expensive to implement on current hardware. Instead, they employ controlled state preparation followed by unrestricted time evolution.

The real-time framework we present in Section 3 relies on acquiring *system observables*, which we define to be scalar-valued (either real or complex) functions over the system state space. With a judicious selection of observables, we can effectively compress high-dimensional state information into lower-dimensional observable vectors, and therefore encode some of the spectral features of the underlying many-body system. The main system observables that we evaluate in our framework are the expectations $\langle\phi|e^{-iHt}|\phi\rangle$.

3 ODMD method

As we design methods tailored to accommodate noisy quantum data, we now look beyond directly solving ill-conditioned generalized eigenvalue problems. Instead, we introduce a new optimization-free algorithm that solves a least-squares problem to achieve a controllable stability against noise. We show that the real-time observables $\langle\phi|e^{-iHt}|\phi\rangle$ give natural probes of spectral and dynamical information of a many-body quantum system. Building upon the classical machinery of DMD, we develop a noise-resilient measurement-driven framework for estimating eigenenergies from quantum dynamics.

3.1 DMD to ODMD

The standard dynamic mode decomposition (DMD), originally developed within the field of numerical fluid dynamics, is a measurement-driven approximation for the temporal progression of a classical dynamical system [29–33]. Specifically, DMD samples snapshots of the system at some regular time intervals Δt and exploits them to construct an efficient representation

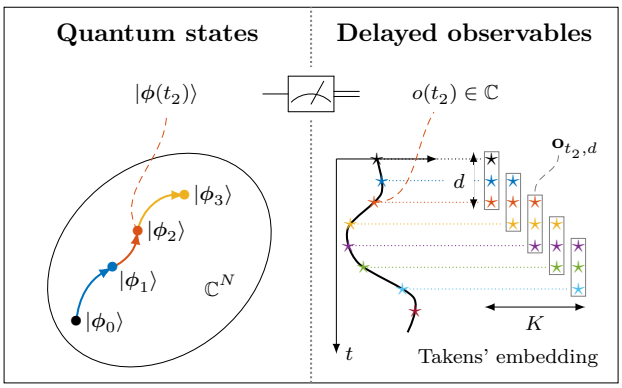

Figure 2: Takens' embedding of scalar observables in the quantum context. The left panel depicts the time evolution of a reference quantum state $|\phi_0\rangle$ in the many-body Hilbert space. Time-evolved states $|\phi_k\rangle = |\phi(t_k)\rangle$ are shown at a sequence of time steps $t_k = k\Delta t$. The right panel tracks the corresponding observables $o(t)$ at the same sequence of time steps, which can be evaluated efficiently on a quantum computer. Takens' embedding establishes a connection between the state evolution and time-delayed observables.

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 least-squares (LS) relation,

$$\boldsymbol{\phi}_{k+1} \overset{\text{LS}}{=} A_k \boldsymbol{\phi}_k, \tag{4}$$

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

$$\begin{bmatrix} | & | & & | \\ \boldsymbol{\phi}_1 & \boldsymbol{\phi}_2 & \cdots & \boldsymbol{\phi}_{K+1} \\ | & | & & | \end{bmatrix} \overset{\text{LS}}{=} A \begin{bmatrix} | & | & & | \\ \boldsymbol{\phi}_0 & \boldsymbol{\phi}_1 & \cdots & \boldsymbol{\phi}_K \\ | & | & & | \end{bmatrix}, \tag{5}$$

where the system matrix A minimizes the sum of squared residuals over the sequence. The optimal linear flow described by this matrix equation naturally generates approximate dynamics $\boldsymbol{\phi}_{\text{DMD}}(t) = A^{\frac{t}{\Delta t}}\boldsymbol{\phi}_0$ governed by the 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 [34–38].

The standard DMD approach described above for classical dynamics cannot be immediately translated to quantum dynamics. The DMD approximation of the system evolution requires complete knowledge of the 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 [1]. To address this challenge, we employ a technique motivated by Takens' embedding theorem [50–52] 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. (An alternative approach involves forming a state vector from multiple distinct operators [53].)

Takens' embedding theorem [50, 51] 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 $d_\star \leq 2N+1$ time-delayed observables. The correspondence reads

$$\boldsymbol{\phi}(t) \leftrightarrow \boldsymbol{o}_{t,d_\star} = \begin{bmatrix} o(t) \\ o(t+\Delta\tau) \\ \vdots \\ o(t+(d_*-1)\Delta\tau) \end{bmatrix}, \qquad (6)$$

where $\Delta\tau > 0$ is the time delay, $\boldsymbol{\phi}(t)$ is the system state, $o(t) = o[\boldsymbol{\phi}(t)]$ is the measured observable, and $\boldsymbol{o}_t$ is the $d_\star$-dimensional "observable trajectory" containing the dynamical information. The RHS of Eq. (6) is known as a $d_\star$-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. This adaptation to the quantum realm is illustrated in Fig. 2.

In anticipation of efficiently leveraging near-term quantum resources in the hybrid quantum-classical approach that we present in Section 3.2, 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 $\{\boldsymbol{o}_{t_k,d}\}_{k=0}^{K+1}$, each of some length $d \leq d_\star$,

$$\boldsymbol{o}_{t_k,d} = \begin{bmatrix} o(t_k) \\ o(t_{k+1}) \\ \vdots \\ o(t_{k+d-1}) \end{bmatrix}, \qquad 0 \leq k \leq K+1. \qquad (7)$$

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 $\boldsymbol{o}_{t_k}$ as columns

$$\mathbf{X}_{k_1:k_2} = \begin{bmatrix} \boldsymbol{o}_{t_{k_1},d} & \boldsymbol{o}_{t_{k_1+1},d} & \cdots & \boldsymbol{o}_{t_{k_2},d} \end{bmatrix}, \qquad (8)$$

has a Hankel form, *i.e.*, the matrix elements on each anti-diagonal are equal (see Fig. 3). In the embedding space, we can then identify the closest linear flow,

$$\mathbf{X}_{1:K+1} \overset{\text{LS}}{=} A\mathbf{X}_{0:K} \implies A = \mathbf{X}_{1:K+1}(\mathbf{X}_{0:K})^+, \qquad (9)$$

where $(\cdot)^+$ denotes the Moore–Penrose pseudo-inverse. The system matrix A assumes a companion structure with just d free parameters (see Eq. (73) in Appendix B). The approximation to the system dynamics is then stored in d parameters inferred from measurements of a total of $K+d+1$ delayed observables. We hence name our least-squares embedding in the observable space the *observable dynamic mode decomposition* (ODMD).

3.2 Estimating the ground state energy

We now specialize the ODMD method in pursuit of a robust eigenenergy estimator for quantum systems. With Hamiltonian H, the system dynamics are given by the time evolution operator e^{-iHt} (with the convention of $\hbar = 1$). Let $E_0 \leq E_1 \leq \cdots \leq E_{N-1}$ denote the eigenenergies of the Hamiltonian. We consider as our observable the complex-valued overlap,

$$s(k\Delta t) = \langle \boldsymbol{\phi}_0 | \, e^{-iHk\Delta t} \, | \boldsymbol{\phi}_0 \rangle, \quad k = 0, 1, \ldots, \qquad (10)$$

whose real and imaginary parts can be separately measured, *e.g.*, using the Hadamard test [48]. For an integer pair (d, K) specifying our time-delayed embedding, this requires $K+d+1$ observables in total, each corresponding to a distinct quantum evolution circuit.

Upon arranging these measured overlaps into a pair of time-shifted Hankel matrices $\mathbf{X}, \mathbf{X}' \in \mathbb{C}^{d\times(K+1)}$,

$$\mathbf{X} = \mathbf{X}_{0:K}, \qquad \mathbf{X}' = \mathbf{X}_{1:K+1}, \qquad (11)$$

the eigenenergies of the Hamiltonian can then be simply estimated by solving the standard eigenvalue problem,

$$A\boldsymbol{\Psi}_\ell = \tilde{\lambda}_\ell \boldsymbol{\Psi}_\ell, \qquad (12)$$

where the matrix A is defined in Eq. (9). Intuitively, we can read off our eigenenergy estimates $\tilde{E}_\ell$

[1] 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.

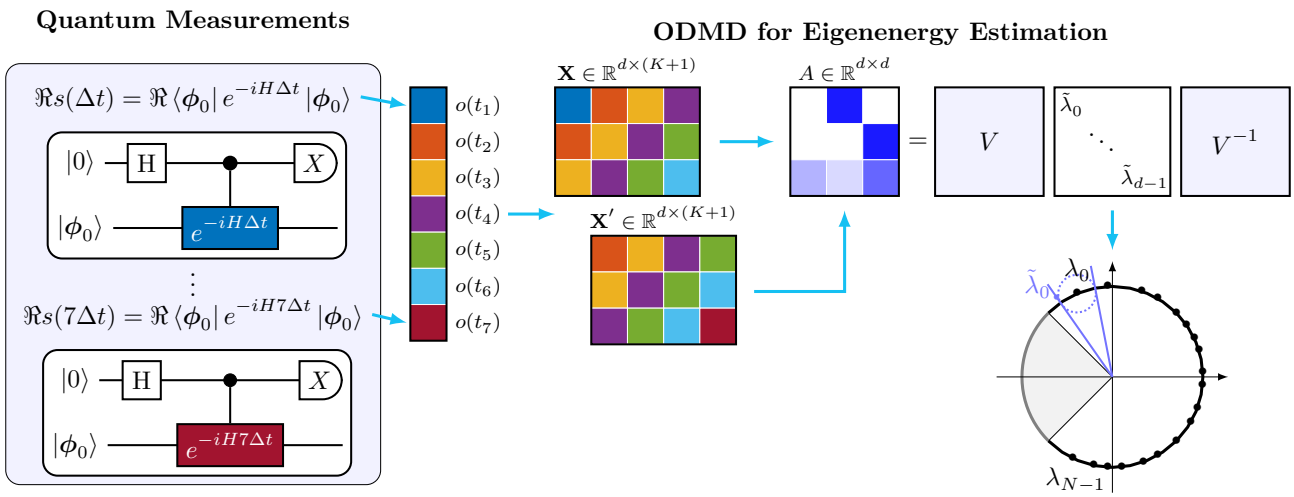

Figure 3: ODMD applied to eigenenergy estimation. ODMD uses the real (or imaginary) part of the expectation value, $s(k\Delta t) = \langle \phi_0 | e^{-iHk\Delta t} | \phi_0 \rangle$, of time evolution with respect to a reference state $|\phi_0\rangle$. This data can be measured efficiently on a quantum processor, for example by means of the Hadamard test. ODMD constructs a pair of Hankel matrices $\mathbf{X}, \mathbf{X}'$ using Takens' embedding (matrix elements of the same color are equal) and generates the DMD system matrix A, which has a companion structure. Under mild assumptions on Δt, the eigenvalue $\tilde{\lambda}_0$ of A with $\arg(\tilde{\lambda}_0) = \max_\ell \arg(\tilde{\lambda}_\ell)$ converges to the true ground state phase λ_0 as the size of $\mathbf{X}, \mathbf{X}'$ increases. The ground state energy is estimated as $\tilde{E}_0 = -\arg(\tilde{\lambda}_0)/\Delta t$.

from the ordering of the phases $\arg(\lambda_\ell)$. The eigenvalue $\tilde{\lambda}_0 \approx e^{-iE_0\Delta t}$ attaining the maximal phase, $\arg(\tilde{\lambda}_0) = \max_\ell \arg(\tilde{\lambda}_\ell) = -\min_\ell \tilde{E}_\ell \Delta t$, encodes the DMD approximation $\tilde{E}_0$ to the true ground state energy E_0. For concreteness, we construct our $d\times(K+1)$ data matrices using $d = \lfloor \alpha(K+1) \rfloor$ for $\alpha = \frac{1}{2}$ fixed throughout the main text. Schematically, an overview of the ODMD approach for the ground state problem is provided in Fig. 3.

To protect against both statistical noise and quantum hardware noise, we exploit the least-squares formulation given in Eq. (9), benefiting from its rigorous and robust regularization capabilities. Since the recovery of the system matrix A requires computing the pseudo-inverse of the data matrix $\mathbf{X}$, its conditioning critically influences the stability of the solution. Here we use the singular value decomposition of $\mathbf{X}$,

$$\mathbf{X} = \sum_{\ell=0}^{d-1} \sigma_\ell \boldsymbol{u}_\ell \boldsymbol{v}_\ell^\dagger, \tag{13}$$

where $\sigma_\ell > 0$ are the singular values, and $(\boldsymbol{u}_\ell, \boldsymbol{v}_\ell)$ specify the associated left and right singular vectors, respectively. To ensure noise-resilience, we filter out perturbative error using a thresholding procedure: we regularize by truncating singular values smaller than the cutoff, $\tilde{\delta}\sigma_{\max}(\mathbf{X})$, with a relative threshold $\tilde{\delta} > 0$ of our choice,

$$\mathbf{X} \mapsto \mathbf{X}_{\tilde{\delta}} = \sum_{\ell:\sigma_\ell > \tilde{\delta}\sigma_{\max}} \sigma_\ell \boldsymbol{u}_\ell \boldsymbol{v}_\ell^\dagger, \tag{14}$$

where $\sigma_{\max} = \max_\ell \sigma_\ell$ denotes the largest singular value of the data matrix $\mathbf{X}$. Incorporating the LS regularization, our measurement-driven estimation of the ground state energy is outlined in Algorithm 1. Our algorithm takes as input the time step Δt and noise threshold $\tilde{\delta}$.

Algorithm 1: ODMD ground state energy estim.

Input: Time step Δt, noise threshold $\tilde{\delta}$.
Output: Estimated ground state energy $\tilde{E}_0$.

$k \leftarrow 0$
while not converged **do**
  $s_k \leftarrow \Re s(k\Delta t)$; /* quantum measurement */
  $\mathbf{X}, \mathbf{X}' \leftarrow \text{Hankel}(s_0, s_1, \ldots, s_k)$; /* Eq. (11) */
  $\mathbf{X}_{\tilde{\delta}} \leftarrow \mathbf{X}$; /* least-squares regularization */
  $A \leftarrow \mathbf{X}'\mathbf{X}_{\tilde{\delta}}^+$; /* update system matrix */
  $\tilde{E}_0 \Delta t \leftarrow -\max_{1\le\ell\le d_k} \Im \log(\tilde{\lambda}_\ell)$; /* update energy */
  $k \leftarrow k+1$
end

The choice of the time step Δt is crucial for the optimal performance of the ODMD algorithm. Small time steps can hinder convergence due to the collection of excessively correlated information. On the other hand, there is a fundamental limitation on the maximum size of Δt, *i.e.*, when we extract $\tilde{\lambda}_0 \approx e^{-iE_0\Delta t}$ from ODMD, we can only solve for the leading eigenangle $E_0\Delta t \bmod 2\pi$. Thus to extract E_0 unambiguously, the angles $\{E_n\Delta t\}_{n=0}^{N-1}$ can wrap around the circle only once, which requires $(E_{N-1} - E_0)\Delta t < 2\pi$, and we must also identify the location of the dividing

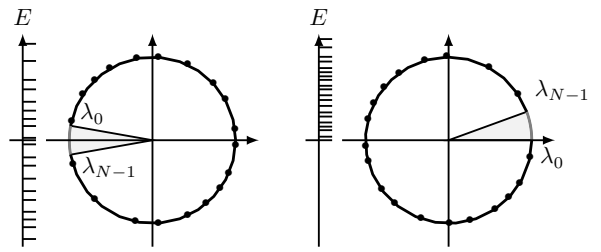

Figure 4: Guideline to select Δt and unambiguously extract $\tilde{E}_0$ from the ODMD eigenvalue $\tilde{\lambda}_0$, depending on known bounds on the eigenenergy spectrum, E_{min} and E_{max}. The region shaded in gray shows where the choice of Δt ensures there are no eigenphases λ_n and separates the top and bottom of the eigenangles $E_n \Delta t$. **Left:** When the spectral bounds are symmetric around 0, *i.e.*, $\|H\|_2 \leq E_{\mathsf{max}} = |E_{\mathsf{min}}|$, we take $\Delta t \lesssim \pi/E_{\max}$ and the dividing boundary between the upper and lower parts of the eigenangles sits at π. **Right:** When the whole energy spectrum is positive with asymmetric spectral bounds $0 = E_{\min} < \|H\|_2 \leq E_{\max}$, taking $\Delta t \lesssim 2\pi/E_{\mathsf{max}}$ places the dividing eigenangle boundary at 0^+.

line between upper and lower parts of the eigenangles. We show two illustrative examples in Fig. 4.

Note that the exact largest and smallest eigenenergies do not need to be known in advance. Rather, if we know that $E_{\min} \leq \langle \boldsymbol{\phi}|H|\boldsymbol{\phi}\rangle \leq E_{\max}$ for any normalized state $|\boldsymbol{\phi}\rangle$, we can guarantee that the phase condition is satisfied by choosing Δt so that $(E_{\max} - E_{\min})\Delta t < 2\pi$. Moreover, Δt must also satisfy a further compatibility condition: as derived in Appendix B,

$$\Delta t < \frac{2\pi}{(E_{N-1} - E_0) + (E_1 - E_0)} \tag{15}$$

in order to ensure the convergence of the ODMD estimate to the ground state energy. Notably, the restriction on Δt for unambiguous determination of E_0 from λ_0 actually follows directly from (15).

In practice, a good choice is to first bound the spectral range of the Hamiltonian and then linearly shift the range to be in $[-C\pi, C\pi]$, for some constant $C < 1$ close to but less than 1; the lower right phase sketch in Fig. 3 shows a restriction with $C = 3/4$. The time step Δt can then be set to 1, giving rapid convergence of the algorithm while satisfying the condition of Eq. (15). Alternatively, if the Hamiltonian has spectral norm $\|H\|_2$, we can simply use $\Delta t < \pi/\|H\|_2$. Either way, the spectrum of $H\Delta t$ will be guaranteed to lie in the window $(-\pi, \pi)$ and there is no ambiguity in estimating the eigenenergy E_0.

