

Quantum-Centric Algorithm for Sample-Based Krylov Diagonalization

Jeffery Yu,^{1, 2, 3,*} Javier Robledo Moreno,^{1, †} Joseph T. Iosue,^{1, 2, 3} Mirko Amico,¹ Luke Bertels,⁴ Daniel Claudino,⁴ Bryce Fuller,¹ Peter Groszkowski,⁵ Travis S. Humble,⁶ Petar Jurcevic,¹ William Kirby,⁷ Thomas A. Maier,⁸ Mario Motta,¹ Bibek Pokharel,¹ Alireza Seif,¹ Amir Shehata,⁹ Kevin J. Sung,¹ Minh C. Tran,¹ Vinay Tripathi,¹ Antonio Mezzacapo,^{1, ‡} and Kunal Sharma^{1, §}

¹*IBM Quantum, IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA*

²*Joint Center for Quantum Information and Computer Science,
NIST/University of Maryland, College Park, Maryland 20742, USA*

³*Joint Quantum Institute, NIST/University of Maryland, College Park, Maryland 20742, USA*

⁴*Quantum Information Science Section, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA*

⁵*National Center for Computational Sciences,
Oak Ridge National Laboratory, Oak Ridge, TN, USA*

⁶*Quantum Science, Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA*

⁷*IBM Quantum, IBM Research Cambridge, Cambridge, MA 02142, USA*

⁸*Computational Sciences and Engineering Division,*

Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA

⁹*Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA*

Approximating the ground state of many-body systems is a key computational bottleneck underlying important applications in physics and chemistry. The most widely known quantum algorithm for ground state approximation, quantum phase estimation, is out of reach of current quantum processors due to its high circuit-depths. Subspace-based quantum diagonalization methods offer a viable alternative for pre- and early-fault-tolerant quantum computers. Here, we introduce a quantum diagonalization algorithm which combines two key ideas on quantum subspaces: a classical diagonalization based on quantum samples, and subspaces constructed with quantum Krylov states. We prove that our algorithm converges in polynomial time under the working assumptions of Krylov quantum diagonalization and sparseness of the ground state. We then demonstrate the scalability of our approach by performing the largest ground-state quantum simulation of impurity models using a Heron quantum processor and the Frontier supercomputer. We consider both the single-impurity Anderson model with 41 bath sites, and a system with 4 impurities and 7 bath sites per impurity. Our results are in excellent agreement with Density Matrix Renormalization Group calculations.

I. INTRODUCTION

A significant bottleneck in physics and chemistry is the efficient estimation of the low-energy spectrum of quantum systems. Fault-tolerant quantum algorithms, including phase estimation, promise advantages over classical methods for this task, but require circuit depths beyond the reach of pre-fault-tolerant devices [1]. Shallow circuit approaches such

as the variational quantum eigensolver [2] are more hardware-efficient, but they rely on parametric optimization and stochastic estimation of complex observables, which limits their scaling because of prohibitive runtime connected to the large number of measurements [3–5]. This motivates the development of new quantum algorithms that can estimate the spectral properties of physical systems on current quantum computers.

Quantum diagonalization methods based on subspaces have emerged as promising algorithms for estimating spectral properties on pre-fault-tolerant devices [6–19]. Notably, an experimental implementation of Krylov quantum diagonalization (KQD) was shown on quantum many-body systems of up to 56 spins [17]. KQD involves constructing a subspace by time-evolving a reference state over various time intervals, followed by classical diagonalization of the Hamiltonian within that subspace. A key advantage of this approach is that convergence is guaranteed when the initial state has polynomial overlap with the ground state, and it relies on simulating quantum dynamics using circuits that can be executed at sizes beyond the reach of exact classical methods [20, 21].

* jey@umd.edu

† j.robledomoreno@ibm.com

‡ mezzacapo@ibm.com

§ kunals@ibm.com

This manuscript has been authored by UT-Battelle, LLC, under Contract No. DE-AC0500OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for the United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan.

Subspace algorithms based on individual quantum samples [15, 16, 22–25] approximate ground state energies by sampling from a quantum state and performing classical post-processing on noisy data and diagonalization in quantum-centric supercomputing environments [26]. Unlike KQD, these sample-based quantum diagonalizations (SQD) do not require time-evolution circuits, making them appealing for chemistry Hamiltonians with a large number of terms. These ideas have been experimentally demonstrated for molecular electronic structure up to sizes not amenable to exact diagonalization [16].

In this work, we introduce a new algorithm that combines key ideas from the KQD and SQD frameworks. We refer to this algorithm as *sample-based Krylov quantum diagonalization* (SKQD), which leverages the advantages of both frameworks: the convergence guarantees of Krylov methods and the noise resilience of sample-based techniques. SKQD constructs a subspace from bitstrings sampled from quantum states obtained by time-evolving a reference state over multiple intervals, and then classically diagonalizes the Hamiltonian within this subspace to approximate the ground state.

We prove that under the sparsity assumptions for the ground state and given a reference state with polynomial overlap, SKQD approximates the ground state energy in polynomial time. We experimentally demonstrate the scalability of SKQD by computing ground state properties of a single-impurity Anderson model (SIAM) [27] with 41 bath sites (42 electrons in 42 orbitals), simulated using 85 qubits and up to $6 \cdot 10^3$ two-qubit gates to prepare Krylov states; and the Anderson model with four impurities and 28 bath sites (32 electrons in 32 orbitals), using 70 qubits of a Heron processor. We show excellent agreement between SKQD and Density Matrix Renormalization Group (DMRG) [28–30] calculations, on system sizes not amenable to exact diagonalization. To the best of our knowledge, these results are the largest simulations of the ground state properties of impurity models using heterogeneous quantum and classical architectures.

II. SAMPLE-BASED KRYLOV QUANTUM DIAGONALIZATION (SKQD)

We are interested in approximating the ground-state energy of a Hamiltonian H defined for n qubits. Let $N = 2^n$. Let $|\psi_0\rangle$ denote an initial (reference) state. Similar to KQD (see Section V), we consider the following time-evolved states, also known as Krylov states:

$$|\psi_k\rangle := e^{-ikH\Delta t} |\psi_0\rangle, \quad (1)$$

where $k \in \{0, 1, \dots, d-1\}$ and Δt is a chosen time step. To implement SKQD, we proceed as follows (see Fig. 1):

1. Prepare a reference state $|\psi_0\rangle$.
2. For each $k \in \{0, 1, \dots, d-1\}$, prepare $M = \mathcal{O}(\text{poly}(n))$ copies of $|\psi_k\rangle$.
3. Measure each $|\psi_k\rangle$ in the computational basis to obtain a sequence of bitstrings $\mathcal{X}_k = \{a_{km} \mid m = 0, 1, \dots, M-1\}$.
4. Classically project H onto $\text{span}(B_{d,M})$ to obtain the matrix $\widehat{\mathbf{H}}$, where $B_{d,M} = \{|a_{km}\rangle \mid k = 0, \dots, d-1; m = 0, \dots, M-1\}$.
5. Diagonalize $\widehat{\mathbf{H}}$ classically to find the approximation to the ground state: $|\Psi\rangle = \sum_j \Psi_j |b_j\rangle$, where b_j is a symbol that summarizes the unique elements a_{km} in \mathcal{X} .