Finally, we leave an important remark that measuring only the real or imaginary part of the overlap, $\Re s(k\Delta t)$ or $\Im s(k\Delta t)$, leads to a real-valued system matrix A. Accordingly, the eigenvalues of A contain pairs of complex conjugates. These ODMD eigenvalues, however, preserve the ground state information due to Euler's identity, for example, $\cos(E_0\Delta t) = \frac{1}{2}(e^{-iE_0\Delta t} + e^{iE_0\Delta t})$. The eigenenergies can be explicitly approximated as the phases of the ODMD eigenvalues just as in the complex case. Because of the phase symmetrization $E_\ell \mapsto \pm E_\ell$, what appears to be the ground state energy could instead come from the symmetrization of the maximum energy eigenvalue, if $|E_{N-1}| > |E_0|$. Thus to ensure that the maximal phase ODMD eigenvalue really corresponds to the ground state energy E_0, we need to shift the energy range of H so that $|E_{N-1}| \leq |E_0|$. This is guaranteed if the applied spectral shift satisfies, $E_0 < 0$ and $|E_0| = \|H\|_2$. Even accounting for this constant shift, the use of $\Re s(k\Delta t)$ can give a reduction in the total measurement cost of up to 50%.

3.3 Accessing the ground state

In addition to accessing the eigenenergies, the ODMD algorithm can also provide an estimate to the corresponding eigenstates. As a first step, let us recall that the companion system matrix A is non-Hermitian, and it admits a spectral decomposition $A = \sum_\ell \tilde{\lambda}_\ell \boldsymbol{\Psi}_\ell \boldsymbol{\Phi}_\ell$ with right and left eigenvectors, $\boldsymbol{\Psi}_\ell$ and $\boldsymbol{\Phi}_\ell$, respectively. The left eigenvector $\boldsymbol{\Phi}_0 = [z_0, z_1, \ldots, z_{d-1}]$ corresponding to the eigenvalue $\tilde{\lambda}_0$ satisfies

$$\boldsymbol{\Phi}_0 A = \tilde{\lambda}_0 \boldsymbol{\Phi}_0 \implies \boldsymbol{\Phi}_0 \mathbf{X}' = \tilde{\lambda}_0 \boldsymbol{\Phi}_0 \mathbf{X}. \tag{16}$$

Focusing on observable snapshots from, for example, the first columns of our data matrices, Eq. (16) implies

$$\boldsymbol{\Phi}_0 \begin{bmatrix} s(\Delta t) \\ s(2\Delta t) \\ \vdots \\ s(d\Delta t) \end{bmatrix} = \tilde{\lambda}_0 \boldsymbol{\Phi}_0 \begin{bmatrix} s(0) \\ s(\Delta t) \\ \vdots \\ s((d-1)\Delta t) \end{bmatrix}, \tag{17}$$

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

$$\begin{aligned} \langle \boldsymbol{\phi}_0 | \, e^{-iH\Delta t} \left(\sum_{\ell=0}^{d-1} z_\ell e^{-iH\ell\Delta t} \, |\boldsymbol{\phi}_0\rangle \right) \\ = \tilde{\lambda}_0 \, \langle \boldsymbol{\phi}_0 | \left(\sum_{\ell=0}^{d-1} z_\ell e^{-iH\ell\Delta t} \, |\boldsymbol{\phi}_0\rangle \right). \end{aligned} \tag{18}$$

Since $\tilde{\lambda}_0 \approx e^{-iE_0\Delta t}$, the dynamic mode above closely follows the ground state oscillation driven at the desired frequency, $\langle \boldsymbol{\phi}_0 | e^{-iH\Delta t} | \psi_0 \rangle = e^{-iE_0\Delta t} \langle \boldsymbol{\phi}_0 | \psi_0 \rangle$, where $|\psi_0\rangle$ is the actual ground state. By Eq. (18), we thus approximate the ground state as the weighted sum,

$$|\psi_0\rangle \approx |\tilde{\psi}_0\rangle = \frac{\sum_{\ell=0}^{d-1} z_\ell e^{-iH\ell\Delta t} \, |\boldsymbol{\phi}_0\rangle}{\|\sum_{\ell=0}^{d-1} z_\ell e^{-iH\ell\Delta t} \, |\boldsymbol{\phi}_0\rangle\|_2}, \tag{19}$$

where the set of coefficients z_ℓ assign the relative weights to the real-time states $e^{-iH\ell\Delta t} \, |\boldsymbol{\phi}_0\rangle$. For simplicity, we assume that the coefficients z_ℓ produce a

normalized state so that the denominator of Eq. (19) is unity. The ground state properties can hence be extracted in terms of the real-time states,

$$\langle\psi_0|O|\psi_0\rangle \approx \sum_{k,\ell=0}^{d-1} z_k^* z_\ell \langle\phi_0|e^{iHk\Delta t} O e^{-iH\ell\Delta t}|\phi_0\rangle\,, \tag{20}$$

for an arbitrary operator O. As the mode approximation typically admits small d at convergence, we remark that ODMD enables us to extract a highly compact representation of the eigenstates through real-time evolution.

3.4 Practical considerations

For practical implementation, we discuss the feasibility of executing the ODMD algorithm on near-term devices, particularly when the time evolution of the initial state is inexact. Here we assess two sources of inexactness, the algorithmic error due to approximation(s) of the operator exponentials for time evolution and the device error from noisy quantum hardware.

For algorithmic error, we consider the Trotter-Suzuki product formula [54] as one of the standard approaches for approximating the time evolution operator. Since our algorithm is based on time-delayed embedding, it suffices to accurately approximate the unit-timestep propagator $e^{-iH\Delta t}$. For a first-order Trotter decomposition, we have

$$e^{-iH\Delta t} \approx \left[\prod_\nu e^{-iH_\nu \Delta t/M}\right]^M, \tag{21}$$

where $H = \sum_\nu H_\nu$ gives a decomposition of the problem Hamiltonian into Pauli terms $H_\nu = c_\nu P_\nu$ (with P_ν being a Pauli string and c_ν a scaling coefficient satisfying $|c_\nu| = \|H_\nu\|_2$). In particular, the Trotter error $\epsilon_{\rm TS} = \mathcal{O}(\frac{\Delta t^2 \Gamma}{M})$ is controlled by a constant Γ dependent on commutators of the terms H_ν. Accordingly, the inexact evolution can be performed with circuit depth $\mathcal{O}(\frac{N_H\sqrt{M\Gamma}t_{\max}}{\sqrt{\epsilon_{\rm TS}}})$, where N_H is the number of Pauli terms in the Hamiltonian and $t_{\max}$ is the maximal evolution time.

Recall that the timestep Δt in the ODMD algorithm is selected as $\Delta t = \mathcal{O}(\|H\|_2^{-1})$. For a typical many-body Hamiltonian, it suffices to set $M = \frac{1}{\mathrm{polylog}(N)\epsilon_{\rm TS}}$ in the Trotterized evolution to ensure an approximation error of $\mathcal{O}(\epsilon_{\rm TS})$, introducing a spectral perturbation of $\mathcal{O}(\epsilon_{\rm TS})$. With time steps $t_k = k\Delta t$ taken at regular intervals, the total evolution time scales as $t_{\rm tot} = \mathcal{O}(t_{\max}^2)$ and the total gate count grows accordingly. In addition, the total shot count is proportional to

$$k_{\max} \times \text{Measurement cost per circuit} = \frac{t_{\max}}{\Delta t \epsilon^2}, \tag{22}$$

where ϵ quantifies the sampling error due to finite shots per overlap measurement using the Hadamard test. The $\mathcal{O}(\epsilon^{-2})$ dependence reflects the standard measurement scaling, where mitigating statistical uncertainty requires quadratically more shots. Provided that the eigenenergy perturbation due to the Trotterized Δt-evolution is sufficiently small, the ODMD estimates remain robust. As an example, choosing $t_{\max} = \mathcal{O}(\epsilon_{\rm TS}^{-\eta})$ with $\frac{2}{3} \le \eta \le 1$ and $\epsilon = \mathcal{O}(1)$ guarantees a final energy error of $\mathcal{O}(\epsilon_{\rm TS})$ [55]. Notably, the stable convergence of ODMD under inexact time evolution has been confirmed directly in recent numerical and experimental simulations of frustrated spin lattices [56].

For device error, we first analyze the elementary case of a depolarizing channel, perhaps the simplest and most ubiquitous error models. The global depolarizing channel acts as,

$$\Lambda_t[\rho;\zeta] = e^{-\zeta|t|}\mathcal{U}_t\rho\mathcal{U}_t^\dagger + \frac{1-e^{-\zeta|t|}}{2N} I_{2N}, \tag{23}$$

where ρ is any input state consisting of the ancilla and register qubits, $\mathcal{U}_t = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{-iHt}$ denotes the controlled time evolution queried in a Hadamard test, and the positive parameter $\zeta > 0$ characterizes the noise rate. In the ODMD setup, we have $\rho = |+\rangle\langle +| \otimes |\phi_0\rangle\langle\phi_0|$ so that the overlap, *e.g.*, $\Re\langle\phi_0|e^{-iHt}|\phi_0\rangle$ for $t > 0$, under noisy time evolution becomes

$$\mathrm{Tr}\left[(X\otimes I)\Lambda_t[\rho]\right] = e^{-\zeta t}\Re\langle\phi_0|e^{-iHt}|\phi_0\rangle\,. \tag{24}$$

The ODMD observables under depolarizing noise exhibit exponential damping, which is reflected in the real part $\Re\log(\tilde{\lambda}_\ell)$ of the computed ODMD eigenphases. Because the real and imaginary parts of the ODMD eigenphases can be extracted separately, our energy estimate remains accurate despite such noise-induced spectral broadening. The generalization to arbitrary Markovian noise channels is explored in Appendix A.

4 Convergence

Alongside its algorithmic elegance, ODMD also converges rapidly. Here we present our main convergence analysis that demonstrates efficient estimation of the ground state energy via an algebraic argument.

We first note that the overlap can always be expressed as a sum of complex sinusoids,

$$s_k = s(k\Delta t) = \sum_{n=0}^{N-1} p_n \lambda_n^k, \tag{25}$$

where N is the size of Hilbert space, $\lambda_n = e^{-iE_n\Delta t}$ are the eigenphases accumulated from a single time step, and $p_n = |\langle\psi_n|\phi_0\rangle|^2$ specify the eigenstate probabilities in the reference state $|\phi_0\rangle$. For the extreme

case of sufficient sample data with $d = K + 1 = N$,

$$A = V \begin{bmatrix} \lambda_0 & & & \\ & \lambda_1 & & \\ & & \ddots & \\ & & & \lambda_{N-1} \end{bmatrix} V^{-1}, \tag{26}$$

where $(V_{ij})_{i,j=1}^{N} = \lambda_{i-1}^{j-1}$ is the Vandermonde matrix containing the eigenphases. In this case, the eigenvalue problem of Eq. (12) resolves the entire Hamiltonian spectrum, echoing classic techniques of Prony [57] and Padé [58, 59]. Here we highlight the utility of our estimator as a powerful extension when working with limited data samples $d < K \ll N$. Under a mild assumption that we regulate our time step Δt to keep the eigenphases in a 2π-window satisfying Eq. (15), we can show a rapid convergence of the energy estimation,

$$|\tilde{E}_0 - E_0| \leq \mathcal{O}\left(Ke^{-\beta d}\right), \tag{27}$$

where

$$\beta > -\ln \frac{\sin((E_{N-1} - E_1)\Delta t/2)}{\sin((E_{N-1} - E_0)\Delta t/2)}, \tag{28}$$

denotes the leading-order exponent that depends on the normalized spectral gap $(E_1 - E_0)/(E_{N-1} - E_0)$.

A proof of this error bound exploits the minimal residual property of DMD solvers. We elaborate the proof in its entirety in Appendix B, though sketch the main ideas here.

Our proof relies on the companion structure of the system matrix $A \in \mathbb{C}^{d\times d}$. The matrix is uniquely determined by its last row denoted as $\vec{a} = (a_0, a_1, \cdots, a_{d-1}) \in \mathbb{C}^d$. The eigenvalues $\tilde{\lambda}_\ell$ of A, corresponding to the roots of the following characteristic polynomial,

$$\mathcal{C}_{\mathrm{DMD}}(z) = \sum_{\ell=0}^{d} a_\ell z^\ell = \prod_{\ell=0}^{d-1} (z - \tilde{\lambda}_\ell), \tag{29}$$

exhibit a smooth dependence on $\vec{a}$. We note that $a_d \equiv 1$. The matrix elements are determined by the LS regression minimizing the total residual. In particular, we prove that the ODMD solution admits a residual at most exponentially small in the size d of the system matrix.

On the other hand, we can also construct a new companion matrix, C, that is of the same size as A and has the two following key properties: (1) its maximal phase eigenvalue is precisely $\lambda_0 = e^{-iE_0\Delta t}$, where E_0 is the true ground state energy; and (2) the difference between A and C can be effectively bounded from residual analysis. The crux of the proof is that such a C can indeed be constructed, which we show by extending Prony's idea [57] from the full Hilbert space dimension N to the much smaller dimension d.

In particular, the eigenvalues of C, which has last row $\vec{c} = (c_0, c_1, \cdots, c_{d-1}) \in \mathbb{C}^d$, are determined by the characteristic polynomial,

$$\mathcal{C}_\Omega(z) = \sum_{\ell=0}^{d} c_\ell z^\ell = (z - \lambda_0) \prod_{\ell=1}^{d-1} (z - \lambda_{n_\ell}), \tag{30}$$

where λ_0 is, again, the exact ground state phase and $(\lambda_{n_1}, \lambda_{n_2}, \cdots, \lambda_{n_{d-1}})$ are $d-1$ eigenphases different from λ_0. We show in Appendix B that there exists some choice of $\{\lambda_{n_\ell}\}_{\ell=1}^{d-1}$ such that C satisfies the above conditions (1) and (2). Then based on an elementary smoothness argument, proximity in the characteristic polynomials $\vec{a}$ and $\vec{c}$ implies proximity in their respective roots as the eigenvalues of the associated companion matrices. Specifically, we incur, using perturbation analysis, a nearly exponentially small error when approximating the exact ground state phase with the maximal phase ODMD eigenvalue $\tilde{\lambda}_0 = e^{-i\tilde{E}_0\Delta t}$.

As an aside, we observe that in the idealized case of a perfect initial state, where $|\phi_0\rangle = |\psi_0\rangle$, ODMD requires only two time steps to recover the ground state energy, as the overlap signal contains a single frequency component. This reveals how a more informative reference state can further enhance the energy convergence. However, unlike alternative approaches that rely more critically on initial state preparation, ODMD does not require a dominantly high ground state overlap to perform effectively.

5 Connections to classical techniques

The ODMD approach is closely related to prominent eigenvalue methods from the classical literature. In this section, we explore the connections to Krylov and matrix pencil methods to gain insight into the stable convergence of our algorithm and its ability to consistently determine the eigenenergies.

5.1 Oblique two-sided Krylov method

The ODMD algorithm can be interpreted as a hybrid quantum-classical oblique projection method onto two-sided Krylov subspaces [60]. Krylov subspace approaches [61] variationally target the ground state via searching in a systematically constructed low-energy subspace, and are thus especially powerful for large-scale eigenvalue problems found in quantum physics.

Defining $\mathcal{A} := e^{-iH\Delta t}$, we first observe that the shifted Hankel matrices in Eq. (11) can be combined linearly to form what is known in the numerical linear algebra community as a "matrix pencil," $\mathbf{X}' - \lambda\mathbf{X}$. The matrix pencil can be decomposed as follows:

$$\mathbf{X}' - \lambda\mathbf{X} = \mathbf{W}_d^\dagger (\mathcal{A} - \lambda I) \mathbf{V}_{K+1}, \tag{31}$$

where

$$\begin{aligned} \mathbf{W}_d &:= \left[|\phi_0\rangle \quad \mathcal{A}^\dagger |\phi_0\rangle \quad \cdots \quad (\mathcal{A}^{d-1})^\dagger |\phi_0\rangle\right], \\ \mathbf{V}_{K+1} &:= \left[|\phi_0\rangle \quad \mathcal{A} |\phi_0\rangle \quad \cdots \quad \mathcal{A}^K |\phi_0\rangle\right]. \end{aligned} \tag{32}$$

When $d = K + 1 = N$, Eq. (31) represents a linear transformation that preserves the operator spectrum provided that $\mathbf{W}_N, \mathbf{V}_N$ possess full rank. In this case, we have that the spectrum of the matrix pencil $\mathbf{X}' - \lambda\mathbf{X}$ corresponds exactly to the eigenfrequencies of $e^{-iH\Delta t}$. However, we typically do not run our algorithm to the full dimension N and instead extract eigenvalue approximations for $d, K \ll N$. From Eqs. (31) and (32), we observe that in this setting, the rectangular Hankel matrix pencil, $\mathbf{X}' - \lambda\mathbf{X}$, corresponds to a two-sided projection of $\mathcal{A} - \lambda I$ onto the left and right Krylov subspaces,

$$\begin{aligned} \mathcal{L}_d &= \text{span}\left\{|\phi_0\rangle, \mathcal{A}^\dagger |\phi_0\rangle, \ldots, (\mathcal{A}^{d-1})^\dagger |\phi_0\rangle\right\}, \\ \mathcal{R}_{K+1} &= \text{span}\left\{|\phi_0\rangle, \mathcal{A} |\phi_0\rangle, \ldots, \mathcal{A}^K |\phi_0\rangle\right\}, \end{aligned} \tag{33}$$

respectively, for which $\mathbf{W}_d$ and $\mathbf{V}_{K+1}$ form a pair of non-orthogonal bases. Specifically, $\mathbf{X}'$ is the oblique projection of $\mathcal{A}$, the unitary evolution operator, onto $\mathcal{R}_{K+1}$ along $\mathcal{L}_d$, while $\mathbf{X} = \mathbf{W}_d^\dagger \mathbf{V}_{K+1}$ gives the two-sided overlap matrix. In the case of $\mathbf{X} = I$, *i.e.*, the bases for $\mathcal{L}_d$ and $\mathcal{R}_{K+1}$ are mutually orthogonal, this corresponds to a biorthogonal Krylov process [60]. The procedure to compute the system matrix A from Eq. (9) using the LS regularization gives a robust recipe to convert the generalized eigenvalue problem of Eq. (31) to a well-conditioned standard eigenvalue problem of Eq. (12).