Note that the matrix $\widehat{\mathbf{H}}$ is of polynomial size and thus can be diagonalized classically, and that the Ψ_j components are the entries of the eigenvector of $\widehat{\mathbf{H}}$ with smallest eigenvalue. In the presence of noise, we also perform configuration recovery based on the $U(1)$ symmetry of the problem, as introduced in [16]. A key feature of SKQD is its robustness to noise, which comes with additional classical cost in performing configuration recovery and diagonalization, making it well-suited for noisy quantum processors. We demonstrate this capability on IBM quantum processors in Section II B. Moreover, in the Supplementary Information we show evidence of SKQD outperforming the standard KQD approach under a fixed shot budget.

Before presenting our experimental results, we argue that SKQD algorithm converges efficiently under a notion of sparsity of the ground state $|\phi_0\rangle$ of H . We define sparsity as follows:

Definition 1 ((α_L, β_L)-sparsity). For any state $|\psi\rangle$, let

$$|\psi\rangle = \sum_{j=1}^N g_j |b_j\rangle, \quad (2)$$

where (b_1, \dots, b_N) is some ordering of length- N bitstrings such that $|g_1| \geq |g_2| \geq \dots \geq |g_N|$. We say that $|\psi\rangle$ exhibits (α_L, β_L) -sparsity on $|b_1\rangle$ through $|b_L\rangle$ if

$$\sum_{j=1}^L |g_j|^2 \geq \alpha_L \quad (3)$$

and

$$|g_1|^2, \dots, |g_L|^2 \geq \beta_L. \quad (4)$$

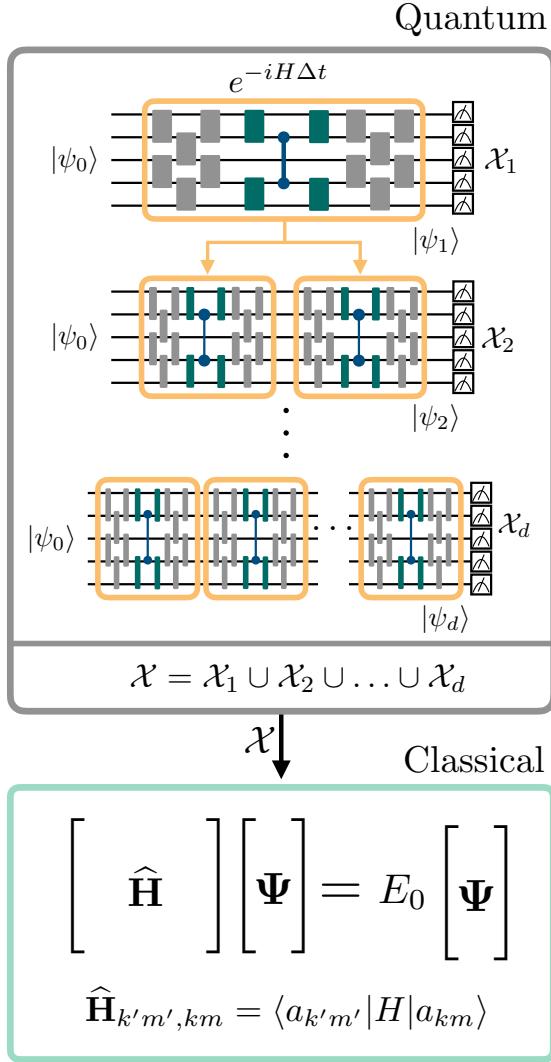


Figure 1. **SQKD algorithm.** A Krylov subspace is constructed by the time evolution of a reference state $|\psi_0\rangle$ to d different times. At the end of each circuit, the state is measured in the computational basis, yielding a set of measurement outcomes \mathcal{X} . The computational basis states sampled in \mathcal{X} are used to span an approximation to the ground state of the system $|\Psi\rangle = \sum_j \Psi_j |b_j\rangle$ for $b_j \in \mathcal{X}$. The components Ψ_j are obtained in closed form by the diagonalization of the projection of H in the subspace spanned by the sampled bitstrings.

Let L be the smallest possible integer and $\alpha_L^{(0)}, \beta_L^{(0)}$ be the largest possible parameters such that the ground state $|\phi_0\rangle$ exhibits $(\alpha_L^{(0)}, \beta_L^{(0)})$ -sparsity.

We introduce the following notation for the spectrum of H and the initial state used in SKQD. Let $E_0 \leq E_1 \leq \dots \leq E_{N-1}$ denote eigenvalues of H with corresponding orthonormal eigenstates $|\phi_0\rangle, \dots, |\phi_{N-1}\rangle$. Let $\Delta E_j = E_j - E_0$ for each

$0 < j < N$ and let $|\psi_0\rangle = \sum_{k=0}^{N-1} \gamma_k |\phi_k\rangle$ be the eigenstate decomposition of the initial state.

A. Convergence guarantees

We argue that if the ground state exhibits sparsity as defined in Definition 1, then SKQD can efficiently estimate the ground-state energy with bounded error. In particular, our main convergence result is as follows:

Theorem 1. *Let H be a Hamiltonian whose ground state $|\phi_0\rangle$ exhibits $(\alpha_L^{(0)}, \beta_L^{(0)})$ -sparsity. Let $|\Psi\rangle$ be the lowest energy state supported on the L important bitstrings in $|\phi_0\rangle$. The error in estimating the ground state energy of H using SKQD is bounded by*

$$\langle \Psi | H | \Psi \rangle - \langle \phi_0 | H | \phi_0 \rangle \leq \sqrt{8} \|H\| \left(1 - \sqrt{\alpha_L^{(0)}}\right)^{1/2},$$

provided all L important bitstrings are sampled. The success probability of sampling all L important bitstrings is at least $1 - \eta$ as long as the number of samples from each Krylov basis state exceeds $(d^2 \log(L/\eta)) / (|\gamma_0|^2 (\beta_L^{(0)} - 2\sqrt{\tilde{\varepsilon}}))$, where

$$\tilde{\varepsilon} = 2 - 2\sqrt{1 - \varepsilon/\Delta E_1}$$

with

$$\varepsilon = 8\Delta E_{N-1} \left(\frac{1 - |\gamma_0|^2}{|\gamma_0|^2} \right) \left(1 + \frac{\pi \Delta E_1}{\Delta E_{N-1}} \right)^{-(d-1)}, \quad (5)$$

where $|\gamma_0|^2 = |\langle \phi_0 | \psi_0 \rangle|^2$ denotes the overlap of the initial state $|\psi_0\rangle$ with the true ground state $|\phi_0\rangle$, and the timestep in Eq. (1) is chosen as $\Delta t = \pi/\Delta E_{N-1}$.

Thus, our analytical bound on the additive error in approximating the ground state energy of H depends on the sparsity parameter $\alpha_L^{(0)}$. Additionally, the number of samples required to find all L important bitstrings is inversely proportional to $|\gamma_0|^2$, a requirement similar to the KQD method (see Section V).

To prove Theorem 1, we develop several key results. We provide a brief proof for Theorem 1 in Section V and a detailed proof in Sections I and II; here we summarize the main ideas. First, we recall from [10] that the KQD method achieves an additive error ε in the energy as in Eq. (5). It then implies that the error in approximating $|\phi_0\rangle$ using $|\Psi\rangle$ is also small. We invoke the sparsity of the true ground state to show that $|\Psi\rangle$ is also sparse. Using this, we prove that each relevant bitstring in the ground state has

an overlap proportional to $|\gamma_0|^2$ with at least one of the Krylov basis states. Combining these results, we argue that all L relevant bitstrings can be obtained with high probability, and thus the ground state and the corresponding ground state energy can be approximated with small error.

We note that if each time-evolution unitary $e^{-ikH\Delta t}$ incurs γ Trotter error, then the number of samples needed from each Krylov basis state in Theorem 1 scales as $(\log(L/\eta)) / (|\gamma_0|^2(\beta_L^{(0)} - 2\sqrt{\tilde{\varepsilon}})/d^2 - \gamma)$. We provide a proof in Section VI.

B. Experiments on impurity models

To assess the accuracy of SKQD applied, we consider impurity models with increasing number of impurities as a testbed [27, 31–33]. The generic Hamiltonian for an impurity model is given by:

$$H = H_{\text{imp.}} + H_{\text{bath}} + H_{\text{hyb.}}, \quad (6)$$

The impurity term describes electrons that can hop between L different impurities, with a Hubbard-like onsite repulsive interaction of strength U :

$$H_{\text{imp.}} = \sum_{\substack{l,l'=1 \\ \sigma \in \{\uparrow, \downarrow\}}}^L t_{ll'} \hat{d}_{l\sigma}^\dagger \hat{d}_{l'\sigma} + U \sum_{l=1}^L \hat{d}_{l\uparrow}^\dagger \hat{d}_{l\uparrow} \hat{d}_{l\downarrow}^\dagger \hat{d}_{l\downarrow}, \quad (7)$$

where $\hat{d}_{l\sigma}^\dagger / \hat{d}_{l\sigma}$ are the creation/annihilation operators for impurity mode l with spin σ . The symmetric matrix with $t_{ll'}$ elements describes the hopping amplitudes between impurities, and its diagonal part is a chemical potential, which is set to $t_{ll} = U/2$.

H_{bath} describes a number K of non-interacting fermionic modes per impurity. Given its non-interacting nature, the bath can always be written in the single-particle basis where it is diagonal:

$$H_{\text{bath}} = \sum_{\substack{l=1 \\ \sigma \in \{\uparrow, \downarrow\}}}^L \sum_{\mathbf{k}=1}^K \varepsilon_{\mathbf{k}} \hat{c}_{\mathbf{k}l\sigma}^\dagger \hat{c}_{\mathbf{k}l\sigma}, \quad (8)$$

where $\hat{c}_{\mathbf{k}l\sigma}^\dagger / \hat{c}_{\mathbf{k}l\sigma}$ are the creation/annihilation operators for mode k and spin σ , associated to impurity l . $\varepsilon_{\mathbf{k}}$ is the energy of each bath mode.

The hybridization term describes the hopping of electrons between impurities and their corresponding bath sites:

$$H_{\text{hyb.}} = \sum_{\substack{l=1 \\ \sigma \in \{\uparrow, \downarrow\}}}^L \sum_{\mathbf{k}=1}^K V_{\mathbf{k}} (\hat{c}_{\mathbf{k}l\sigma}^\dagger \hat{d}_{l\sigma} + \hat{d}_{l\sigma}^\dagger \hat{c}_{\mathbf{k}l\sigma}) \quad (9)$$

where $V_{\mathbf{k}}$ is the so-called hybridization function. Given band-width of the bath $D = \max_{\mathbf{k}}(\varepsilon_{\mathbf{k}}) - \min_{\mathbf{k}}(\varepsilon_{\mathbf{k}})$, we consider semicircle-like hybridization functions $V_{\mathbf{k}} = V \sqrt{(D/2)^2 - \varepsilon_{\mathbf{k}}^2}$, with V a parameter that controls the hybridization amplitude. We only consider the subspace corresponding to half filling and zero total magnetization.

In this manuscript we study two families of impurity models. The first is the single-impurity Anderson model (SIAM) whose $\varepsilon_{\mathbf{k}}$ and values of $V_{\mathbf{k}}$ are derived from a 1D bath geometry with open boundary conditions (see Sec. V for further details). We choose a system size $K = 41$ bath modes, making a total of 42 spinful modes and, values of $U = 1, 3, 7, 10$.

The second is a 4-impurity model with $K = 7$ bath modes per impurity, making a total of 32 spin-full fermionic modes. The impurity modes are arranged in a square geometry, with hopping amplitudes $t_{l,l+1} = t = -1$ and $t_{l,l+2} = t' = -0.5$, and a value of $U = 10$. The values of $\varepsilon_{\mathbf{k}}$ are sampled from a uniform distribution with $\max_{\mathbf{k}}(\varepsilon_{\mathbf{k}}) = 2$ and $\min_{\mathbf{k}}(\varepsilon_{\mathbf{k}}) = -2$ (see Sec. V for the specific values). The hybridization amplitude $V = 0.16$ is chosen as the case where Heat Bath Configuration Interaction (HCl) and DMRG required the largest computational resources to converge, as detailed in Sec V. Figure 2 shows an schematic representation of the 4-impurity model.