Moreover, ODMD exploits the freedom of splitting the subspace dimensions d and $K + 1$ for the left and right Krylov spaces, respectively, in order to choose a combination such that each subspace individually is in general better-conditioned compared to other subspace methods exploiting real-time evolution. The most natural comparison is with the unitary variational quantum phase estimation algorithm (UVQPE) [22]. In UVQPE, $\mathcal{A}$ is projected *one-sided* on $\mathbf{V}_m$, i.e., UVQPE solves the reduced generalized eigenvalue problem $\mathbf{V}_m^\dagger(\mathcal{A} - \lambda I)\mathbf{V}_m$ which is constructed from measuring m observables. ODMD uses the same m observables and instead generates a two-sided projection using $\mathbf{W}_{m/3}$ and $\mathbf{V}_{2m/3}$. As the conditioning of D-dimensional non-orthogonal Krylov bases (32) typically grows as $\mathcal{O}(\gamma^D)$ for some $\gamma > 1$ [62], we see that ODMD effectively reduces the problem of ill-conditioning compared to UVQPE by reducing the subspace dimensions ($D = m/3$ or $2m/3$ instead of m) while simultaneously preserving the information and distributing it asymmetrically over a pair of subspaces. The connection with projection methods provides an intuitive explanation for the validity of ODMD as Krylov subspace methods are known to extract accurate eigenvalue approximations from low-dimensional subspaces.

5.2 Function Krylov method

Here we highlight some additional theoretical intuition about the efficacy of data Hankelization for ODMD, further elucidating its connection to the Krylov subspace approaches. We establish this connection through the Koopman operator analysis [34–38]. We consider a physical Hilbert space $\mathcal{H} = \mathbb{C}^N$. The Koopman (super)operator, $\mathcal{K}_{\Delta t}$, probes the underlying dynamics of the system by mapping scalar functions, $f : \mathcal{H} \to \mathbb{C}$, or observables in our language, to scalar functions. For a function f, $\mathcal{K}_{\Delta t}$ is defined by

$$\mathcal{K}_{\Delta t}[f](|\phi\rangle) = f(e^{-iH\Delta t} |\phi\rangle), \tag{34}$$

which gives a push-forward of the dynamics via the time evolution $e^{-iH\Delta t}$. For an eigenstate $|\psi_n\rangle$ of the quantum system, we define an associated function f_n by

$$f_n(|\phi\rangle) = \langle\psi_n|\phi\rangle, \tag{35}$$

which is an eigenfunction of $\mathcal{K}_{\Delta t}$ with eigenvalue $\lambda_n = e^{-iE_n\Delta t}$,

$$\mathcal{K}_{\Delta t}[f_n] = \lambda_n f_n. \tag{36}$$

Similarly, we define a function f_ψ from a state $|\psi\rangle$ by

$$f_\psi(|\phi\rangle) = \langle\psi|\phi\rangle = \sum_{n=0}^{N-1} \langle\psi|\psi_n\rangle \langle\psi_n|\phi\rangle. \tag{37}$$

While the Koopman operator itself is infinite-dimensional, the functions f_ψ reside in an invariant subspace of $\mathcal{K}_{\Delta t}$ spanned by $\{f_n\}_{n=0}^{N-1}$. If we use a state $|\psi\rangle$ that maintains nonzero overlap with our target energy eigenstates, then we can translate the estimation of the Hamiltonian spectrum to a dual eigenvalue problem defined on the function space. For a delay embedding dimension d, we consider the collection of d functions,

$$\boldsymbol{g} = \left[f_\psi \quad \mathcal{K}_{\Delta t}[f_\psi] \quad \cdots \quad (\mathcal{K}_{\Delta t})^{d-1}[f_\psi]\right]^\top, \tag{38}$$

within the invariant subspace. For any such function f_ψ, we can approximate $(\mathcal{K}_{\Delta t})^d[f_\psi]$ as a linear combination of the functions in $\boldsymbol{g}$, and this becomes exact as d grows to the full Hilbert space dimension N. When $d < N$, we can look for a linear operator $\mathbf{K}_{\Delta t} \in \mathbb{C}^{d\times d}$ that gives an optimal approximation on the invariant subspace. To find $\mathbf{K}_{\Delta t}$, we formulate the dual problem,

$$\underbrace{\begin{bmatrix} | & & | \\ \boldsymbol{g}(e^{-iH\Delta t}|\phi_0\rangle) & \cdots & \boldsymbol{g}(e^{-iH\Delta t}|\phi_K\rangle) \\ | & & | \end{bmatrix}}_{\mathbf{G}'\in\mathbb{C}^{d\times(K+1)}} \overset{\text{LS}}{=} \mathbf{K}_{\Delta t} \underbrace{\begin{bmatrix} | & & | \\ \boldsymbol{g}(|\phi_0\rangle) & \cdots & \boldsymbol{g}(|\phi_K\rangle) \\ | & & | \end{bmatrix}}_{\mathbf{G}\in\mathbb{C}^{d\times(K+1)}}, \tag{39}$$

relaxing the pointwise equality on $\mathcal{H}$ to a LS equality on the states $\{|\phi_k\rangle\}_{k=0}^K$. This matrix approximation yields a factorization, $\mathbf{K}_{\Delta t} = \mathbf{G}'(\mathbf{G})^+ = \mathbf{Z}'(\mathbf{Z})^+$, with

$$\mathbf{Z}' = \frac{1}{K+1}\sum_{k=0}^{K} \boldsymbol{g}(e^{-iH\Delta t}|\phi_k\rangle)\boldsymbol{g}^\dagger(|\phi_k\rangle), \tag{40}$$

$$\mathbf{Z} = \frac{1}{K+1}\sum_{k=0}^{K} \boldsymbol{g}(|\phi_k\rangle)\boldsymbol{g}^\dagger(|\phi_k\rangle), \tag{41}$$

from the definition $\mathbf{G}^+ = (\mathbf{G}^\dagger\mathbf{G})^{-1}\mathbf{G}^\dagger$. This factorization, as we shall show, has an intimate connection to the Krylov subspace approaches. [60] For the ODMD setting, we use $|\psi\rangle = |\phi_0\rangle$ and $|\phi_k\rangle = e^{-iHk\Delta t}|\phi_0\rangle$ where $|\phi_0\rangle$ is our reference state. To proceed, we define the functional inner product $\langle f, f'\rangle_\mu$ for functions f and f' as,

$$\langle f, f'\rangle_\mu = \int_{|\phi\rangle\in\mathcal{H}} d\mu\, f^*(|\phi\rangle) f'(|\phi\rangle), \tag{42}$$

where μ is the empirical measure,

$$\mu = \frac{1}{K+1}\sum_{k=0}^{K} \delta(|\phi\rangle - |\phi_k\rangle), \tag{43}$$

on the state trajectory $\{|\phi_k\rangle\}_{k=0}^K$ composed of $(K+1)$ Dirac δ-functions $\delta(|\phi\rangle - |\phi_k\rangle)$. Equations (40) and (41) can then be expressed using the $L^2(\mu)$ inner product defined above,

$$\mathbf{Z}'_{ij} = \langle g_j, \mathcal{K}_{\Delta t}[g_i]\rangle_\mu, \tag{44}$$

$$\mathbf{Z}_{ij} = \langle g_j, g_i\rangle_\mu, \tag{45}$$

where we interpret $\mathbf{Z}'$ and $\mathbf{Z}$ as matrix representations of the functional "target" and "overlap" operators, respectively, projected onto a Krylov subspace spanned by the functions $\{\mathcal{K}^i_{\Delta t}[f_\psi]\}_{i=0}^{d-1}$. An eigenpair $(\lambda, \vec{v})$ of $\mathbf{K}_{\Delta t}$ thus solves the generalized eigenvalue equation, $\mathbf{Z}'\vec{v} = \lambda\mathbf{Z}\vec{v}$. Equations (40) and (41) already point to the enhanced noise-resilience of our approach, since statistical errors on the observables $g_i(|\phi_k\rangle)$ become tighter when passed onto the functional matrix elements $\mathbf{Z}'_{ij}$ and $\mathbf{Z}_{ij}$. In other words, the variance in the matrix elements $\mathbf{Z}'_{ij}$ and $\mathbf{Z}_{ij}$ is reduced by a factor of $\mathcal{O}(K^{-1})$ relative to the variance in the measured observables due to central-limit scaling. For example, independent Gaussian errors $\mathcal{N}(0, \epsilon^2)$ on each observable combine to make an error of leading order $\mathcal{N}(\epsilon^2, K^{-1}\epsilon^2)$ and $\mathcal{N}(0, K^{-1}\epsilon^2)$ on $\mathbf{Z}_{ij}$ for the respective cases $i = j$ and $i \neq j$.

We remark that for a general reference state $|\phi_0\rangle$ and any finite K, the Koopman operator is non-normal and $\langle f_n, f_m\rangle_\mu \neq 0$. The canonical Krylov bounds on spectral convergence do not strictly apply here as they would require properties of orthogonal polynomials; nevertheless, we showed a rapid convergence of the ODMD estimation using pure algebraic tools in Section 4.

5.3 Matrix pencil methods

We delve deeper into a common framework connecting our data-driven strategy to various well-known methods in signal processing such as the generalized matrix pencil method [63], the Eigensystem Realization Algorithm (ERA) [64, 65] and Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) [66, 67]. We do so by explicitly showcasing an isomorphism between the Hankel matrix pencil of Eq. (31) used in ODMD and other matrix pencil approaches that are extensively used across scientific fields.

To elucidate the direct correspondence between DMD and matrix pencil methods used for solving Eq. (31), we consider the singular value decomposition (SVD) of $\mathbf{X} = U\Sigma V^\dagger$ from Eq. (13). In terms of the SVD, we obtain a new matrix isospectral with the system matrix A,

$$A = \mathbf{X}'\mathbf{X}^+ \cong \Sigma^{-1/2}U^\dagger\mathbf{X}'V\Sigma^{-1/2}, \tag{46}$$

where $\cong$ denotes a spectral congruence that preserves the eigenvalues as any similarity transform leaves the spectrum invariant. The formulation of an eigenvalue problem with respect to the RHS of Eq. (46) is known as the ERA method [64, 65, 68–70] for system identification, which shows an equivalence to the DMD approach on the LHS.

Next, we consider a simultaneous factorization of the data and shifted data matrix $\mathbf{X}$ and $\mathbf{X}'$ using the SVD of the base Hankel matrix,

$$\mathbf{X}_\text{b} = \begin{bmatrix}\mathbf{X}_0 & \mathbf{X}_{1:K} & \mathbf{X}_{K+1}\end{bmatrix} = U_\text{b}\Sigma_\text{b}\begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix}^\dagger, \tag{47}$$

for which $\mathbf{X} = [\mathbf{X}_0\ \mathbf{X}_{1:K}]$, $\mathbf{X}' = [\mathbf{X}_{1:K}\ \mathbf{X}_{K+1}]$, and Σ_b constitutes three singular value blocks consistent with

the partition. Since $\mathbf{X} = U_\mathrm{b}\Sigma_\mathrm{b}V_\mathbf{X}^\dagger = U_\mathrm{b}\Sigma_\mathrm{b}[V_1^\dagger \; V_2^\dagger]$ and $\mathbf{X}' = U_\mathrm{b}\Sigma_\mathrm{b}V_{\mathbf{X}'}^\dagger = U_\mathrm{b}\Sigma_\mathrm{b}[V_2^\dagger \; V_3^\dagger]$, we have

$$A = U_\mathrm{b}\Sigma_\mathrm{b}V_{\mathbf{X}'}^\dagger(V_\mathbf{X}^\dagger)^+\Sigma_\mathrm{b}^{-1}U_\mathrm{b}^\dagger, \tag{48}$$

$$\cong \left[V_\mathbf{X}^\dagger(V_{\mathbf{X}'}^\dagger)^+\right]^\top = \left[(V_{\mathbf{X}'}^\dagger)^\top\right]^+(V_\mathbf{X}^\dagger)^\top, \tag{49}$$

where the congruence follows from invariance of the spectrum under matrix transpose. This SVD-based formulation of the eigenvalue problem associated with the RHS of Eq. (49) is widely recognized as the ESPRIT method [66, 67] for signal recovery.

Algorithms	Real-time data	Tunable hyperparameters
ODMD	$d \times (T-d)$ Hankel	SVD threshold $\tilde{\delta}$
QCELS [24]	$T \times 1$ time series	Number of target modes
(U)VQPE [22]	$T \times T$ Toeplitz	SVD threshold $\tilde{\delta}$

Table 1: Summary of post-processing data structures and tunable hyperparameters for ODMD, QCELS, VQPE, and UVQPE. Each method processes the real-time data, assumed to be overlap matrix elements collected over T time steps, in different formats and involves different hyperparameters that influence accuracy and stability. For ODMD, we set $d = \lfloor \frac{T}{3} \rfloor$.

6 Numerical demonstration

We now empirically demonstrate ODMD's suitability for extracting spectral information from noisy measurements obtained on quantum hardware. Specifically, we demonstrate that our method achieves systematic convergence to the target many-body ground states and corresponding energies even with a large degree of perturbative noise. Our numerical experiments follow the prescription outlined in Sections 3.2 and 3.3 above.

Our test problems belong to two classes of many-body systems. First, we consider the Heisenberg model of interacting quantum spins arranged on a one-dimensional chain. This is a prototypical model describing the effects of strong interactions between electronic spins, and the model is commonly used in condensed matter physics for benchmarking new computational methods. Second, we estimate the ground state energies of strongly-correlated molecular Hamiltonians, a task of substantial pragmatic importance in chemical and materials sciences. We consider molecules of varying complexity, ranging from the relatively simple LiH to the more challenging Cr_2 (displayed in Appendix D and Section 6.2, respectively).

In addition to investigating the performance of ODMD, we also examine recently-developed methods that use similar types of measurements and resources, including subspace expansion techniques exemplified by the Variational Quantum Phase Estimation (VQPE) and its unitary modification (UVQPE) [22], as well as optimization-based techniques represented by the Quantum Complex Exponential Least Squares (QCELS) [24, 25]. A numerical comparison of the different approaches is a crucial initial step in quantifying the relative performance of different real-time algorithm classes, each developed for unique purposes and aimed at different strategies to leverage the capabilities of current quantum hardware. Out of the four methods considered, we illustrate that ODMD consistently reaches the target accuracy with both the fewest time steps and the fewest measurements, especially when the initial state has a small overlap with the target eigenstate.

For numerical efficiency, in each case we select the time step Δt (to be specified within the captions of the following figures) and/or scale our target Hamiltonian H such that $\|H\|_2\Delta t \leq 3\pi/4$. We recall that the ODMD data matrices constructed from a total of T time steps are of dimension $d \times (T-d)$ with $d = \lfloor \frac{T}{3} \rfloor$, a particular choice considered in Section 3.2. For comparison, VQPE and UVQPE process the same real-time data as a pair of $T \times T$ Toeplitz matrices, while QCELS treats it as a length-T time series. To simulate the noise effect, we first compute the exact overlap of Eq. (10), then we introduce Gaussian errors $\mathcal{N}(0, \epsilon^2)$ on the overlap measurements. According to the central limit theorem, the Gaussian noise model correctly captures effects of the shot noise, *i.e.*, the statistical error due to taking a finite number of measurements that scale as $\mathcal{O}(\epsilon^{-2})$ in the Hadamard test. Some types of hardware noise may also exhibit approximately Gaussian behavior, though we do not investigate specific hardware noise models in this section. To mitigate the impact of the added noise, we apply singular-value thresholding as described at the end of Section 3.2 for ODMD; an SVD-based thresholding strategy is also applicable to VQPE and UVQPE [22], which we adopt to promote optimized numerical stability across all real-time methods.

We summarize these structural and parametric aspects in Table 1. We remark that the precise noise variance ϵ^2 is not directly accessible in practice, as this would require an infinite number of measurement shots. However, if ϵ is known up to its order of magnitude, we find empirically that choosing $\tilde{\delta}/\epsilon = 10$ for $\epsilon \lesssim 10^{-2}$ typically leads to an optimal performance for ODMD, VQPE, and UVQPE. Moreover, throughout this section, we primarily consider reference states that do not overlap dominantly with the low energy eigenstates, where the dominance condition is given by $\min_{n\in\mathcal{S}} p_n > 1 - \sum_{n\in\mathcal{S}} p_n$ for target eigenindices $\mathcal{S} \subset \{0, 1, \ldots, N-1\}$. This condition is satisfied only if $\mathcal{S}$ approaches the full set of eigenindices, which corresponds to an intractable classical cost because of the exponential scaling of the nonlinear optimization with the number of target modes. In the few-mode regime, QCELS lacks a convergence guarantee. Hence, we restrict our attention to the single-mode QCELS for ground state estimation without loss of generality.