The similarity transformation of the Hamiltonian by a fermionic Gaussian unitary (orbital rotation, or single-particle basis change) can impact the accuracy of many-body methods in the approximation of the ground-state properties of the problem [31, 34–43]. In the limit of vanishing V , the ground state of the impurity models is sparse in the basis where the bath is diagonal, since the state of each bath can be described by a single Slater determinant, and the ground state of the impurities can be described by a small number of basis states. As the value of V increases, it becomes energetically favorable to allow the hopping of electrons between the impurity and corresponding bath, resulting in an increased number of Slater determinants required to obtain an accurate description of the ground state. Motivated by the observation that the basis of so called natural orbitals (NOs) that diagonalizes the one-body reduced density matrix Γ yields the set of orbitals in which the wave function is closest to a single Slater determinant, we perform a two-step SKQD experiment. The first SKQD calculation is performed in the basis where the bath is diagonal, obtaining an approximation to Γ . The reduced density matrix is diagonalized in blocks in order to find approximate NOs that only mix impurity modes with two or three bath modes for the 4-impurity and single-impurity models respectively. The bath modes that are al-

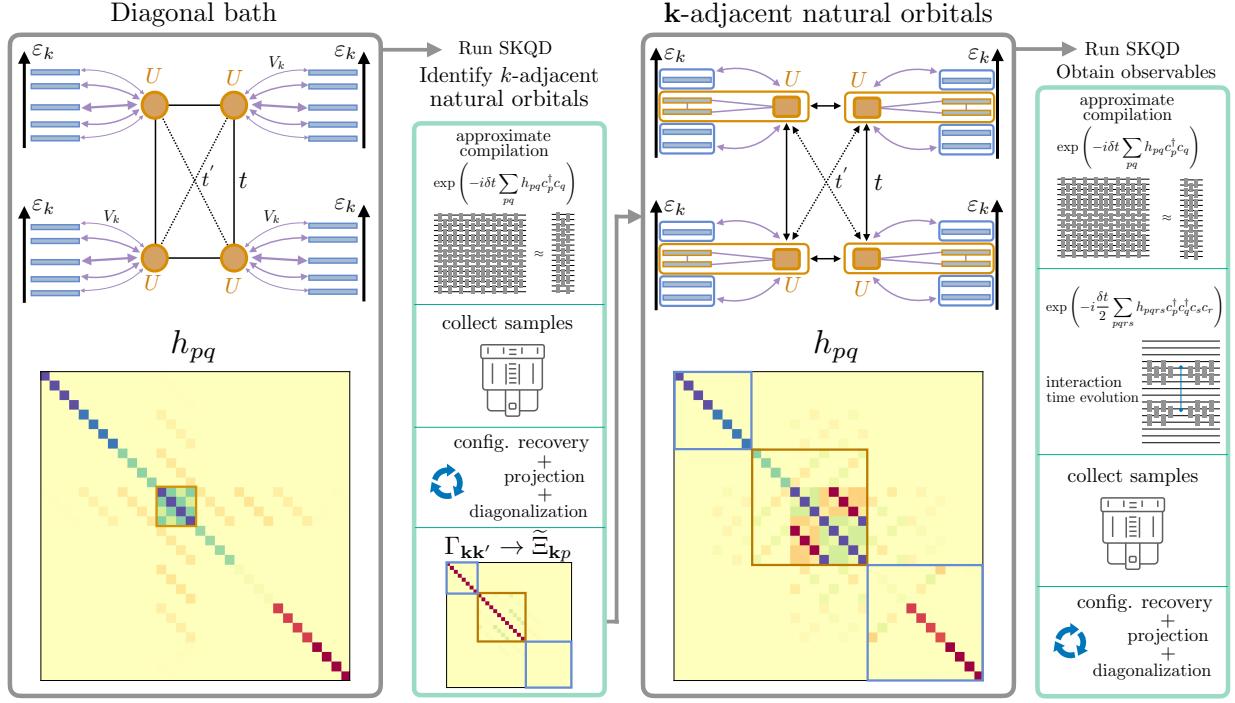


Figure 2. SQKD experimental workflow for the ground state of four-impurity model. From left to right: the 4-impurity model in the basis where the baths are diagonal, with the corresponding one-body matrix elements of the Hamiltonian h_{pq} . The brown box shows the impurity modes. SKQD is first run in this basis. The first step is the compilation of the free-fermion time evolution into a shallow circuit of Givens rotations. Then, measurement realizations are collected from the quantum device at each Trotter step, followed by an SQD ground-state estimation that uses the configuration recovery procedure, as introduced in Ref. [16]. The one-body reduced density matrix $\Gamma_{kk'}$ is used to identify k -adjacent natural orbitals, where the impurity mode is only allowed to be mixed with the bath modes corresponding to k_f and $k_f - 1$. The resulting Hamiltonian is one where the one-body matrix elements h_{pq} are close to diagonal deep in the Fermi sea and for large values of k , and with off-diagonal two-body matrix elements. SKQD is run in this new basis, requiring the approximate compilation of the free-fermion evolution, and the efficient compilation into a constant-depth circuit of the off-diagonal two-body terms.

lowed to be mixed with the impurity mode are those closest to the Fermi level ($\varepsilon_{\mathbf{k}_f} = 0$). This procedure is summarized in Fig. 2.

We use the Jordan-Wigner [44] encoding to map the fermionic degrees of freedom into the quantum processor. The second-order Trotter-Suzuki decomposition is used to realize each $|\psi_k\rangle$, as described in Sec. V. The mappings into the quantum processors for the SIAM and 4-impurity models are shown in panels (a) and (d) of Fig. 3.

To benchmark the accuracy of the quantum experiments, we choose DMRG as a reference classical method, since it is one of the *state-of-the-art* approximate methods for single-band, single-impurity models [31, 33, 45]. Several physical properties are compared between the DMRG and SKQD estimations. The first is the relative error in the SKQD ground-state energy estimation, defined as $|(\langle E_{\text{SKQD}} \rangle - \langle E_{\text{DMRG}} \rangle)/\langle E_{\text{DMRG}} \rangle|$. Additionally, we compare the estimation of other relevant physical prop-

erties, such as two-point correlation functions.

Figure 3 (b) shows the relative error in the SKQD ground-state energy estimation as a function of the subspace dimension on the SKQD eigenstate solver D . The SKQD relative error decreases from values $\sim 10^{-4}$ to $\sim 10^{-5}$ as U increases from $U = 1$ to $U = 10$, which is the opposite behavior as compared to RHF and CCSD solutions. The SKQD estimations of the ground state energy become more accurate with increasing correlations in the system. This is a consequence of the increased ground-state sparsity for larger values of U . Panel (c) of Fig. 3 compares the values of the two-point spin $\bar{C}_S(\mathbf{j})$ (Eq. 22) and density $\bar{C}_n(\mathbf{j})$ (Eq. 23) correlation functions obtained from SKQD to those obtained with DMRG. The SKQD estimations are in excellent agreement with the DMRG values for all values of \mathbf{j} , the distance between the impurity spin and the bath spin. Additionally, in the Supplementary Materials we show that the SKQD error in the SIAM does not in-

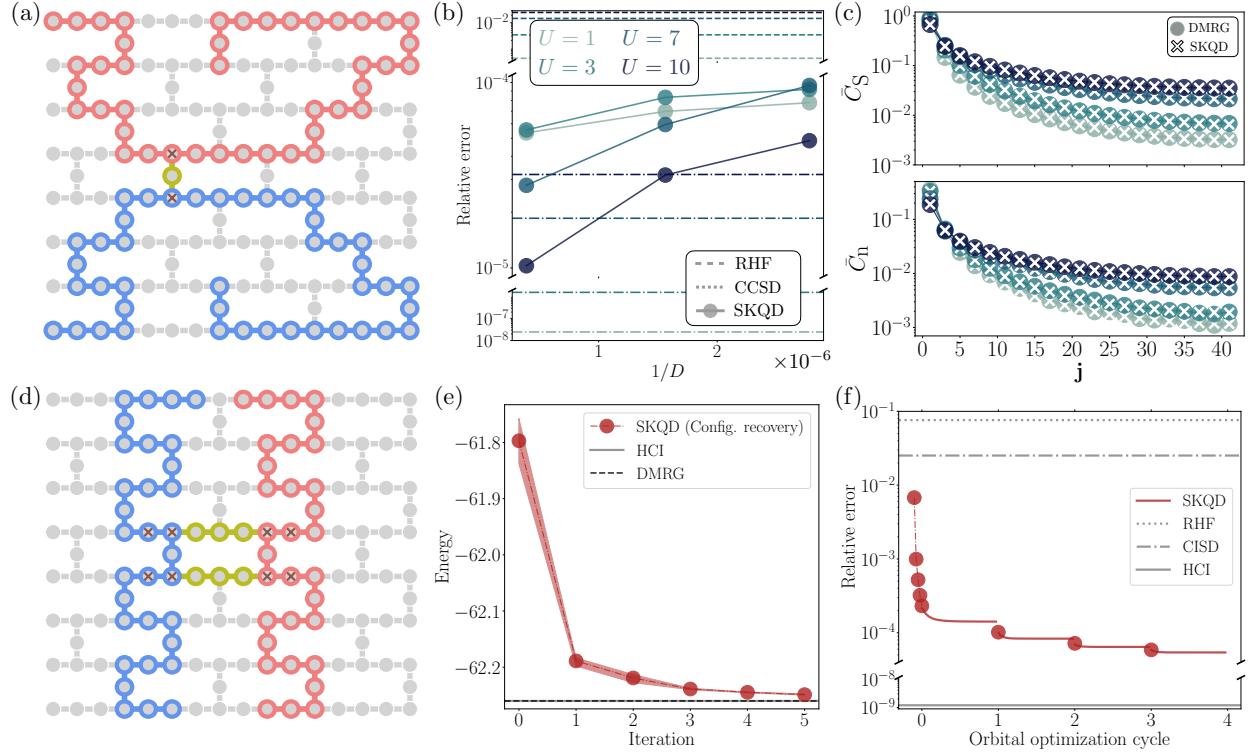


Figure 3. Experiments on quantum processors, and comparison against DMRG. (a)-(c) SIAM with 41 bath sites (85-qubit experiment). (a) Shows the qubit layout, where red and blue qubits correspond to spin-up and spin-down degrees of freedom respectively. The qubit marked with a cross is the qubit representing the impurity. The green qubit is an auxiliary qubit used to implement the time evolution of the Hubbard interaction. (b) Ground state energy error relative to the DMRG estimation as a function of the inverse of the SKQD subspace dimension D . Different colors correspond to different values of the onsite repulsion U as indicated in the legend. The Hartree-Fock energies (RHF) and Coupled Cluster with Singles and Doubles (CCSD) errors are shown for reference. (c) Comparison of the two-point spin \bar{C}_S and density \bar{C}_n correlation functions (see Eqs. 22 and 23) obtained with DMRG and SKQD. (d)-(f) 4-impurity model ($L = 4$) with $K = 7$ bath models per impurity (70-qubit experiment). (d) Shows the qubit layout, where red and blue qubits correspond to spin-up and spin-down degrees of freedom respectively. The qubits marked with a cross correspond to the qubits representing the impurity. The green qubits are auxiliary qubits used to implement the time evolution of the Hubbard interaction. (e) SKQD ground state energy estimation in the \mathbf{k} -adjacent NOs basis as a function of the configuration recovery step. The HCI and DMRG energy estimations are shown for reference. (f) Effect of orbital optimizations applied to the converged SKQD ground state estimation. The error in the ground-state energy relative to the converged DMRG energy is shown as a function of the self-consistent orbital optimization cycle. The dots connected by the dashed lines correspond to the configuration recovery trajectory shown in panel (e). The solid lines show the improvement of the energy error when the orbitals are optimized to minimize the energy for fixed wave function coefficients. The dots show the error after re-diagonalizing the Hamiltonian in the new basis found by the orbital optimization procedure. The RHF, CISD and HCI errors are shown for reference.

crease as the number of bath sites is increased from $K = 29$ to $K = 41$.

Figure 3 (e) shows the SKQD energy convergence as a function of the self-consistent configuration recovery iteration [16], in the basis of \mathbf{k} -adjacent NOs. An additional retention and carry-over of the most relevant configurations is considered as compared to the procedure introduced in Ref. [16] (see Sec. V for additional details). The final energy estimation is on good agreement with the DMRG and HCI estimates. Once the recovery procedure has identified the rel-

evant electronic configurations and corresponding wave-function components, we perform a final set of 5 self-consistent orbital optimizations cycles, where we alternatively optimize over single-particle basis transformations and wave function coefficients. The improvement on the energy error as a function of the orbital optimization cycle is shown in Fig. 3 (f). The SKQD energy estimation has a relative error in the ground state energy of $\sim 10^{-4}$, in the basis of \mathbf{k} -adjacent NOs, while the relative error in the basis of optimized orbitals decreases to $\sim 6 \cdot 10^{-5}$.

III. DISCUSSION

We obtained our ground-state energy estimations via sampling from Krylov states. Consequently and in order to keep low circuit depths, the class of problems which are amenable to SKQD are those that can be mapped easily using the layout of the quantum processor. Thus, *ab-initio* use cases are out of reach for the SKQD implementation as presented in this manuscript in near term processors. We leave it to future investigations the adaptation of SQKD to *ab-initio* problems with reduced-depth circuits. On lattice problems, SKQD resolves practically every issue of algorithms for ground states: it does not require optimization of ansatzaes, it does not incur in the quantum measurement problem, and it is robust to noisy samples since one can use configuration recovery and a classical diagonalization overhead to effectively remove the effect of noise.

SKQD shares many convergence properties with the standard Krylov quantum diagonalization [17]. However, it requires reduced circuit depths since it doesn't need to execute Hadamard tests as subroutines, and has improved noise resistance properties, as mentioned before. While our convergence proofs requires sparsity of the ground state, in practice SKQD could be used to study non-sparse ground states, in scenarios where basis states are captured by the Krylov circuits.