In Sections 6.1 to 6.3, we present numerical results

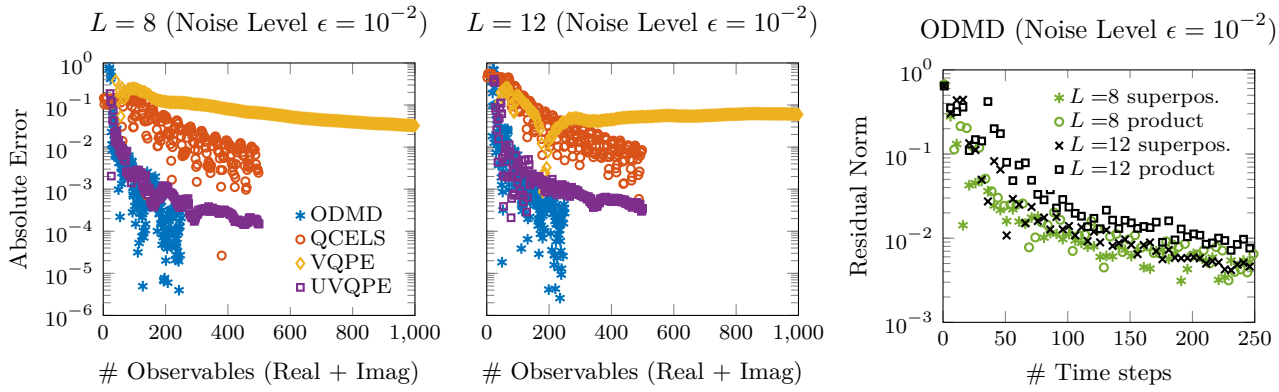

Figure 5: Convergence of ODMD compared with state-of-the-art hybrid algorithms based on real-time evolution (ODMD, QCELS, VQPE, UVQPE) for extracting ground state information of Heisenberg spin chain from overlap observables, $\langle\phi_0|e^{-iHk\Delta t}O|\phi_0\rangle$ with $O = I$ or H. The maximum number of time steps shown is the same for each algorithm; however, the maximum number of observables can differ, for example because we count real and imaginary parts separately and ODMD measures only the real part. **Left**: Convergence of the ground state energy for $L = 8$. For the reference state $|\phi_0\rangle$ used for generating the overlaps, we use the product state $|01010101\rangle$. The noise level ϵ, singular value threshold $\tilde{\delta}$ (for ODMD, VQPE, UVQPE), reference-ground overlap probability $p_0 = |\langle\psi_0|\phi_0\rangle|^2$, and timestep Δt are $(10^{-2}, 10^{-1}, 0.16, 0.15)$. **Center**: Convergence of the ground state energy for $L = 12$ with $(\epsilon, \tilde{\delta}, p_0, \Delta t) = (10^{-2}, 10^{-1}, 0.15, 0.1)$. For the initial reference state, we use the superposition of two product states specified in Section 6.1. **Right**: Convergence of ground state eigenvector found with ODMD, quantified by the residual norm. We show the convergence for four cases: $L = 8$ with product and superposition initial states (16% and 32% overlaps, respectively), and $L = 12$ with product and superposition initial states (7% and 15%).

on spin and molecular systems for benchmarking the performance of ODMD. All dataset and codes are available through the repository: ODMD.

6.1 Eigenstates and eigenenergies of spin systems

To demonstrate the usefulness of ODMD in condensed matter physics, we consider the Heisenberg spin model in 1D,

$$H = J \sum_{i=0}^{L-1} \mathbf{S}_i \cdot \mathbf{S}_{i+1}, \tag{50}$$

where $J > 0$ is the antiferromagnetic coupling strength and $\mathbf{S}_i$ is the spin-1/2 operator $\frac{\hbar}{2}(\sigma^x, \sigma^y, \sigma^z)$ on site i, with the standard Pauli operators $\boldsymbol{\sigma}$. We take periodic boundary conditions and examine systems of size $L = 8$ and $L = 12$ spins.

The spectrum of the 1D Heisenberg Hamiltonian is bounded by $-3LJ < \langle\phi|H|\phi\rangle \leq LJ$ for any state $|\phi\rangle$ [2]. Accordingly, we compute real-time overlaps in the two cases with time step sizes of $\Delta t = 0.15$ and 0.1, respectively, satisfying the bound from Eq. (15). For numerical convenience we have set $J = 4$ and $\hbar = 1$, so that $J\mathbf{S} \cdot \mathbf{S}$ becomes simply $\boldsymbol{\sigma} \cdot \boldsymbol{\sigma}$. We use two different types of initial states: an antiferromagnetic product state $|\phi_0\rangle = |01\rangle^{\otimes L/2}$ and an equal superposition of two such product states, $|\phi_0\rangle = (|01\rangle^{\otimes L/2} + |10\rangle^{\otimes L/2})/\sqrt{2}$. Both have reasonable overlap with the true ground state; the overlap for the superposition state is exactly twice that of the product state. We compute the expectation values $s(k\Delta t) = \langle\phi_0|e^{-iHk\Delta t}|\phi_0\rangle$ for $0 \leq k \leq 250$, and after adding Gaussian errors, we feed the noisy data into ODMD as well as several of the advanced hybrid methods discussed in Section 2. For SVD thresholding, we adopt $\tilde{\delta}/\epsilon = 10$ for a given noise level ϵ.

We show the convergence of the approximate ground state energies $\tilde{E}_0$ in Fig. 5, left and center panels, where the reference energies come from exact Hamiltonian diagonalization. First, we explain the horizontal axis. All four algorithms we test, namely ODMD, QCELS, VQPE, and UVQPE, take as input a list of measured observables $\langle\phi_0|e^{-iHk\Delta t}|\phi_0\rangle$ for $k = 1, 2, \ldots$; VQPE additionally uses the Hamiltonian matrix elements $\langle\phi_0|e^{-iHk\Delta t}H|\phi_0\rangle$. One way to measure the cost of running one of these algorithms would be to count the number of time steps used, *i.e.*, the number of integers k at which the expectation values are measured. However, a more relevant estimate of the cost is the total number of circuits that need to be run on a quantum computer.

If we use the Hadamard test to measure the expectation values $\langle U_k\rangle$ of $U_k = e^{-iHk\Delta t}$, we need to run repeated shots of two distinct circuits, one for the real part and the other for the imaginary part. Hence ODMD, which needs only the real part, actually requires half as many circuits as QCELS or UVQPE. We show this in the figure by plotting against the number of observables, separating the real and imaginary parts. We show the same number of time steps

[2] The lower bound comes from the sum of singlet energies on every bond, which is lower than the true ground state energy. The upper bound comes from a fully ferromagnetic state with all spins up.

for each algorithm, which is why the ODMD results (blue stars) terminate only at 250 on the horizontal axis while QCELS (red circles) and UVQPE (purple squares) go to 500. Thus our figure contains both metrics of convergence speed. By comparing in a single vertical column, we see the relative speed of convergence of the algorithms as a function of the number of circuits to be run, *i.e.*, the practical cost. By comparing the end of the blue markers at 250 to the end of the purple or red markers at 500, we see the relative speed of convergence as a function of the number of time steps.

Because VQPE additionally calculates the expectation values of $U_k H$, we count it as requiring four observables per time step: the real and imaginary parts of $\langle U_k \rangle$ and of $\langle U_k H \rangle$. In fact, this underestimates the true cost as the many-body Hamiltonians are generally non-unitary. Thus $\langle U_k H \rangle$ will require either multiple circuits for each of the real and imaginary parts or multiple ancillas for methods such as block encoding [71]. Since QCELS and UVQPE double the number of observables, while VQPE increases it by at least a factor of four, the total number of measurement shots required to achieve ϵ-precision across all observables scales accordingly.

ODMD results show rapid convergence despite a significant level of noise. The convergence is slightly faster than with UVQPE, though the latter converges reasonably well. In contrast, QCELS, which is more reliant on the overlap between the initial state and the true ground state, converges more slowly. VQPE may not converge at all. The comparison with VQPE in particular highlights the noise-resilience of ODMD; with no noise (including no statistical noise from sampling the wavefunction, *i.e.*, infinite shots), VQPE also converges to the correct energy. However, VQPE involves an ill-conditioned generalized eigenvalue problem, so that even with thresholding strategies it fails to converge in the presence of moderate noise. Such instability is not observed for ODMD. We also notice that ODMD's performance is robust when the initial state overlap is relatively small.

Now we investigate the convergence of the corresponding ODMD state approximation $|\tilde{\psi}_0\rangle$ to the ground state $|\psi_0\rangle$. We measure convergence using the residual norm, $\|H|\tilde{\psi}_0\rangle - \tilde{E}_0|\tilde{\psi}_0\rangle\|_2/\|H\|_2$. We note that although ODMD does not directly prepare the ground state, the residual (up to the normalization factor $\|H\|_2$) can be accessed on the quantum computer. Expanding the square of the residual norm gives

$$\langle\tilde{\psi}_0|H^2|\tilde{\psi}_0\rangle - 2\tilde{E}_0\langle\tilde{\psi}_0|H|\tilde{\psi}_0\rangle + \tilde{E}_0^2, \qquad (51)$$

and each term can be computed from Eq. (20) using real-time evolution. We observe that the residual norm is smaller for initial state with a better ground state overlap; see, for example, the black crosses (15% overlap) versus black squares (7%). On the other hand, the system size does not seem to have a strong impact: the black crosses and green circles correspond to different system sizes, with approximately the same ground state overlap, and behavior similarly as a function of time step.

6.2 Eigenenergies of molecular systems

We focus on the ground state energy calculation of Cr_2, a transition metal dimer known for exhibiting among the strongest electronic correlations. Here we work with a truncated sector of the molecular Hilbert space by employing the adaptive sampling configuration interaction algorithm (ASCI) [72, 73] (see Appendix D for further details). Figure 6 shows the convergence of the real-time algorithms and illuminates the efficacy and noise-resilience of the ODMD approach. For these calculations, we have rescaled the Hamiltonian so that its eigenenergies fall in the range $[-\pi/4, \pi/4]$, enabling us to choose $\Delta t < 4$. The three panels display the convergence of estimated ground state energy; note that in the figures we use the original (unscaled) energy scale in order to show where the energy estimate reaches chemical accuracy. We examine both small ($\Delta t = 1$, left and center panels) and large ($\Delta t = 3$, right panel) time step regimes, with the same number of time steps in each case, favoring shorter time evolution and thus lower circuit depth on the one hand and longer time evolution and thus richer dynamical information on the other.

The left panel of Fig. 6 shows a lower noise level and a reference state with small ground state probability $p_0 = |\langle\psi_0|\boldsymbol{\phi}_0\rangle|^2 = 0.05 = \max_n p_n$. The blue stars marking the ODMD estimation exhibit satisfying performance compared to the purple squares marking the UVQPE estimation that converges variationally to the true ground state energy in the noiseless limit. The slower convergence of QCELS estimation, fitted to a uniform - time step model, results from a rugged optimization landscape as the reference $|\boldsymbol{\phi_0}\rangle$ is almost orthogonal to the target state $|\psi_0\rangle$. VQPE fails to converge to high precision even with this small level of noise due to the ill-conditioning of the generalized eigenvalue problem when the initial state overlap is small.

On the other hand, the right panel shows numerical simulations for a Hartree-Fock reference state with large ground state overlap $p_0 > 0.7$. In this case, even though the noise level is also higher, all the algorithms converge. However, ODMD still has a significant advantage coming from the fact that it needs only half as many observables/circuits per time step.

Finally, the center panel shows an intermediate case with higher noise and small ground state probability. The ODMD estimation is the only one converging past chemical accuracy. Due to a relatively low measurement cost and robust continuous convergence, the ODMD approach emerges as a promising hybrid

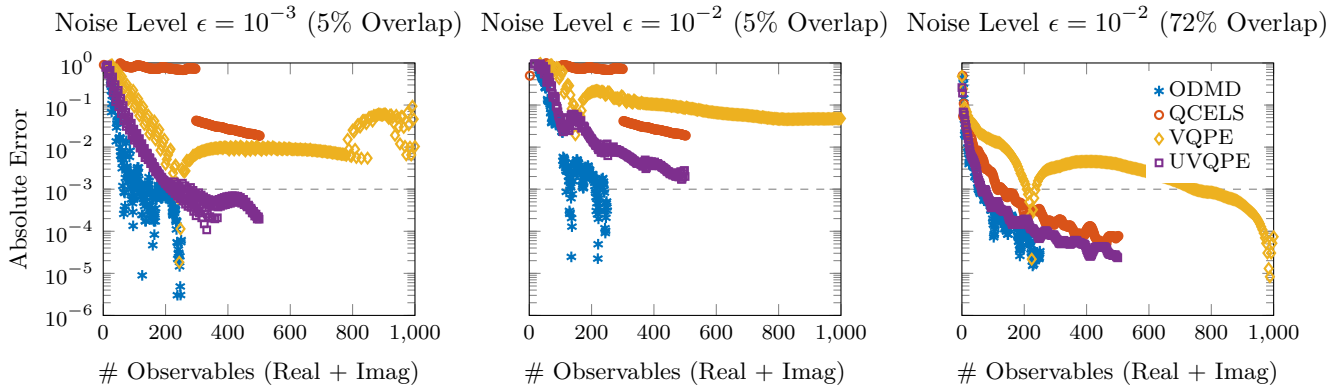

Figure 6: Absolute error in the Cr_2 ground state energy computed from four state-of-the-art hybrid algorithms based on real-time evolution (ODMD, QCELS, VQPE, UVQPE) as a function of the number of observables, $\langle\phi_0|e^{-iHk\Delta t}O|\phi_0\rangle$ with $O = I$ or H, measured. The maximum number of timesteps shown is the same for each algorithm; the maximum number of observables can differ, for example because we count real and imaginary parts separately and ODMD uses only the real part. The three panels exhibit convergence for different values $(\epsilon, \tilde{\delta}, p_0, \Delta t)$ of the noise level ϵ, singular value threshold $\tilde{\delta}$ (for ODMD, VQPE, UVQPE), reference-ground overlap probabilities $p_0 = |\langle\psi_0|\phi_0\rangle|^2$, and timestep Δt. **Left**: $(10^{-3}, 10^{-2}, 5\%, 1)$; **Center**: $(10^{-2}, 10^{-1}, 5\%, 1)$; and **Right**: $(10^{-2}, 10^{-1}, 72\%, 3)$ (for a Hartree-Fock reference state). The horizontal dashed line indicates chemical accuracy.

eigensolver on near-term platforms.

Figure 7 further shows the energy convergence of ODMD for varying noise levels, singular value truncation thresholds, and ground state probabilities. The left panel plots the estimation for different reference states with increasing ground state probability. We observe that for any initial overlap size, the ODMD result progressively improves with additional observables. The right panels show the noise-resilience of ODMD for specific ground state probabilities $p_0 = 0.2$ and $p_0 > 0.7$. The two panels present the number of time steps needed to achieve chemical accuracy, highlighting the compactness of our measurement-driven approach.

6.3 Eigenenergies beyond the ground state

So far we have numerically demonstrated the utility of ODMD for ground state calculations. In particular, we observe that our ODMD ground state estimate converges much more rapidly than VQPE when noise on the overlap matrix elements becomes large, and than QCELS when the initial state lacks a dominant overlap with the ground state. In both regimes, ODMD's significant performance advantage is robust to fine-tuning of the post-processing hyperparameters of the real-time algorithms.

To further substantiate the performance of ODMD relative to UVQPE, we proceed to examine their extended applicability to excited state problems. We first note that although ODMD and UVQPE are constructed distinctly, both can be regarded as a subspace-based approach and thereby face a common challenge: convergence to excited states tends to be more gradual than to the ground state – a limitation typical of many subspace methods [60]. For VQPE and UVQPE to access excited state information, we require multiple reference states $\{|\phi_0^{(i)}\rangle : 1 \leq i \leq D\}$, which readily extends the single-reference method. This leads to the generalized eigenvalue problem $\mathbf{H}\Psi = \tilde{E}\mathbf{S}\Psi$ with a block structure, *i.e.*,

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}^{(11)} & \cdots & \mathbf{H}^{(1D)} \\ \vdots & \ddots & \vdots \\ \mathbf{H}^{(D1)} & \cdots & \mathbf{H}^{(DD)} \end{bmatrix}, \tag{52}$$

and

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}^{(11)} & \cdots & \mathbf{S}^{(1D)} \\ \vdots & \ddots & \vdots \\ \mathbf{S}^{(D1)} & \cdots & \mathbf{S}^{(DD)} \end{bmatrix}, \tag{53}$$

where the individual matrix blocks arising from the pairing of different reference states are defined as,

$$\mathbf{H}_{jk}^{(ii')} = \langle\phi_0^{(i)}|He^{iH(t_j-t_k)}|\phi_0^{(i')}\rangle\,, \tag{54}$$

$$\mathbf{S}_{jk}^{(ii')} = \langle\phi_0^{(i)}|e^{iH(t_j-t_k)}|\phi_0^{(i')}\rangle\,, \tag{55}$$

for $1 \leq i, i' \leq D$. The multi-reference (U)VQPE thus accelerates excited state convergence by initializing with a D-dimensional subspace spanned by the reference states, indicating a runtime *quadratic* in the number of reference states.