We performed experiments with circuits up to 85 qubits and $\sim 6 \cdot 10^3$ two-qubit gates simulating the ground-state properties of the single-impurity Anderson model for different values of the onsite repulsion strength U , obtaining excellent agreement with DMRG and HCI calculations of the same system. Experiments were also performed in circuits with 70 qubits and $\sim 6 \cdot 10^3$ two-qubit gates studying the ground-state of generic 4-impurity models, again achieving excellent agreement with DMRG and HCI calculations. This confirms that SKQD can be used as to probe ground state physics on pre-fault-tolerant quantum computers, exceeding the reach of existing quantum methods for lattice problems, for system sizes well beyond the reach of exact diagonalization methods.

Note added. While finalizing our paper, we noticed two independent papers on arXiv that share some of the ideas presented here [46, 47].

IV. CODE AND DATA AVAILABILITY

The simulation of the time evolution of fermionic Hamiltonians is carried out with the library `ffsim` [48], while configuration recovery, projection

and diagonalization are carried out with the python package `qiskit-addon-sqd` [49]. Quantum circuits are generated and transpiled using `qiskit` [50]. DMRG calculations are performed using the `block2` package [51]. HF and CCSD calculations are performed using the PySCF library [52, 53].

ACKNOWLEDGMENTS

We acknowledge helpful discussions with Jay Gambetta, Toshinari Itoko, and Caleb Johnson. We acknowledge the use of `tacokit` and are grateful to Zlatko Minev for development of the package and troubleshooting support. This research was supported by the Quantum Science Center, a National Quantum Science Initiative of the Department of Energy (DOE), managed by the Oak Ridge National Laboratory (ORNL). The work by T.A.M. (model selection and analysis) was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division. D.C. and L.B. acknowledge support from the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the US Department of Energy. This material is based upon work supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Quantum Science Center (QSC). This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.

AUTHOR CONTRIBUTIONS

Design of the workflow and experiments: J.Y., J.R.-M., P.G., W.K., T.A.-M., M.M., K.J.-S., M.T., A.M., and K.S. Implementation of the workflow: J.R.-M., M.A., P.G., P.J., M.M., B.P., and V.T. Execution of the quantum part of the workflow on Heron: B.P., P.J., and V.T. Execution of the classical part of the workflow on Frontier: P.G. Numerical benchmarks: J.Y., J.R.-M., J.T.-I., M.M., and K.J.S. Analytical derivations: J.Y., J.R.-M., J.T.-I., W.K., M.T., and K.S. All authors contributed to the manuscript writing and data analysis.

V. METHODS

Before providing proof details of Theorem 1, we recall prior results on KQD and SQD. As discussed in Section II, let $|\psi_0\rangle$ denote an initial (reference)

state. In KQD, the quantum Krylov subspace \mathcal{S} is generated by the time-evolved states $|\psi_k\rangle$ as in Eq. (1), where $k \in \{0, 1, \dots, d-1\}$. This construction reduces the exponentially large N -dimensional Hilbert space to a subspace of dimension d . By projecting the Hamiltonian onto \mathcal{S} , we get the following generalized eigenvalue problem:

$$\tilde{\mathbf{H}}v = \tilde{E}\tilde{\mathbf{S}}v, \quad (10)$$

where v is a coordinate vector corresponding to the state $|\tilde{\phi}\rangle = \sum_{k=0}^{d-1} v_k |\psi_k\rangle \in \mathcal{S}$, and

$$\tilde{\mathbf{H}}_{jk} := \langle \psi_j | H | \psi_k \rangle, \quad \tilde{\mathbf{S}}_{jk} := \langle \psi_j | \psi_k \rangle. \quad (11)$$

Let \tilde{E} denote the ground-state energy approximation obtained from the Krylov quantum subspace approach. From the results of Epperly *et al.* in [10], it follows that in the ideal, noise-free case, $0 \leq \tilde{E} - E_0 \leq \varepsilon$, as in Eq. (5), where E_0 is the true ground-state energy.

From Eq. (5), it follows that if the initial state has a nontrivial ground-state overlap $|\gamma_0|^2 = \Theta(1)$, and if H has a well-behaved spectrum (i.e., ΔE_{N-1} not growing too quickly and ΔE_1 not too small), then a constant error ε in approximating the ground-state energy can be achieved by setting $d = O(\log(1/\varepsilon))$ (see Section I for a review of this analysis). Moreover, for an analysis of noisy KQD, we refer readers to [10, 54].

We now summarize the SQD framework for approximating ground-state energies, analyzing its behavior against a notion of sparsity of the ground state $|\phi_0\rangle$ of H . Similar to Section II, we consider L to be the smallest possible integer and $\alpha_L^{(0)}, \beta_L^{(0)}$ be the largest possible parameters such that the ground state $|\phi_0\rangle$ exhibits $(\alpha_L^{(0)}, \beta_L^{(0)})$ -sparsity. If L bitstrings $|b_1\rangle, \dots, |b_L\rangle$ can be sampled with high probability from a quantum circuit [16], then H can be represented in the subspace spanned by $\{|b_j\rangle\}_{j=1}^L$, yielding a projected matrix $\hat{\mathbf{H}}$. By classically diagonalizing $\hat{\mathbf{H}}$, we obtain an approximation of E_0 . In [16], it was shown that using this approach, one can approximate E_0 with an additive error of up to $\varepsilon' = 2\sqrt{2}\|H\| \left(1 - \sqrt{\alpha_L^{(0)}}\right)^{1/2}$ and a success probability of at least $1-\eta$, provided that the number of samples obtained from the state exceeds $2/\beta_L^{(0)} \log(1/\eta)$.

The KQD approach can achieve good accuracy in estimating the ground-state energy, provided the initial state has a nontrivial overlap with $|\phi_0\rangle$ and the spectrum of H is well-behaved. However, this approach requires estimating the matrix elements of $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{S}}$, as defined in Eq. (11), necessitating $O(1/\varepsilon^2)$ samples to achieve error ε , and additional sample overhead for noise mitigation. While these challenges

do not rule out near-term implementation, as shown by [17], finding ways to circumvent them is clearly desirable. On the other hand, the SQD approach is more natural for near-term devices and well-suited for problems where the target ground state energy can be captured by a sparse wavefunction.

A. Convergence proof

We prove the convergence as follows. We begin with the sparsity assumption for ground states. Let the ground state $|\phi_0\rangle$ of H exhibits $(\alpha_L^{(0)}, \beta_L^{(0)})$ -sparsity, in the sense of Definition 1. For our convergence proof, we assume the ground state is sparse in the sense that $L = \mathcal{O}(\text{poly}(n))$. We provide a comparison of our sparsity assumption with those proposed in [55, 56] in Section IV.