Similarly, ODMD admits a novel extension to a multi-observable setting, mirroring a multi-reference subspace approach. To introduce the multi-observable framework, we solve the LS problem $\mathbf{X}' \overset{\text{LS}}{=} A\mathbf{X}$ with a vectorial signal,

$$\mathbf{X} = \begin{bmatrix} \mathbf{s}_0 & \cdots & \mathbf{s}_K \\ \vdots & \ddots & \vdots \\ \mathbf{s}_{d-1} & \cdots & \mathbf{s}_{K+d} \end{bmatrix}, \tag{56}$$

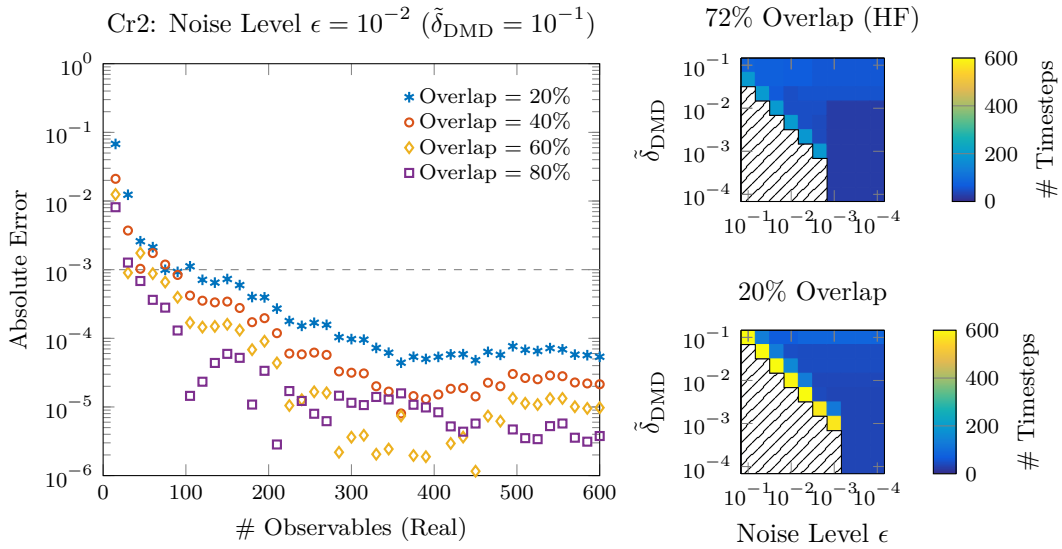

Figure 7: **Left**: Convergence of the ODMD estimate of the ground state energy for Cr_2, with overlap measurements given by the exact result $s(t)$ plus Gaussian noise $\mathcal{N}(0, \epsilon^2)$ with standard deviation $\epsilon = 10^{-2}$. A timestep of $\Delta t = 3$ is used for generating the results. The energy converges rapidly even when the reference state $|\phi_0\rangle$ has only 20% overlap probability with the true ground state. In solving the least-squares problem, we filter singular values below the threshold $\tilde{\delta}_{\text{DMD}} = 10^{-1}$. **Right**: The number of DMD timesteps needed to reach chemical accuracy (horizontal dashed line) as a function of the noise level (horizontal axis) and DMD truncation (vertical axis). We consider ground state probabilities of 72% (Hartree-Fock reference state) and 20%. In the black-dotted regions, absolute error remains above chemical accuracy after 750 timesteps.

and

$$\mathbf{X}' = \begin{bmatrix} \mathbf{s}_1 & \cdots & \mathbf{s}_{K+1} \\ \vdots & \ddots & \vdots \\ \mathbf{s}_d & \cdots & \mathbf{s}_{d+K+1} \end{bmatrix}, \tag{57}$$

where the vector $\mathbf{s}_k$ at the k-th timestep has components,

$$\mathbf{s}_k^{(i)} = \langle \phi_0 | O_i e^{-iHt_k} | \phi_0 \rangle \,, \tag{58}$$

for a set of Hermitian operators $\{O_i = O_i^\dagger : 1 \leq i \leq D\}$, each determining a sequence of time-delayed observables. These operators define a D-dimensional "observable subspace". Under our generalized framework, the observable space formulation implies a runtime *linear* in the number of operators O_i. This key difference in runtime is directly reflected in the number of time evolution circuits required for generating the real-time data.

To evaluate the excited state performance of the multi-observable ODMD and multi-reference UVQPE methods, we consider the transverse-field Ising spin model in 1D,

$$H = -J \sum_{i=0}^{L-1} \sigma_i^z \sigma_{i+1}^z - \sum_{i=0}^{L-1} \sigma_i^x, \tag{59}$$

with ferromagnetic coupling of $J > 0$ and system size of $L = 12$. For numerical convenience, we have set $J = 1$ and rescaled the Hamiltonian spectrum such that $\Delta t = 1$. We start with the single reference state as a superposition of two product states, $|\phi_0\rangle = (3|+\rangle^{\otimes L} + 2|-\rangle^{\otimes L})/\sqrt{13}$ where $|\pm\rangle$ are the ± 1 eigenstates of the Pauli X operator, σ^x. For ODMD, we consider the operators

$$\left\{ O_1 = I,\ O_2 = \sigma_i^z,\ O_3 = \sigma_j^y : i, j \text{ random} \right\}, \tag{60}$$

which compose a 3-dimensional observable subspace. For UVQPE, we initialize with the associated reference states $\{|\phi_0^{(i)}\rangle = O_i|\phi_0\rangle\}_{i=1}^3$. We compute the expectation values using Eqs. (55) and (58) for time steps $0 \leq k \leq 200$, and after adding Gaussian errors, we feed the noisy real-time data into ODMD and UVQPE. For SVD thresholding, we adopt $\tilde{\delta}/\epsilon = 10$ for a given noise level ϵ.

We illustrate the convergence of the approximate first excited state energy $\tilde{E}_1$ in Fig. 8, where we compare it against a reference energy obtained by exact Hamiltonian diagonalization. To assess the quantum resource cost, we follow the same metric used in Sections 6.1 and 6.2 by counting the total number of circuits required on a quantum computer. Specifically, the multi-observable ODMD measures $D = 3$ observables, the real part of $\langle O_i U_k \rangle$, per time step, whereas the multi-reference UVQPE measures $2D^2 = 18$ observables, the real and imaginary parts of $\langle O_i U_k O_{i'} \rangle$, per time step. As visualized in Fig. 8, ODMD converges with fewer total observables – and hence fewer quantum circuits – highlighting its efficiency for excited state estimation.

Finally, we comment that the runtime of multi-observable ODMD may be further reduced to *logarithmic* in the number of observables D. Such favorable reduction is achievable with randomization tech-

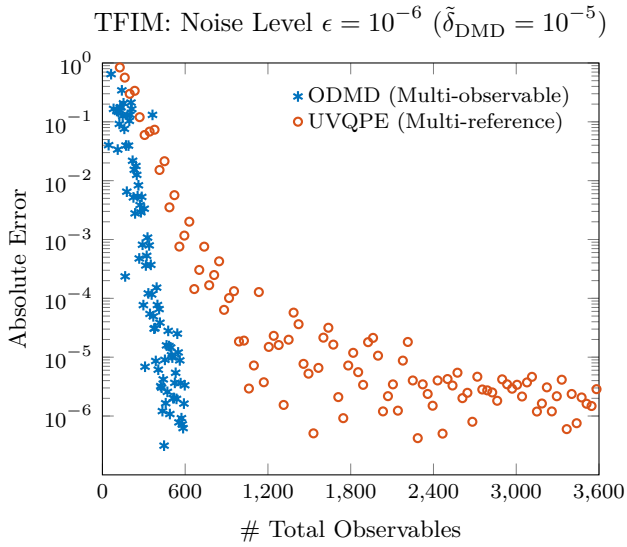

Figure 8: Absolute error in the TFIM first excited state energy computed using two real-time algorithms (ODMD and UVQPE) as a function of the total number of observables, $\langle \phi_0 | O_i e^{-iHk\Delta t} | \phi_0 \rangle$ (with $O_1 = I$ for ODMD) and $\langle \phi_0^{(i)} | e^{-iHk\Delta t} | \phi_0^{(i')} \rangle$ (with $|\phi_0^{(1)}\rangle = |\phi_0\rangle$ for UVQPE), measured. The maximum number of timesteps shown is the same for each algorithm; the maximum number of observables can differ, due to ODMD's observable-based formulation versus UVPQE's state-based formulation. The noise level ϵ, singular value threshold $\tilde{\delta}$, and timestep Δt are $(\epsilon, \tilde{\delta}, \Delta t) = (10^{-6}, 10^{-5}, 1)$. We start with the reference state $|\phi_0\rangle = (3\,|+\rangle^{\otimes L} + 2\,|-\rangle^{\otimes L})/\sqrt{13}$. For ODMD we use the additional observables $O_2 = \sigma_i^z$ and $O_3 = \sigma_j^y$ with random sites i and j while for multi-reference UVQPE we use the corresponding additional reference states $|\phi_0^{(2)}\rangle = O_2\,|\phi_0\rangle$ and $|\phi_0^{(3)}\rangle = O_3\,|\phi_0\rangle$.

niques, as carefully developed in our recent work [74]. In contrast, applying similar randomization techniques to UVQPE still results in a runtime *superlinear* in D. This is because each initial state $|\phi_0^{(i)}\rangle$ requires its own state preparation. Consequently, the ODMD framework lays the foundation for a novel subspace approach capable of simultaneously resolving many-body ground and excited state properties at a significantly lower cost.

7 Discussion

In this work, we developed a hybrid quantum-classical algorithm, ODMD, that leverages real-time evolution to extract ground state information of quantum many-body systems. We proved rigorously that the ODMD estimate converges rapidly to the ground state energy, and we numerically demonstrated robust convergence against noise using Hamiltonians drawn from both condensed matter physics and quantum chemistry. We also related ODMD to a variety of established methods in the numerical linear algebra community.

ODMD presents several distinct advantages over previous hybrid eigensolvers. Most notably, ODMD is more versatile and resilient to perturbative noise. While our framework shares salient perspectives with state-of-the-art subspace and signal processing approaches, it remains effective in regimes where those methods converge slowly or even fail to converge. By formulating a least-squares problem in the space of time-delayed observables, ODMD improves the numerical conditioning and admits a powerful multi-observable generalization best tailored for excited state computation. Moreover, it does not presume a dominant overlap between the initial state and ground state, alleviating the resource overhead associated with state preparation. Consequently, the algorithm converges more reliably as we increase the number of time steps, and also requires fewer shots to sample each measured observable. Additionally, ODMD allows us to recover the eigenstate information and desired eigenstate properties along side the corresponding eigenenergies. Finally, our algorithm only samples the real or imaginary part of the observables, making it even more measurement-efficient for near-term implementation. This last advantage may be adaptable to other real-time based algorithms.

We have further established a unified theoretical foundation for ODMD. We explicitly showed that it can be understood as a Krylov subspace method, which provides an intuitive variational picture to explain convergence to the ground state energy. Also, we demonstrated the connection between ODMD and matrix factorization methods, highlighting the signal processing nature of our algorithm and suggesting a clear interpretation of its noise-resilience.

The ODMD framework hence marks a significant step towards realizing viable real-time evolution methods on quantum hardware, and can be extended

to applications beyond ground state problems.

Acknowledgments

This work was funded by the U.S. Department of Energy (DOE) under Contract No. DE-AC0205CH11231, through the Office of Science, Office of Advanced Scientific Computing Research (ASCR) Exploratory Research for Extreme-Scale Science (YS, DC, KK, SD, RVB) and Accelerated Research for Quantum Computing Programs (AS, DBWY). 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. NMT is grateful for support from NASA Ames Research Center and from the NASA-DOE interagency agreement SAA2-403862.

References

[1] Leon Balents. "Spin liquids in frustrated magnets". Nature **464**, 199–208 (2010).

[2] Yi Zhou, Kazushi Kanoda, and Tai-Kai Ng. "Quantum spin liquid states". Rev. Mod. Phys. **89**, 025003 (2017).

[3] Joel E. Moore. "The birth of topological insulators". Nature **464**, 194–198 (2010).

[4] M. Z. Hasan and C. L. Kane. "Colloquium: Topological insulators". Rev. Mod. Phys. **82**, 3045–3067 (2010).

[5] J. G. Bednorz and K. A. Müller. "Possible hightc superconductivity in the ba-la-cu-o system". Z. Phys. B. Con. Mat. **64**, 189–193 (1986).

[6] Yoichi Kamihara, Takumi Watanabe, Masahiro Hirano, and Hideo Hosono. "Iron-based layered superconductor La[$O_{1-x}F_x$]FeAs (x = 0.05-0.12) with T_c = 26 K". J. Am. Chem. Soc. **130**, 3296–3297 (2008).

[7] B. Keimer, S. A. Kivelson, M. R. Norman, S. Uchida, and J. Zaanen. "From quantum matter to high-temperature superconductivity in copper oxides". Nature **518**, 179–186 (2015).

[8] Markus Reiher, Nathan Wiebe, Krysta M. Svore, Dave Wecker, and Matthias Troyer. "Elucidating reaction mechanisms on quantum computers". Proc. Natl. Acad. Sci. **114**, 7555–7560 (2017).

[9] Akshay Ajagekar and Fengqi You. "Molecular design with automated quantum computing-based deep learning and optimization". Npj Comput. Mater. **9**, 143 (2023).

[10] Ulrich Schollwöck. "The density-matrix renormalization group in the age of matrix product states". Ann. Physics **326**, 96–192 (2011).

[11] Román Orús. "A practical introduction to tensor networks: Matrix product states and projected entangled pair states". Ann. Physics **349**, 117–158 (2014).

[12] A. Yu. Kitaev. "Quantum measurements and the Abelian stabilizer problem" (1995). arXiv:9511026 [quant-ph].

[13] Daniel S. Abrams and Seth Lloyd. "Simulation of many-body fermi systems on a universal quantum computer". Phys. Rev. Lett. **79**, 2586–2589 (1997).

[14] Daniel S. Abrams and Seth Lloyd. "Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors". Phys. Rev. Lett. **83**, 5162–5165 (1999).

[15] Jarrod R. McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. "The theory of variational hybrid quantum-classical algorithms". New J. Phys. **18**, 023023 (2016).

[16] Robert M. Parrish, Edward G. Hohenstein, Peter L. McMahon, and Todd J. Martínez. "Quantum computation of electronic transitions using a variational quantum eigensolver". Phys. Rev. Lett. **122**, 230401 (2019).

[17] Marco Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, et al. "Variational quantum algorithms". Nat. Rev. Phys. **3**, 625–644 (2021).

[18] Bauman Nicholas P, Jaroslav Chládek, Libor Veis, Jiří Pittner, and Kowalski Karol. "Variational quantum eigensolver for approximate diagonalization of downfolded hamiltonians using generalized unitary coupled cluster ansatz". Quantum Sci. Technol. **6**, 034008 (2021).

[19] Jarrod R. McClean, Sergio Boixo, Vadim N Smelyanskiy, Ryan Babbush, and Hartmut Neven. "Barren plateaus in quantum neural network training landscapes". Nat. Commun. **9**, 4812 (2018).

[20] Samson Wang, Enrico Fontana, M. Cerezo, Kunal Sharma, Akira Sone, Lukasz Cincio, and Patrick J. Coles. "Noise-induced barren plateaus in variational quantum algorithms". Nat. Commun. **12**, 6961 (2021).

[21] Zoë Holmes, Andrew Arrasmith, Bin Yan, Patrick J. Coles, Andreas Albrecht, and Andrew T. Sornborger. "Barren plateaus preclude learning scramblers". Phys. Rev. Lett. **126**, 190501 (2021).

[22] Katherine Klymko, Carlos Mejuto-Zaera, Stephen J Cotton, Filip Wudarski, Miroslav Urbanek, Diptarka Hait, Martin Head-Gordon, K. Birgitta Whaley, Jonathan Moussa, Nathan Wiebe, et al. "Real-time evolution for ultracompact Hamiltonian eigenstates on quantum hardware". PRX Quantum **3**, 020323 (2022).

[23] Yizhi Shen, Katherine Klymko, James Sud, David B. Williams-Young, Wibe A. de Jong, and Norm M. Tubman. "Real-Time Krylov Theory

for Quantum Computing Algorithms". Quantum **7**, 1066 (2023).

[24] Zhiyan Ding and Lin 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).

[25] Zhiyan Ding and Lin Lin. "Simultaneous estimation of multiple eigenvalues with short-depth quantum circuit on early fault-tolerant quantum computers". Quantum **7**, 1136 (2023).

[26] William Kirby, Mario Motta, and Antonio Mezzacapo. "Exact and efficient Lanczos method on a quantum computer". Quantum **7**, 1018 (2023).

[27] William J. Huggins, Joonho Lee, Unpil Baek, Bryan O'Gorman, and K. Birgitta Whaley. "A non-orthogonal variational quantum eigensolver". New J. Phys. **22**, 073009 (2020).

[28] Unpil Baek, Diptarka Hait, James Shee, Oskar Leimkuhler, William J. Huggins, Torin F. Stetina, Martin Head-Gordon, and K. Birgitta Whaley. "Say no to optimization: A nonorthogonal quantum eigensolver". PRX Quantum **4**, 030307 (2023).

[29] Igor Mezić and Andrzej Banaszuk. "Comparison of systems with complex behavior". Phys. D: Nonlinear Phenom. **197**, 101–133 (2004).

[30] Igor Mezić. "Spectral properties of dynamical systems, model reduction and decompositions". Nonlinear Dyn. **41**, 309–325 (2005).

[31] Clarence W. Rowley, Igor Mezić, Shervin Bagheri, Philipp Schlatter, and Dan S. Henningson. "Spectral analysis of nonlinear flows". J. Fluid Mech. **641**, 115 – 127 (2009).

[32] PETER J. Schmid. "Dynamic mode decomposition of numerical and experimental data". J. Fluid Mech. **656**, 5–28 (2010).

[33] Igor Mezić. "Analysis of fluid flows via spectral properties of the Koopman operator". Annu. Rev. Fluid Mech. **45**, 357–378 (2013).

[34] Bernard O. Koopman. "Hamiltonian systems and transformation in Hilbert space.". Proc. Natl. Acad. Sci. USA **17 5**, 315–8 (1931).

[35] G. D. Birkhoff and B. O. Koopman. "Recent contributions to the ergodic theory". Proc. Natl. Acad. Sci. **18**, 279–282 (1932).

[36] Steven. L Brunton, Bingni. W Brunton, Joshua. L Proctor, and J. Nathan Kutz. "Koopman observable subspaces and finite linear representations of nonlinear dynamical systems for control". PloS one**11** (2015).

[37] Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. "Discovering governing equations from data by sparse identification of nonlinear dynamical systems". Proc. Natl. Acad. Sci. **113**, 3932–3937 (2016).

[38] Hassan Arbabi and Igor Mezić. "Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the Koopman operator". SIAM J. Appl. Dyn. Syst. **16**, 2096–2126 (2017).

[39] Zhiyan Ding, Haoya Li, Lin Lin, HongKang Ni, Lexing Ying, and Ruizhe Zhang. "Quantum Multiple Eigenvalue Gaussian filtered Search: an efficient and versatile quantum phase estimation method". Quantum **8**, 1487 (2024).

[40] Robert M. Parrish and Peter L. McMahon. "Quantum filter diagonalization: Quantum eigendecomposition without full quantum phase estimation" (2019). arXiv:1909.08925 [quant-ph].

[41] Nicholas H. Stair, Renke Huang, and Francesco A Evangelista. "A multireference quantum Krylov algorithm for strongly correlated electrons". J. Chem. Theory Comput. **16**, 2236–2245 (2020).

[42] Cristian L. Cortes and Stephen K. Gray. "Quantum Krylov subspace algorithms for ground-and excited-state energy estimation". Phys. Rev. A **105**, 022417 (2022).

[43] Nicholas H. Stair, Cristian L. Cortes, Robert M. Parrish, Jeffrey Cohn, and Mario Motta. "Stochastic quantum Krylov protocol with double-factorized Hamiltonians". Phys. Rev. A **107**, 032414 (2023).

[44] Kazuhiro Seki and Seiji Yunoki. "Quantum power method by a superposition of time-evolved states". PRX Quantum **2**, 010333 (2021).

[45] Mario Motta, Chong Sun, Adrian TK Tan, Matthew J O'Rourke, Erika Ye, Austin J Minnich, Fernando GSL Brandão, and Garnet Kin-Lic Chan. "Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution". Nat. Phys. **16**, 205–210 (2020).

[46] Ethan N. Epperly, Lin Lin, and Yuji Nakatsukasa. "A theory of quantum subspace diagonalization". SIAM J. Matrix Anal. Appl. **43**, 1263–1290 (2022).

[47] Dean Lee. "Quantum techniques for eigenvalue problems". The European Physical Journal A **59**, 275 (2023).

[48] Richard Cleve, Artur Ekert, Chiara Macchiavello, and Michele Mosca. "Quantum algorithms revisited". Proc. Roy. Soc. Lond. A **454**, 339 (1998).

[49] Vojtěch Havlíček, Antonio D. Córcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, and Jay M. Gambetta. "Supervised learning with quantum-enhanced feature spaces". Nature **567**, 209–212 (2019).

[50] David Ruelle and Floris Takens. "On the nature of turbulence". Comm. Math. Phys. **20**, 167–192 (1971).

[51] Floris Takens. "Detecting strange attractors in turbulence". Pages 366–381. Springer Berlin Heidelberg. (1981).

[52] Irene López Gutiérrez, Felix Dietrich, and Christian B. Mendl. "Quantum process tomography of

unitary maps from time-delayed measurements". Quant. Inf. Proc. **22**, 251 (2023).

[53] Niladri Gomes, Jia Yin, Siyuan Niu, Chao Yang, and Wibe Albert de Jong. "A hybrid method for quantum dynamics simulation" (2023). arXiv:2307.15231 [quant-ph].

[54] H. F. Trotter. "On the product of semi-groups of operators". Proc. Am. Math. Soc. **10**, 545–551 (1959).

[55] Zhiyan Ding, Ethan N. Epperly, Lin Lin, and Ruizhe Zhang. " The ESPRIT Algorithm Under High Noise: Optimal Error Scaling and Noisy Super-Resolution ". In Proc. Annu. IEEE Symp. Found. Comput. Sci. FOCS. Pages 2344–2366. Los Alamitos, CA, USA (2024). IEEE Computer Society.

[56] Aaron Szasz, Ed Younis, and Wibe Albert de Jong. "Ground state energy and magnetization curve of a frustrated magnetic system from real-time evolution on a digital quantum processor". Quantum **9**, 1704 (2025).

[57] Gaspard Riche de Prony. "Essai experimental et analytique sur les lois de la dilatalrlite de fluids elastiques et sur cells de la vapeur de l'alcool, à différents tempoeratures". Journal de l'Ecole Polytechnique (París) **1**, 24–76 (1795).

[58] H. Padé. "Sur la représentation approchée d'une fonction par des fractions rationnelles". Annales scientifiques de l'École Normale Supérieure **9**, 3–93 (1892). url: `http://eudml.org/doc/81047`.

[59] Gerlind Plonka and Manfred Tasche. "Prony methods for recovery of structured functions". GAMM-Mitt. **37**, 239–258 (2014).

[60] Yousef Saad. "Numerical methods for large eigenvalue problems". Society for Industrial and Applied Mathematics. (2011).

[61] Roger A. Horn and Charles R. Johnson. "Matrix analysis". Cambridge University Press. (1985).

[62] Bernhard Beckermann. "The condition number of real vandermonde, krylov and positive definite hankel matrices". Numer. Math. **85**, 553–577 (2000).

[63] Yingbo Hua and Tapan K. Sarkar. "Generalized pencil-of-function method for extracting poles of an EM system from its transient response". IEEE Trans. Antennas Propag. **37**, 229–234 (1989).

[64] Richard S. Pappa and Jer-Nan Juang. "Galileo spacecraft modal identification using an eigensystem realization algorithm". American Institute of Aeronautics and Astronautics. (1984).

[65] Jer-Nan Juang and Richard S. Pappa. "An eigensystem realization algorithm for modal parameter identification and model reduction". J. Guid. Control Dyn. **8**, 620–627 (1985).

[66] Richard H. Roy and Thomas Kailath. "ESPRIT-estimation of signal parameters via rotational invariance techniques". IEEE Trans. Acoust. Speech Signal Process. **37**, 984–995 (1989).

[67] Tapan K. Sarkar and Odilon Pereira. "Using the matrix pencil method to estimate the parameters of a sum of complex exponentials". IEEE Antennas Propag. Mag. **37**, 48–55 (1995).

[68] Elmer G. Gilbert. "Controllability and observability in multivariable control systems". SIAM J. Appl. Math., Series A: Control **1**, 128–151 (1963).

[69] Rudolf E. Kálmán. "Mathematical description of linear dynamical systems". SIAM J. Appl. Math., Series A: Control **1**, 152–192 (1963).

[70] Juan J. Sanchez-Gasca and Joe H. Chow. "Computation of power system low-order models from time domain simulations using a Hankel matrix". IEEE Trans. Power Syst. **12**, 1461–1467 (1997).

[71] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. "Quantum singular value transformation and beyond: Exponential improvements for quantum matrix arithmetics". In Proc. Annu. ACM Symp. Theory Comput. Page 193–204. STOC 2019New York, NY, USA (2019). Association for Computing Machinery.

[72] Norm M. Tubman, Joonho Lee, Tyler Y. Takeshita, Martin Head-Gordon, and K. Birgitta Whaley. "A deterministic alternative to the full configuration interaction quantum Monte Carlo method". J. Chem. Phys. **145**, 044112 (2016).

[73] Norm M. Tubman, C. Daniel Freeman, Daniel S. Levine, Diptarka Hait, Martin Head-Gordon, and K. Birgitta Whaley. "Modern approaches to exact diagonalization and selected configuration interaction with the adaptive sampling CI method". J. Chem. Theory Comput. **16**, 2139–2159 (2020).

[74] Yizhi Shen, Alex Buzali, Hong-Ye Hu, Katherine Klymko, Daan Camps, Susanne F. Yelin, and Roel Van Beeumen. "Efficient measurement-driven eigenenergy estimation with classical shadows" (2024). arXiv:2409.13691 [quant-ph].

[75] Daniel Greenbaum. "Introduction to quantum gate set tomography" (2015). arXiv:1509.02921 [quant-ph].

[76] Daniel Potts and Manfred Tasche. "Parameter estimation for nonincreasing exponential sums by Prony-like methods". Linear Algebra Appl. **439**, 1024–1039 (2013).

[77] N. Akhiezer. "Über die functionen die in gegebenen intervallen am wenigsten von null abweichen". Izv. Kaz. fiz. mat. ob-va**3** (1928). url: `https://www.mathnet.ru/eng/im/y1933/i4/p499`.

[78] Klaus Schiefermayr. "Chebyshev polynomials on circular arcs". Acta Sci. Math. **85**, 629–649 (2019).

[79] Yingbo Hua and Tapan K. Sarkar. "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise".

IEEE Trans. Acoust. Speech Signal Process. **38**, 814–824 (1990).

[80] P. Jordan and E. Wigner. "Über das Paulische Äquivalenzverbot". Z. Phys. **47**, 631–651 (1928).

[81] M. J. Seaton. "Hartree–fock method". Nature **269**, 631–631 (1977).

[82] Florian Weigend and Reinhart Ahlrichs. "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy". Phys. Chem. Chem. Phys. **7**, 3297–3305 (2005).

[83] J. Stephen Binkley, John A. Pople, and Warren J. Hehre. "Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements". J. Am. Chem. Soc. **102**, 939–947 (1980).

[84] Warren J. Hehre, Robert F. Stewart, and John A. Pople. "Self-consistent molecular-orbital methods. I. Use of Gaussian expansions of Slater-type atomic orbitals". J. Chem. Phys. **51**, 2657–2664 (1969).

A ODMD under general noise model

The depolarizing noise model discussed within Section 3.4 can be extended to more general error models, including Pauli noise, where the channel acts as $\Lambda_t[\rho] = e^{-\mathcal{L}t}[\rho]$ for some Lindbladian,

$$\mathcal{L}[\rho; \{\zeta_i\}] = \mathcal{L}_{\mathrm{U}}[\rho; H] + \mathcal{L}_{\mathrm{D}}[\rho; \{\zeta_i\}], \tag{61}$$

$$= \frac{i}{2}[(I_2 - Z) \otimes H, \rho] - \sum_i \zeta_i \left(M_i \rho M_i^\dagger - \frac{1}{2}\{M_i^\dagger M_i, \rho\} \right), \tag{62}$$

defined by a set of jump operators $\{M_i\}_{i\geq 1}$ with jumping rates $\{\zeta_i\}_{i\geq 1}$. The generators $\mathcal{L}_{\mathrm{U}}$ and $\mathcal{L}_{\mathrm{D}}$ govern the unitary and dissipative components of the dynamics, respectively. For simplicity, we examine the special case of a single Pauli jump operator $M_1 = I_2 \otimes M_{1,\mathrm{sys}}$ where $M_{1,\mathrm{sys}} = M_{1,\mathrm{sys}}^\dagger$ commutes with H. The Lindbladian determines the evolution of the input state as,

$$\Lambda_t[\rho] = \frac{1 + e^{-2\zeta_1 t}}{2} \mathcal{U}_t \rho \mathcal{U}_t^\dagger + \frac{1 - e^{-2\zeta_1 t}}{2} \mathcal{U}_t \tilde{\rho} \mathcal{U}_t^\dagger, \tag{63}$$

where $\tilde{\rho} = |+\rangle\langle+| \otimes |\tilde{\boldsymbol{\phi}}_0\rangle\langle\tilde{\boldsymbol{\phi}}_0|$ represents the noise-corrupted reference state with $|\tilde{\boldsymbol{\phi}}_0\rangle = M_{1,\mathrm{sys}}|\boldsymbol{\phi}_0\rangle$. In this case, the overlap becomes $\frac{1+e^{-2\zeta_1 t}}{2}\Re\langle\boldsymbol{\phi}_0|e^{-iHt}|\boldsymbol{\phi}_0\rangle + \frac{1-e^{-2\zeta_1 t}}{2}\Re\langle\tilde{\boldsymbol{\phi}}_0|e^{-iHt}|\tilde{\boldsymbol{\phi}}_0\rangle$, indicating that the ODMD eigenphases still provide an approximation to the exact eigenenergies. For general jump operator(s), we may instead consider measuring extended observables of the form,

$$o_j(t) = \frac{1}{2}\mathrm{Tr}\left[(X \otimes O_j)\Lambda_t[\rho]\right], \tag{64}$$

for some pool of operators $\{O_j\}_{j\geq 1}$ chosen to efficiently probe the eigenstructure of the Lindbladian $\mathcal{L}$. Suitable choices of $\{O_j\}_{j\geq 1}$ can be guided by the invariant subspaces of the Pauli transfer matrix associated with the Lindbladian [75], which in itself presents a theoretically inspiring direction for future investigation.

B Generalized Extension to Prony's Method

In this appendix, we include a comprehensive proof of the convergence bound from the main text. Prony's method [57] is a pioneering classic approach that has laid the foundation for many modern signal processing and recovery techniques. To understand its underlying idea, let us consider the task of extracting principal frequency components from temporally uniform samples,

$$s_k = \sum_{n=0}^{N-1} p_n e^{-iE_n k\Delta t} = \sum_{n=0}^{N-1} p_n \lambda_n^k = \left(\sum_{\Omega} + \sum_{\Omega^c}\right) p_n \lambda_n^k, \tag{65}$$

where $p_n = |\langle\psi_n|\phi_0\rangle|^2$ so that $\sum_n p_n = 1$, $\Omega \subseteq [N] = \{0, 1, \ldots, N-1\}$ denotes a set of $|\Omega| \ll N$ eigenindices satisfying $\sum_{n\in\Omega} p_n \gg \sum_{n\in\Omega^c} p_n$, and $\Omega^c = [N] \setminus \Omega$ denotes the complement of Ω. For simplicity, let us assume $\Omega = [d]$. Recall that the eigenvalues are the zeros of the characteristic polynomial

$$\mathcal{C}_\Omega(z) = \prod_{\ell\in\Omega}(z - \lambda_\ell) = \sum_{\ell=0}^{d} c_\ell z^\ell, \tag{66}$$

where $\{c_\ell\}_{\ell=0}^d$ are a set of monomial coefficients that depend on $\{\lambda_\ell\}_{\ell\in\Omega}$ with the convention $c_d \equiv 1$. Next, observe that evaluating Eq. (66) at the samples (65), *i.e.*, $\mathcal{C}_\Omega(z = s_k)$ for $k \geq 0$, yields

$$\sum_{\ell=0}^{d} c_\ell s_{\ell+k} = \sum_{\ell=0}^{d} c_\ell \sum_{n=0}^{N-1} p_n \lambda_n^{\ell+k}, \tag{67}$$

$$= \sum_{n=0}^{N-1} p_n \lambda_n^k \sum_{\ell=0}^{d} c_\ell \lambda_n^\ell = \sum_{n=0}^{N-1} p_n \lambda_n^k \mathcal{C}_\Omega(\lambda_n), \tag{68}$$

$$= \sum_{n=d}^{N-1} p_n \lambda_n^k \mathcal{C}_\Omega(\lambda_n) := r_{k,\Omega}, \tag{69}$$

where the sum $\sum_{n<d}$ vanishes since $\mathcal{C}_\Omega(\lambda_n) \equiv 0$ for $n < d$. By using the triangle inequality, the residual term $r_{k,\Omega}$ can be bounded uniformly from above

$$|r_{k,\Omega}| \leq \sum_{n\in\Omega^c} p_n |\mathcal{C}_\Omega(\lambda_n)| \leq \max_{n\in\Omega^c} |\mathcal{C}_\Omega(\lambda_n)| \sum_{n\in\Omega^c} p_n \leq \|\vec{c}\|_1 \sum_{n\in\Omega^c} p_n, \tag{70}$$

so

$$\sum_{\ell=0}^{d-1} c_\ell s_{\ell+k} = -s_{d+k} + r_{k,\Omega} \approx -s_{d+k}, \tag{71}$$

when $\|\vec{c}\|_1 \sum_{n\in\Omega^c} p_n \ll 1$. In matrix form, the evolution of $\{s_k\}_{d\leq k\leq K+d}$ for some $K \in \mathbb{N}$ therefore follows

$$\begin{bmatrix} s_0 & s_1 & \cdots & s_{d-1} \\ s_1 & s_2 & \cdots & s_d \\ \vdots & \vdots & \ddots & \vdots \\ s_K & s_{K+1} & \cdots & s_{K+d-1} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{d-1} \end{bmatrix} = -\begin{bmatrix} s_d \\ s_{d+1} \\ \vdots \\ s_{K+d} \end{bmatrix} + \begin{bmatrix} r_{0,\Omega} \\ r_{1,\Omega} \\ \vdots \\ r_{K,\Omega} \end{bmatrix}, \tag{72}$$

which describes the linear time-invariant dynamics,

$$\underbrace{\begin{bmatrix} s_1 & s_2 & \cdots & s_{K+1} \\ s_2 & s_3 & \cdots & s_{K+2} \\ \vdots & \vdots & \ddots & \vdots \\ s_d & s_{d+1} & \cdots & s_{K+d} \end{bmatrix}}_{\mathbf{X}'} \approx \underbrace{\left[\begin{array}{c|ccc} 0 & & & \\ \vdots & & I_{d-1} & \\ 0 & & & \\ \hline -c_0 & -c_1 & \cdots & -c_{d-1} \end{array}\right]}_{A} \underbrace{\begin{bmatrix} s_0 & s_1 & \cdots & s_K \\ s_1 & s_2 & \cdots & s_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_{d-1} & s_d & \cdots & s_{K+d-1} \end{bmatrix}}_{\mathbf{X}}, \tag{73}$$

where $\mathbf{X}, \mathbf{X}' \in \mathbb{C}^{d\times(K+1)}$ are time-delayed Hankel matrices, $A \in \mathbb{C}^{d\times d}$ is a system matrix of companion structure, and $I_{d-1} \in \mathbb{C}^{(d-1)\times(d-1)}$ denotes the identity matrix. For interested readers, the connection between the

characteristic polynomial $\mathcal{C}_\Omega$ and the associated linear recurrence relation above are explicitly discussed in Lemma 2.2 from Ref [76]. Prony's method essentially solves for the monomial coefficients $\vec{c}$ and thus the eigenfrequencies $\{\lambda_n\}_{n\in\Omega}$ as the roots of $\mathcal{C}_\Omega$ when Eqs. (72) and (73) hold exact for $N = d$ (so that $r_{k,\Omega} \equiv 0$).