We first prove that if the KQD method achieves an additive error ε in the energy as in Eq. (5), then the error in the corresponding approximate ground state $|\tilde{\phi}\rangle$ is

$$\left\| |\tilde{\phi}\rangle - |\phi_0\rangle \right\|^2 \leq \tilde{\varepsilon} = O\left(\frac{\varepsilon}{\Delta E_1}\right). \quad (12)$$

Next, we show that if the true ground state $|\phi_0\rangle$ exhibits $(\alpha_L^{(0)}, \beta_L^{(0)})$ -sparsity, then $|\tilde{\phi}\rangle$ also exhibits (α_L, β_L) -sparsity, with

$$\alpha_L = \alpha_L^{(0)} - 2\sqrt{\tilde{\varepsilon}}, \quad \beta_L = \beta_L^{(0)} - 2\sqrt{\tilde{\varepsilon}}. \quad (13)$$

We denote the set of important bitstrings defining the ground state as $B = \{|b_i\rangle\}_{i=1}^L$, where a bitstring's importance is determined by $|g_i|$, as defined in Definition 1. We prove that each bitstring in B has overlap proportional to $|\gamma_0|^2$ with at least one of the Krylov basis states. We express each $|\psi_k\rangle$ in the computational basis as $|\psi_k\rangle = \sum_{j=1}^N c_j^{(k)} |b_j\rangle$. Assume that $|\tilde{\phi}\rangle = \sum_{k=0}^{d-1} d_k |\psi_k\rangle$, where $|\psi_k\rangle$ is defined in Eq. (1) and $|d_k| \leq \frac{1}{|\gamma_0|}$ for each $k = 0, 1, \dots, d-1$ (note that the latter condition is nontrivial because the $|\psi_k\rangle$ are nonorthogonal, so the squared norms of the d_k need not sum to unity). We then show that for each $1 \leq j \leq L$, there exists some k such that

$$|c_j^{(k)}|^2 \geq \frac{|\gamma_0|^2 \beta_L}{d^2}, \quad (14)$$

where one should note that β_L also depends on d via Eqs. (12) and (13) and the fact that the error ε from KQD converges with increasing d . Thus, one can efficiently obtain the L important bitstrings by sampling from Krylov basis states, provided the initial state $|\psi_0\rangle$ has overlap $|\gamma_0|^2 \in \mathcal{O}(1/\text{poly}(n))$ with the exact ground state.

B. Impurity model parameters

a. Single-impurity Anderson model.— In this case the impurity Hamiltonian is simplified to:

$$H_{\text{imp}} = \frac{U}{2} (\hat{n}_{d\uparrow} + \hat{n}_{d\downarrow}) + U \hat{n}_{d\uparrow} \hat{n}_{d\downarrow}. \quad (15)$$

We consider a bath corresponding to a 1D chain of length K with open boundary conditions, sharing the topology of the impurity model considered in Ref. [31]:

$$H_{\text{bath}} = -t \sum_{\substack{\mathbf{j}=0 \\ \sigma \in \{\uparrow, \downarrow\}}}^{K-1} \left(\hat{c}_{\mathbf{j},\sigma}^\dagger \hat{c}_{\mathbf{j}+1,\sigma} + \hat{c}_{\mathbf{j}+1,\sigma}^\dagger \hat{c}_{\mathbf{j},\sigma} \right), \quad (16)$$

Where the symbol \mathbf{j} labels the bath sites in the position basis. The hybridization term describes the hopping between the first bath site and the impurity:

$$H_{\text{hyb}} = V \sum_{\sigma \in \{\uparrow, \downarrow\}} \left(\hat{d}_\sigma^\dagger \hat{c}_{0,\sigma} + \hat{c}_{0,\sigma}^\dagger \hat{d}_\sigma \right). \quad (17)$$

Due to the approximate translational symmetry present in the bath, we do not expect the ground state to be sparse in the position basis (Eq. 16). Since the bath is non-interacting, there exists a single-particle basis transformation that diagonalizes H_{bath} . We refer to the transformed basis as the basis where the bath is diagonal, whose fermionic modes are labeled by the sub-index \mathbf{k} :

$$\hat{c}_{\mathbf{k},\sigma}^\dagger = \sum_{\mathbf{j}=0}^{K-1} \Xi_{\mathbf{j}\mathbf{k}} \hat{c}_{\mathbf{j},\sigma}^\dagger, \quad (18)$$

where $\mathbf{k} = 0, \dots, K-1$ and $\Xi \in \mathbb{R}^{L \times L}$ the orthonormal matrix that diagonalizes the $L \times L$ hopping matrix T :

$$T = \begin{bmatrix} 0 & -t & 0 & 0 & \dots & 0 \\ -t & 0 & -t & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & & 0 \\ 0 & \dots & -t & 0 & & -t \\ 0 & \dots & 0 & -t & 0 \end{bmatrix}. \quad (19)$$

While this is close to a momentum basis, each column in Ξ does not exactly correspond to a basis vector of the discrete Fourier transform in a one-dimensional domain with L points, due to the choice of open boundary conditions. In the basis that diagonalizes the bath, the bath and hybridization Hamiltonians take the form:

$$H_{\text{bath}} = \sum_{\substack{\mathbf{k}=0 \\ \sigma \in \{\uparrow, \downarrow\}}}^{L-1} \varepsilon_{\mathbf{k}} \hat{n}_{\mathbf{k},\sigma}, \quad (20)$$

with $\varepsilon_{\mathbf{k}}$ the eigenvalues of T , and:

$$H_{\text{hyb}} = \sum_{\substack{\mathbf{k}=0 \\ \sigma \in \{\uparrow, \downarrow\}}}^{L-1} V_{\mathbf{k}} \left(\hat{d}_\sigma^\dagger \hat{c}_{\mathbf{k},\sigma} + \hat{c}_{\mathbf{k},\sigma}^\dagger \hat{d}_\sigma \right). \quad (21)$$

With $V_{\mathbf{k}} = V \cdot \Xi_{0\mathbf{k}}$. Note that the locality of the hybridization term is lost in favor of a sparse representation of the ground state.