Here, we consider the dynamic mode decomposition (DMD) approach formulated in Eq. (5) as a powerful extension of Prony's method when $d \ll N$, which solves the linear least-squares (LLSQ) problems of the same type, *i.e.*, $\mathbf{X}' = A\mathbf{X}$, where $A = \arg\min_{\tilde{A}\in\mathbb{C}^{d\times d}} \|\mathbf{X}' - \tilde{A}\mathbf{X}\|_F^2$. In other words, the DMD post-processing identifies the set of monomial coefficients $\vec{a} \approx \vec{c}$ that minimize the total residual. Let

$$\mathcal{C}_{\mathrm{DMD}}(z) = \prod_{\ell=0}^{d-1}(z - \tilde{\lambda}_\ell) = \sum_{\ell=0}^{d} a_\ell z^\ell, \tag{74}$$

represent the characteristic polynomial corresponding to the system matrix A, with $a_d \equiv 1$. Then, we define the DMD residual as follows

$$r_{k,\mathrm{DMD}} = \sum_{n=0}^{N-1} p_n \lambda_n^k \mathcal{C}_{\mathrm{DMD}}(\lambda_n), \tag{75}$$

p_n and λ_n are given in Eq. (65). Note that in case $\tilde{\lambda}_n \approx \lambda_n$, for $n = 0, 1, \ldots, d-1$, the contribution of the d first terms in Eq. (75) will remain minimal. In fact, we may now show that the eigenvalues $\{\tilde{\lambda}_\ell\}_{\ell=0}^{d-1}$ of the DMD system matrix A well approximate the target eigenvalues $\{\lambda_n\}_{n\in\Omega}$ for some Ω.

Let $\mathcal{P}_d$ be the set of polynomials of at most degree d and $\mathcal{P}'_d \subset \mathcal{P}_d$ the set of monic polynomials, *i.e.*, $\mathcal{P}'_d = \{\mathcal{C} \in \mathcal{P}_d : \mathcal{C} - z^d \in \mathcal{P}_{d-1}\}$. For any polynomial $\mathcal{C} \in \mathcal{P}_d$, the triangle inequality implies

$$\left|\sum_{n=0}^{N-1} p_n \lambda_n^k \mathcal{C}(\lambda_n)\right| \leq \sum_{n=0}^{N-1} p_n |\mathcal{C}(\lambda_n)|, \tag{76}$$

which allows us to establish a uniform upper bound on the DMD residual

$$\sum_{k=0}^{K} |r_{k,\mathrm{DMD}}|^2 = \min_{\mathcal{C}\in\mathcal{P}'_d} \sum_{k=0}^{K} \left|\sum_{n=0}^{N-1} p_n \lambda_n^k \mathcal{C}(\lambda_n)\right|^2, \tag{77}$$

$$\leq (K+1) \min_{\mathcal{C}\in\mathcal{P}'_d} \left(\sum_{n=0}^{N-1} p_n |\mathcal{C}(\lambda_n)|\right)^2, \tag{78}$$

$$\leq (K+1) \min_{\mathcal{C}\in\mathcal{P}'_d} \max_{0\leq n\leq N-1} |\mathcal{C}(\lambda_n)|^2. \tag{79}$$

We now proceed by saturating a tight bound on the DMD residual through the min-max inequality above. Given our freedom to pick the time increment Δt such that $(E_{N-1} - E_0)\Delta t = 2\Theta \leq 2\pi$, we further bound the RHS of Eq. (79) via polynomials over the circular arc $\mathcal{A}_\Theta = \{z \in \mathbb{C} : |z| = 1, -\Theta \leq \arg(z) \leq \Theta\}$. An example of such circular arc with $\Theta = \pi/2$ is sketched in Fig. 9 below.

In particular, let us consider the choice of complex-valued Chebyshev polynomials,

$$\hat{\mathcal{C}}(z) = \arg\min_{\mathcal{C}\in\mathcal{P}'_d} \sup_{z\in\mathcal{A}_\Theta} |\mathcal{C}(z)|, \tag{80}$$

whose parametric representations can be explicitly constructed from tools such as Jacobi's elliptic and theta functions [77]. For notational convenience, we define $\|\mathcal{C}(z)\|_\Theta = \sup_{z\in\mathcal{A}_\Theta} |\mathcal{C}(z)|$. Similar to real-valued Chebyshev polynomials over the unit interval, $\hat{\mathcal{C}}(z)$ retains the minimal norm property on $\mathcal{A}_\Theta$ with

$$\|\hat{\mathcal{C}}\|_\Theta \sim 2\sin^d(\Theta/2)\cos^2(\Theta/4), \tag{81}$$

asymptotically for large d [78]. Consequently, the difference in residuals associated with the Prony-like and DMD linear systems of the type Eq. (72) can be bounded as,

$$\|\vec{r}_\Omega - \vec{r}_{\mathrm{DMD}}\|_2^2 = \sum_{k=0}^{K} \left|r_{k,\Omega} - r_{k,\mathrm{DMD}}\right|^2, \tag{82}$$

$$\leq 2\sum_{k=0}^{K} |r_{k,\Omega}|^2 + 2\sum_{k=0}^{K} |r_{k,\mathrm{DMD}}|^2, \tag{83}$$

$$\leq 2\sum_{k=0}^{K} \max_{1\leq n\leq N-1} |\mathcal{C}_\Omega(\lambda_n)|^2 + 2\sum_{k=0}^{K} |r_{k,\mathrm{DMD}}|^2, \tag{84}$$

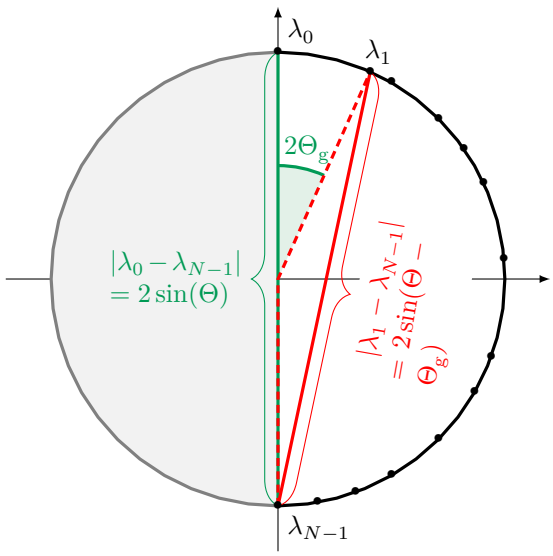

Figure 9: Characteristic polynomial $\mathcal{C}_\Omega$ with roots along the arc $\mathcal{A}_{\Theta=\pi/2}$ and $0 \in \Omega$. The eigenphases $\{\lambda_n\}_{n=0}^{N-1}$ are represented schematically by black dots on the unit circle. Geometrically, $\|\tilde{\mathcal{C}}_{\Omega\setminus\{0\}}\|_{\Theta-\Theta_g}$ is bounded by the solid red chord of length $|\lambda_1 - \lambda_{N-1}| = 2\sin(\Theta - \Theta_g)$, whereas $\|\mathcal{C}_\Omega\|_\Theta$ is bounded by the green solid chord of length $|\lambda_0 - \lambda_{N-1}| = 2\sin(\Theta)$. Our strategy for selecting appropriate reference eigenindex set Ω is aimed at minimizing $\|\tilde{\mathcal{C}}_{\Omega\setminus\{0\}}\|_{\Theta-\Theta_g}/|\partial_z\mathcal{C}_\Omega(\lambda_0)|$, where $|\partial_z\mathcal{C}_\Omega(\lambda_0)| \approx \|\mathcal{C}_\Omega\|_\Theta^{d-1}$ for $\{\lambda_{n_\ell}\}_{\ell=1}^{d-1}$ far away from λ_0. Note that $\tilde{\mathcal{C}}_{\Omega\setminus\{0\}}$ undergoes a global phase shift of $e^{-i\Theta_g}$ relative to $\mathcal{C}_{\Omega\setminus\{0\}}$, which center the arc $\mathcal{A}_{\Theta-\Theta_g}$ around $z = 1$.

where we now work with a generic set $\Omega = \{n_0 = 0, n_1, \cdots, n_{d-1}\}$ of d eigenindices. Eq. (83) follows from the application of triangle and Cauchy–Schwarz inequalities $|z - z'|^2 \le |z|^2 + 2|z||z'| + |z'|^2 \le 2|z|^2 + 2|z'|^2$ while Eq. (84) exploits the upper bound derived in Eq. (70). Using the minimal norm property of the DMD residual, we have

$$\|\vec{r}_\Omega - \vec{r}_{\mathrm{DMD}}\|_2^2 \le 2(K+1) \max_{1\le n\le N-1} |\mathcal{C}_\Omega(\lambda_n)|^2 + 2(K+1) \min_{\mathcal{C}\in\mathcal{P}_d'} \max_{0\le n\le N-1} |\mathcal{C}(\lambda_n)|^2, \tag{85}$$

$$\le 2(K+1)\left[\max_{1\le n\le N-1} |\lambda_n - \lambda_0|^2 \max_{1\le n\le N-1}\left|\prod_{\ell=1}^{d-1}(\lambda_n - \lambda_{n_\ell})\right|^2 + \|\hat{\mathcal{C}}\|_\Theta^2\right], \tag{86}$$

$$\le 2(K+1)\left[\left(2\max_{1\le n\le N-1}\left|\prod_{\ell=1}^{d-1}(\lambda_n - \lambda_{n_\ell})\right|\right)^2 + \|\hat{\mathcal{C}}\|_\Theta^2\right], \tag{87}$$

$$\le 2(K+1)\left[2\|\tilde{\mathcal{C}}_{\Omega\setminus\{0\}}\|_{\Theta-\Theta_g} + \|\hat{\mathcal{C}}\|_\Theta\right]^2, \tag{88}$$

where $\Theta_g = (E_1 - E_0)\Delta t/2$ represents the gap between the dimensionless ground state and first excited state angle. It is worth noting that we can always assume $\Theta_g > 0$ since otherwise $p_0\lambda_0^k + p_1\lambda_1^k = (p_0+p_1)\lambda_0^k \implies N \mapsto N-1$ and $\{E_n\}_{n=1}^{N-1} \mapsto \{E_{n-1}\}_{n=1}^{N-1}$. For notational consistency, we let $\tilde{\mathcal{C}}_{\Omega\setminus\{0\}}(z) = \prod_{\ell=1}^d (e^{-i\Theta_g}z - \lambda_{n_\ell}) \in \mathcal{P}_{d-1}$ denote the polynomial of degree $(d-1)$ with a proper phase shift included. Since $\|\hat{\mathcal{C}}\|_\Theta$ is optimally bounded, next we show that $\|\tilde{\mathcal{C}}_{\Omega\setminus\{0\}}\|_{\Theta-\Theta_g}$ can also be bounded in a controlled way. Specifically, we claim that there exists $(\Theta_0, \Theta_1) \subseteq [0, \pi]$ so that $\forall\Theta \in [\Theta_0, \Theta_1]$, $\|\tilde{\mathcal{C}}_{\Omega\setminus\{0\}}\|_{\Theta-\Theta_g} \ll 2^{d-1}$ asymptotically for a collection of index sets: for concreteness, we examine the example $\Theta = \pi/2$ and $\Omega \setminus \{0\} = \{n_\ell\}_{\ell=1}^{d-1}$ which clusters around the antipodal point of λ_0, *i.e.*, $\lambda_{n_\ell} \approx -\lambda_0$, $\forall\ell \ne 0$. A direct calculation shows that

$$|z - \lambda_{n_\ell}| = 2\sin\left|\frac{\arg(z) - \arg(\lambda_{n_\ell})}{2}\right| \implies \|\tilde{\mathcal{C}}_{\Omega\setminus\{0\}}\|_{\pi/2-\Theta_g} \sim 2^{d-1}\sin^{d-1}(\pi/2 - \Theta_g), \tag{89}$$

Graphically, Fig. 9 helps elucidate the bounded behavior of $\mathcal{C}_\Omega$ and thereby $\tilde{\mathcal{C}}_{\Omega\setminus\{0\}}$ in the presence of the spectral gap Θ_g.

Going forward, for the sake of illustration, let us restrict our attention to any index set Ω that admits near maximal product of distances $\prod_{\ell=1}^{d-1}|\lambda_0 - \lambda_{n_\ell}|$. Recall from Eq. (72) that $\vec{a} - \vec{c} = \mathbf{X}^\top(\vec{r}_\Omega - \vec{r}_{\mathrm{DMD}})$. Thus combining

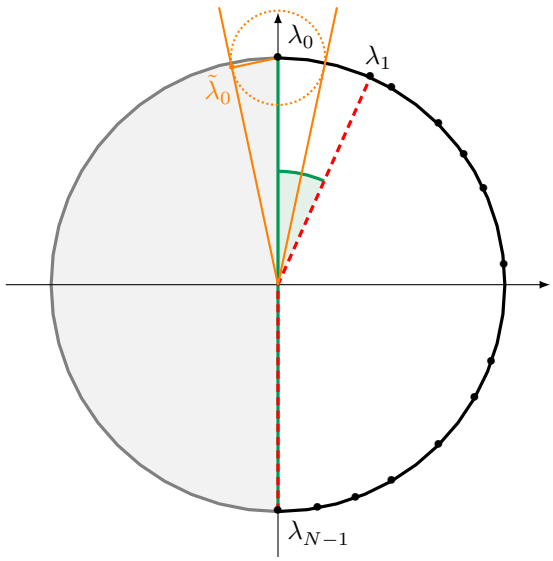

Figure 10: Error in the ground state energy inferred from error in the ground state eigenphase. The eigenphases $\{\lambda_n\}_{n=0}^{N-1}$ are shown schematically as black dots on the unit circle. The dotted orange circle indicates an upper bound on the eigenphase difference $|\tilde{\lambda}_0 - \lambda_0|$. Deviation in the angles $|\tilde{E}_0 - E_0|\Delta t$ can be determined through the construction of tangent lines to the orange circle centered at λ_0.

our previous results, we see that the DMD coefficients $\vec{a}$ and the target coefficients $\vec{c}$ converge towards each other at a rate,

$$\|\vec{a} - \vec{c}\|_2 \lesssim \|(\mathbf{X}^\top)^+\|_2 \|\vec{r}_\Omega - \vec{r}_{\mathrm{DMD}}\|_2, \tag{90}$$

$$\leq \frac{[2(K+1)]^{1/2}\Big\{2\|\tilde{\mathcal{C}}_{\Omega\setminus\{0\}}\|_{\Theta-\Theta_{\mathrm{g}}} + \|\hat{\mathcal{C}}\|_\Theta\Big\}}{\delta}, \tag{91}$$

where recall that $\mathbf{X} \in \mathbb{C}^{d\times(K+1)}$ is the data matrix and $\delta > 0$ denotes a lower bound on its singular values. To monitor the scaling of our DMD approximation, we write $d = \lfloor \alpha(K+1) \rfloor$ for which $0 \leq \alpha \leq 1$ is kept constant throughout (here it suffices to discuss $0 \leq \alpha \leq 1$ due to the Hankel structure of our data matrix). Moreover, previous literature supports the common practice of working with $\alpha \lesssim 1$ as near-optimal choice [63, 79]. The error in the ground state energy can then be estimated and analyzed through first-order perturbation theory,

$$\partial_{c_\ell}\mathcal{C}_\Omega(\lambda_0) = 0,\ \forall \ell \iff \lambda_0^\ell + \sum_{k=1}^{d} c_k k\,(\lambda_0)^{k-1}\,\partial_{c_\ell}\lambda_0 = 0, \tag{92}$$

$$\iff \partial_{c_\ell}\lambda_0 = -\frac{\lambda_0^\ell}{\partial_z \mathcal{C}_\Omega(\lambda_0)}, \tag{93}$$

so we have $\tilde{\lambda}_0 - \lambda_0 = \sum_{\ell=0}^{d-1}(a_\ell - c_\ell)\partial_{c_\ell}\lambda_0$ up to leading order and,

$$|\tilde{\lambda}_0 - \lambda_0| = \left|\frac{\sum_{\ell=0}^{d-1}(a_\ell - c_\ell)\lambda_0^\ell}{\partial_z \mathcal{C}_\Omega(\lambda_0)}\right| \leq \frac{\|\vec{a} - \vec{c}\|_1}{\prod_{\ell\in\Omega\setminus\{0\}} |\lambda_0 - \lambda_\ell|}, \tag{94}$$

$$\leq \frac{(2\alpha)^{1/2}(K+1)\Big\{2\|\tilde{\mathcal{C}}_{\Omega\setminus\{0\}}\|_{\Theta-\Theta_{\mathrm{g}}} + \|\hat{\mathcal{C}}\|_\Theta\Big\}}{\delta \prod_{\ell\in\Omega\setminus\{0\}} 2|\sin(\theta_{0l})|}, \tag{95}$$

where we use the norm inequality $\|\vec{a}-\vec{c}\|_1 \leq d^{1/2}\|\vec{a}-\vec{c}\|_2$ and define the eigenangle variables $\theta_{j\ell} = (E_\ell - E_j)\Delta t/2$. Here we order the DMD eigenvalues $\{\tilde{\lambda}_0, \ldots, \tilde{\lambda}_{d-1}\}$ based on their arguments in a descending order, *i.e.*, $\arg\tilde{\lambda}_0 \geq \arg\tilde{\lambda}_1 \geq \cdots \geq \arg\tilde{\lambda}_{d-1}$, so that $\tilde{\lambda}_0 \approx \lambda_0$ contains the ground state information. For the case $\Theta_0 \leq \Theta \leq \pi/2$, we

note that picking $\theta_{0l} \sim 2\Theta$ gives $|\sin(\theta_{0l})| \sim \sin(\Theta)$. Therefore Eq. (95) can be bounded asymptotically as,

$$|\tilde{\lambda}_0 - \lambda_0| \lesssim \frac{(2\alpha)^{1/2}(K+1)}{\delta 2^{d-1}\sin^{d-1}(\Theta)}\Big\{2\|\tilde{\mathcal{C}}_{\Omega\setminus\{0\}}\|_{\Theta-\Theta_{\rm g}} + \|\hat{\mathcal{C}}\|_{\Theta}\Big\}, \tag{96}$$

$$\leq \frac{2^{3/2}\alpha^{1/2}(K+1)}{\delta 2^{d-1}\sin^{d-1}(\Theta)}\Big\{2^{d-1}\sin^{d-1}(\Theta-\Theta_{\rm g}) + \sin^{d}(\Theta/2)\cos^2(\Theta/4)\Big\}, \tag{97}$$

$$\approx \frac{2^{3/2}\alpha^{1/2}K}{\delta}\left[\frac{\sin^{\alpha}(\Theta-\Theta_{\rm g})}{\sin^{\alpha}(\Theta)}\right]^K + \frac{2^{3/2}\alpha^{1/2}K\sin(\Theta_1/2)\cos^2(\Theta_0/4)}{\delta(2^{\alpha})^K}\left[\frac{\sin^{\alpha}(\Theta/2)}{\sin^{\alpha}(\Theta)}\right]^K, \tag{98}$$

with $K \gg 1$. Since typically $\Theta > 2\Theta_{\rm g}$ for practical Hamiltonians, it follows that the leading term in Eq. (98) decays proportional to $Ke^{-\beta_1 K}$ where $\beta_1(\Theta,\Theta_{\rm g}) = -\alpha\ln[\sin(\Theta-\Theta_{\rm g})/\sin(\Theta)] > 0$ and the subleading term decays proportional to $Ke^{-\beta_2 K}$ where $\beta_2(\Theta) = -\alpha\ln[\sin(\Theta/2)/2\sin(\Theta)] > \beta_1 > 0$. Finally using a simple tangent line argument as sketched in Fig. 10, we have for sufficiently small $|\tilde{\lambda}_0 - \lambda_0|$,

$$|\arg(\tilde{\lambda}_0) - \arg(\lambda_0)| \leq \sin^{-1}|\tilde{\lambda}_0 - \lambda_0| \leq |\tilde{\lambda}_0 - \lambda_0| + \frac{|\tilde{\lambda}_0 - \lambda_0|^3}{2}, \tag{99}$$

leading to an asymptotic error bound on the ground state energy,

$$\delta E_0 = \frac{|\arg(\tilde{\lambda}_0) - \arg(\lambda_0)|}{\Delta t}, \tag{100}$$

$$\leq \frac{2^{3/2}\alpha^{1/2}}{\delta\Delta t}\left[Ke^{-\beta_1 K} + \sin(\Theta_1/2)\cos^2(\Theta_0/4)Ke^{-\beta_2 K}\right] + \text{h.o.t}, \tag{101}$$

which holds for an arbitrary initial state $|\phi_0\rangle$. We remark that the error in the excited state energies can be derived in a similar manner. However, the proof requires careful selection of different index sets Ω (since we want to control $|\partial_z\mathcal{C}_\Omega(z_\ell)|$ from below in the perturbative analysis).

Our theoretical result above suggests that the DMD approach has the potential of adeptly approximating the ground state energy, provided that we have access to the overlap $\{s_k\}_k$ whose real and imaginary parts are measured separately using a quantum computer. In practice, we can trivially replicate the preceding proof by considering only the real (or imaginary) parts of the overlaps, *e.g.*, $\Re s_k = \sum_{n=0}^{N-1} p_n\cos(kE_n\Delta t)$. By Euler's formula $\Re s_k = \sum_{n=0}^{N-1} p_n(\lambda_n^k + \lambda_n^{-k})/2$, we immediately see that,

$$\sum_{\ell=0}^{d} c_\ell s_{\ell+k} = 0 \implies \sum_{\ell=0}^{2d} \underline{c}_\ell \underline{s}_{l+k} = 0, \tag{102}$$

where

$$\sum_{\ell=0}^{2d} \underline{c}_\ell z^\ell = \prod_{\ell=1}^{d}(z-\lambda_{n_\ell})(z-\lambda_{n_\ell}^*), \tag{103}$$

is an extended characteristic polynomial containing conjugate pairs eigenphases (λ_n, λ_n^*) and $\underline{S}_k = \sum_{n=0}^{2N-1}(p_n/2)\lambda_n^k$ is the extended overlap matrix element with $(p_n, E_n) = (p_{n-N}, -E_{n-N})$ for $N \leq n \leq 2N-1$. Hence given the data matrix $\{\Re s_k\}_k$, taking $d \mapsto 2d$ or $\alpha \mapsto 2\alpha$ also effectively recovers the ground state energy in the large K limit.

C Connection to Padé's rational approximation

In this appendix, we present how the DMD approach can be heuristically conceptualized under the framework of Padé approximation [58, 59].

We return to our simple assumption $\Omega = [d]$ and work with the efficient approximating sequence $\{\tilde{s}_k\}_k$ with $\tilde{s}_k = \sum_{n\in\Omega} p_n \lambda_n^k \approx s_k$ (recall that $|\sum_{n\in\Omega^c} p_n \lambda_n^k| \leq \sum_{n\in\Omega^c} p_n \ll 1$). Let us now consider the Z-transform of $\{\tilde{s}_k\}_{k\in\mathbb{N}}$,

$$\mathcal{Z}\big[\{\tilde{s}_k\}\big](z) = \sum_{k=0}^{\infty} \tilde{s}_k z^{-k}, \tag{104}$$

$$= \sum_{n=0}^{d-1} p_n \sum_{k=0}^{\infty} \left(\frac{\lambda_n}{z}\right)^k = \sum_{n=0}^{d-1} p_n \frac{z}{z-\lambda_n}, \tag{105}$$

which converges in the region $\{z \in \mathbb{C} : |z| > \limsup_k |\tilde{s}_k|^{1/k} = 0\}$. The RHS of Eq. (105) consists of a rational function,

$$\sum_{n=0}^{d-1} p_n \frac{z}{z-\lambda_n} = \sum_{n=0}^{d-1} p_n \frac{z \prod_{\ell\neq n\in\Omega}(z-\lambda_\ell)}{\prod_{\ell\in\Omega}(z-\lambda_\ell)} = \frac{\mathcal{B}_\Omega(z)}{\mathcal{C}_\Omega(z)}, \tag{106}$$

where we recognize that $\mathcal{C}_\Omega(z) = \prod_{\ell\in\Omega}(z-\lambda_\ell) = z^d + \sum_{\ell=0}^{d-1} c_\ell z^\ell$ is the characteristic polynomial in Eq. (66) and $\mathcal{B}_\Omega(z) = \sum_{n\in\Omega} p_n z \prod_{\ell\neq n\in\Omega}(z-\lambda_\ell) \approx z^d + \sum_{\ell=1}^{d-1} b_\ell z^\ell$. Associated with these two polynomials is a rational approximation of the power series,

$$\sum_{k=0}^{\infty} \tilde{s}_k z^k = \frac{\mathcal{B}_\Omega(z^{-1})}{\mathcal{C}_\Omega(z^{-1})} \equiv \frac{\mathcal{B}_{\Omega,P}(z)}{\mathcal{C}_{\Omega,P}(z)}, \tag{107}$$

where $\mathcal{B}_{\Omega,P} = z^d \mathcal{B}_\Omega(z^{-1}) \approx 1 + \sum_{\ell=1}^{d-1} b_{d-\ell} z^\ell$ and $\mathcal{C}_{\Omega,P} = z^d \mathcal{C}_\Omega(z^{-1}) = 1 + \sum_{\ell=1}^{d} c_{d-\ell} z^\ell$ constitute the Padé approximant. In our context, Padé approximation finds the best rational approximation $(\mathcal{B}_\Omega, \mathcal{C}_\Omega)$ from the time-accumulated overlap data,

$$\int_0^\infty \frac{dt}{\Delta t} s(t) e^{iEt} \approx \sum_{k=0}^{\infty} \tilde{s}_k \left(e^{iE\Delta t}\right)^k, \tag{108}$$

whose poles contain the target eigenfrequencies $\{\lambda_n^*\}_{n\in\Omega}$ (up to a complex conjugation). Rearranging Eq. (107) above, we have

$$\sum_{k=0}^{\infty} \tilde{s}_k z^k \mathcal{C}_{\Omega,P}(z) = \mathcal{B}_{\Omega,P}(z) \implies \begin{cases} \displaystyle\sum_{\ell=d-k}^{d} c_\ell \tilde{s}_{\ell+k-d} = b_{d-k}, & k \in [d] \\ \displaystyle\sum_{\ell=0}^{d} c_\ell \tilde{s}_{\ell+k} = 0, & k \in \mathbb{N} \end{cases}, \tag{109}$$

upon matching z^k terms on the LHS and the RHS order by order. The first set of derived relations reveals an inherent linkage of the monomial coefficients $\vec{c}$ obtained from the characteristic polynomial $\mathcal{C}_\Omega(z)$ to the set of coefficients $\vec{b}$ used in Padé's construction, whereas the second set of relations simply reiterate the linear dynamics modeled in the DMD approach.

D Ground State Energy Estimation of Molecular Hamiltonians

Here we provide additional data and details for the range of molecular Hamiltonians considered in this study. We recall that a general electronic Hamiltonian takes the form,

$$H = \sum_{ij} h_{ij} a_i^\dagger a_j + \frac{1}{2} \sum_{ijk\ell} a_i^\dagger a_j^\dagger a_k a_\ell, \tag{110}$$

where a_i and $a_i^\dagger$ denote the fermionic annihilation and creation operators respectively in the second quantized picture, while i and j index the sets of virtual and occupied electron spin-orbitals accordingly. The electronic Hamiltonian can be systematically mapped onto its qubit representation via techniques such as the Jordan-Wigner transformation. [80] An advantageous reference state for molecular problems is the Hartree-Fock ground state, $|\phi_0\rangle = |\phi_{\mathrm{HF}}\rangle$, derived from a self-consistent mean-field calculation. [81] Under the Jordan-Wigner transformation,

$$|\phi_{\mathrm{HF}}\rangle = |0\rangle^{\otimes(N_{\mathrm{orb}}-N_{\mathrm{el}})} |1\rangle^{\otimes N_{\mathrm{el}}}, \tag{111}$$

where N_{orb} and N_{el} denote the number of spin-orbitals and electrons. Thus $|\phi_{\mathrm{HF}}\rangle$ takes the simple form of a product state with each qubit indicating the occupancy of a spin-orbital.

For Cr_2 dimer (both here and in the main text), we work with the def2-SVP basis set [82] (bond length 1.5 Å, 30 orbitals, and 24 electrons). We restrict our simulation to the widely studied 30 orbital active space and truncate the Hilbert space, considering only a subset of Slater determinants that we choose using ASCI [72, 73]. All data shown in the main text and here are a truncation of size 4000, selected from a one million determinant ASCI calculation where the 4000 determinants with the largest coefficients are included. In addition to Cr_2, we include LiH (3-21g basis set [83], bond length 1.5 Å, 11 orbitals, and 4 electrons) and H_6 chain (STO-6g basis [84], bond length 1.5 Å, 6 orbitals, and 6 electrons).

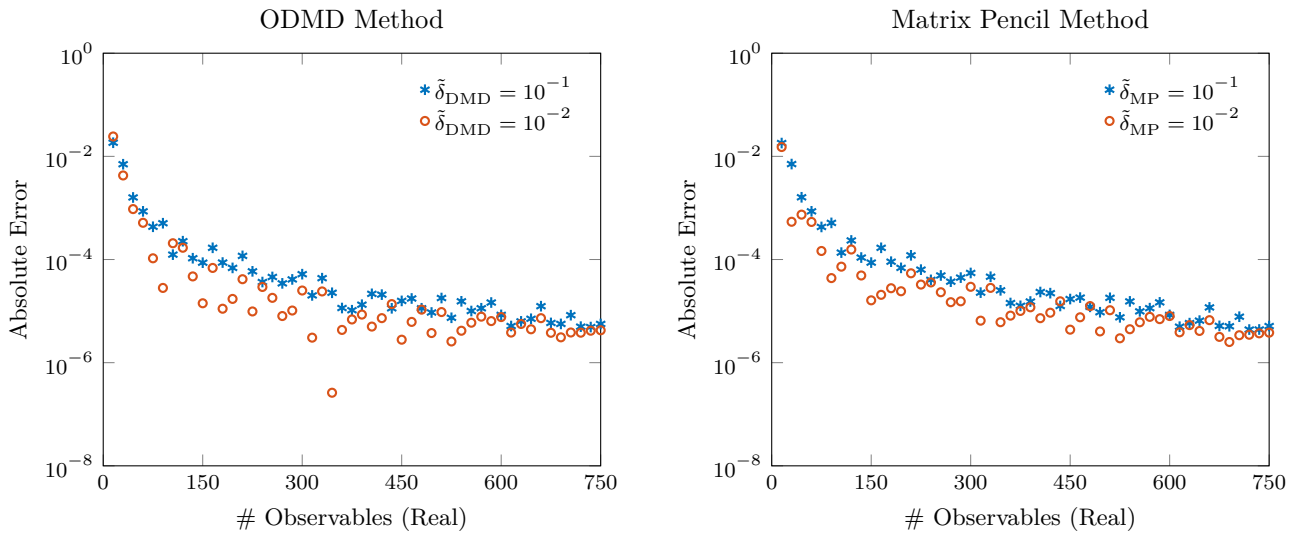

Figure 11: Data comparing ODMD and the matrix pencil method with different singular value truncations ($\tilde{\delta}_{\mathrm{DMD}}$ and $\tilde{\delta}_{\mathrm{MP}}$) for Cr_2 with noise level $\epsilon = 10^{-3}$. A timestep of $\Delta t = 3$ is used for generating the results.

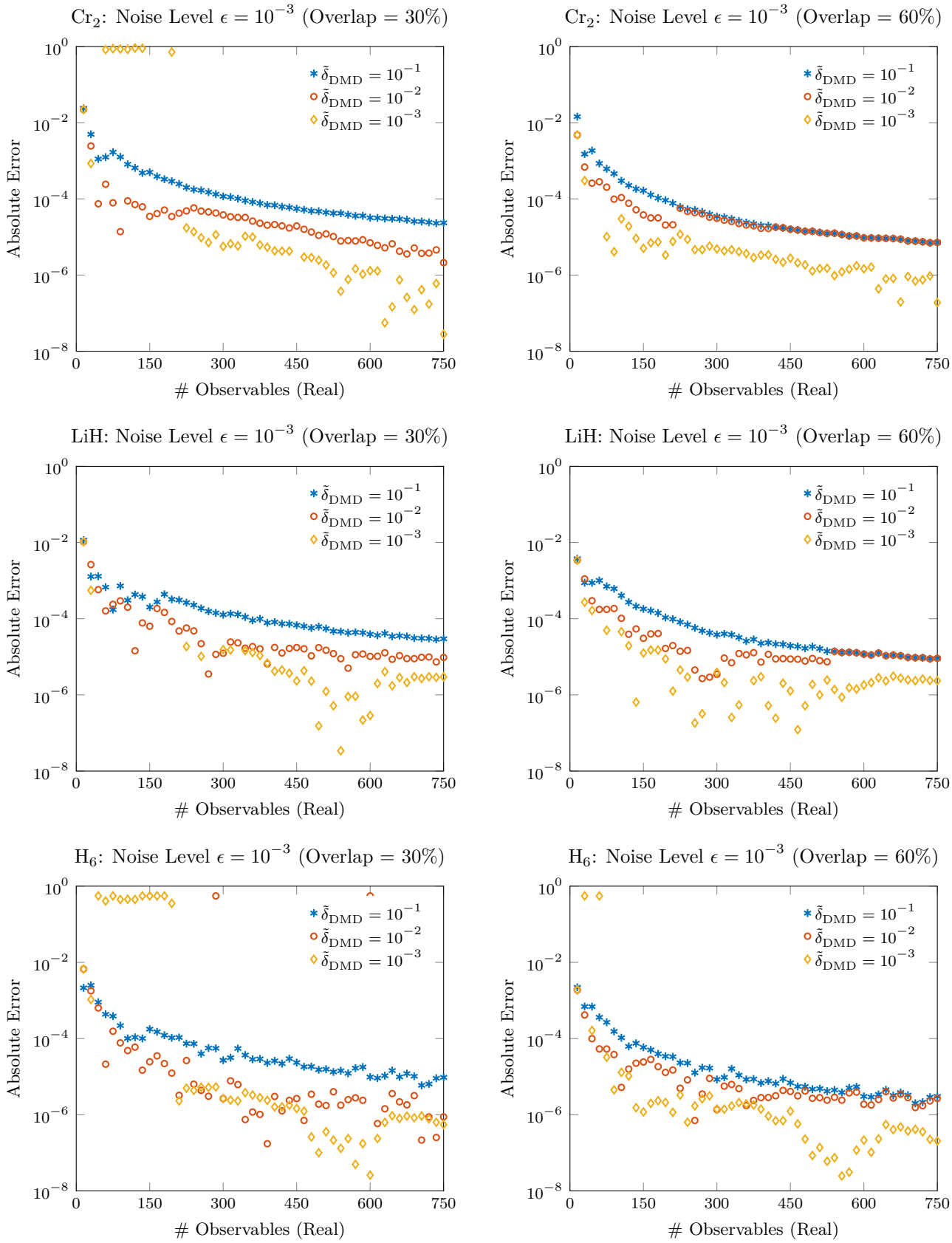

Figure 12: Comparison of ODMD ground state energy convergence for a range of molecules with varying reference-ground state overlap probabilities p_0 and singular value truncations $\tilde{\delta}_{\mathrm{DMD}}$. A timestep of $\Delta t = 3$ is used for generating the results.