In this model the two-point spin and density correlation functions are relevant to study phenomena like the Kondo screening length [32, 33, 57], or their universal collapse for different values of U [31, 33]. The staggered spin-spin correlation function is defined as:

$$C_S(\mathbf{j}) = (-1)^j \left[\langle \hat{S}_d \cdot \hat{S}_j \rangle - \langle \hat{S}_d \rangle \cdot \langle \hat{S}_j \rangle \right], \quad (22)$$

where the spin operators are defined as: $\hat{S}_j^\mu = \sum_{\alpha\beta} \sigma_{\alpha\beta}^\mu \hat{c}_{j,\alpha}^\dagger \hat{c}_{j,\beta}$, with $\mu = x, y, z$, and $\hat{S}_d^\mu = \sum_{\alpha\beta} \sigma_{\alpha\beta}^\mu \hat{d}_\alpha^\dagger \hat{d}_\beta$. The second one is the staggered density-density correlation function:

$$\bar{C}_n(\mathbf{j}) = (-1)^j \sum_{\sigma \in \{\uparrow, \downarrow\}} \left[\langle \hat{n}_{d\sigma} \hat{n}_{j\sigma} \rangle - \langle \hat{n}_{d\sigma} \rangle \langle \hat{n}_{j\sigma} \rangle \right]. \quad (23)$$

b. 4-impurity model.— The values of $\varepsilon_{\mathbf{k}}$ are shown in Fig. 4, corresponding to the values obtained from drawing 7 random real numbers with uniform probability in the range $[-1, 1]$, and translated and scaled such that $\min_{\mathbf{k}} \varepsilon_{\mathbf{k}} = -2$ and $\max_{\mathbf{k}} \varepsilon_{\mathbf{k}} = 2$. The choice of a random dispersion relation for the bath modes is motivated by the desire to eliminate as much as possible the effect of an specific underlying lattice geometry for the bath.

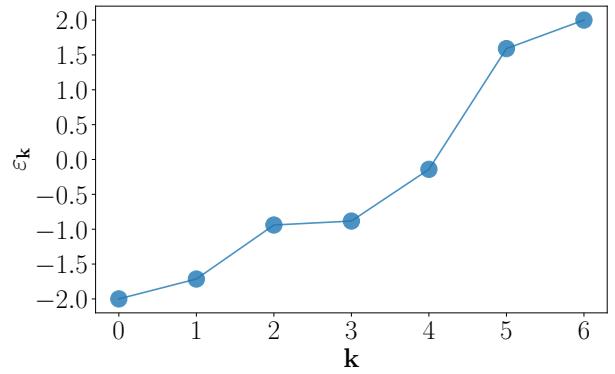


Figure 4. Dispersion relation of the bath for the 4-impurity model. See Equation 8.

C. Computational complexity of the approximation of the ground state of the 4-impurity model as a function of V

One of the goals of this study was to understand the performance of SKQD in relation to typical electronic structure methods for classical computers, as well as to identify the more challenging regimens for the classical methods, and focus on those regimes in the use of SQKD. To this end, we focus on the 4-impurity model and the following values of the hybridization parameter, $V = 0.16, 0.32, 0.40, 0.60, 0.80, 1.60, 2.00$. For these Hamiltonians, we ran standard electronic methods: restricted closed-shell Hartree-Fock (RHF), Moller-Plesset second-order perturbation theory (MP2), configuration interaction singles and doubles (CISD), coupled-cluster singles and doubles (CCSD), and CCSD with perturbative triples (CCSD(T)) as implemented in PySCF [52, 53] (we remark that CCSD does not converge for $V = 0.16$, and here we report the solution leading to the smallest (T) correction). In addition, we ran the density-matrix renormalization group (DMRG) as implemented in Block2 [58], and the heat-bath configuration interaction (HCI) as implemented in DICE [59]. We ran HCI using the basis of CCSD natural orbitals, using truncation thresholds of $\varepsilon = x \cdot 10^{-y}$ with $x = 5, 1$ and $y = 3, 4, 5, 6, 7$. The choice of natural orbitals is compelling for HCI, since natural orbitals may lead to a more compact representation of the ground-state wavefunction as a linear combination of Slater determinants than, e.g., molecular orbitals. We ran DMRG using the basis of CCSD natural orbitals (ordered with a genetic algorithm), enforcing $SU(2)$ symmetry, and using a schedule with bond dimensions $D = 100, 200, 400, 600, 800, 1000, 2000, 3000, 4000$ and 8 sweeps per bond dimension, which allows us to converge the DMRG energy well within 10^{-10} in units of energy.

The total energy from various electronic structure methods is shown in Fig. 5. RHF and CISD overestimate the ground-state energy, and MP2 underestimates it, to an extent that increases with decreasing V . On the other hand, where CCSD and CCSD(T) converge ($V > 0.16$), these methods agree with each other, with HCI, and with DMRG well within 0.001 in energy units. In fact, the largest deviation between the HCI and DMRG energies reported in the figure is $\sim 10^{-7}$ energy units.

In Fig. 6 we explore how the energy error of HCI and DMRG depends on the computational cost, defined as

$$\varepsilon_{\text{method}} = E_{\text{method}}(P) - E_{\text{ref}} \quad (24)$$

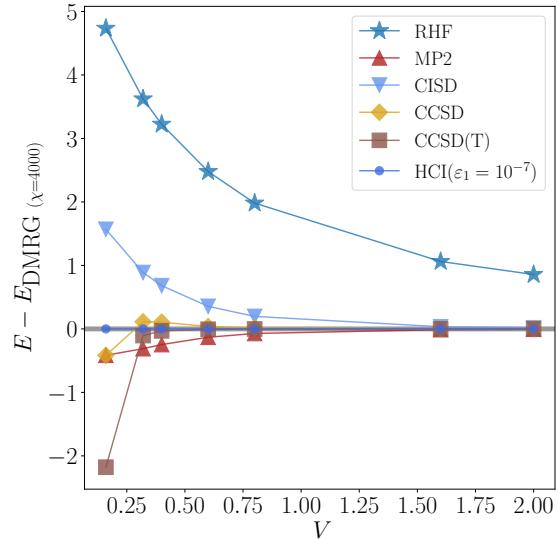


Figure 5. Ground-state energy of the four-impurity model using the classical methods RHF, MP2, CISD, CCSD, and CCSD(T), HCI with truncation threshold 10^{-7} , and DMRG with bond dimension $D = 4000$.

where E_{ref} is a reference value for the ground-state energy (here, the DMRG energy with bond dimension 4000) and P is a parameter that controls the computational cost of the method, i.e. the number d_{HCI} of configurations in HCI and the bond dimension χ of the MPS state in DMRG.

For both methods, the energy error follows a power law dependence on the inverse of the cost control parameter $1/P$ (either the subspace dimension for HCI or the bond dimension for DMRG), with exponents and prefactors that depend on both V and the method. In other words,

$$\begin{aligned} \varepsilon_{\text{DMRG}}(V) &= C_{\text{DMRG}}(V) \left(\frac{1}{D}\right)^{m_{\text{DMRG}}(V)} \\ \varepsilon_{\text{HCI}}(V) &= C_{\text{HCI}}(V) \left(\frac{1}{d_{\text{HCI}}}\right)^{m_{\text{HCI}}(V)}. \end{aligned} \quad (25)$$

Such a power-law scaling, combined with the empirical observation that HCI and DMRG do not appear to be converging on local minima of the energy, means that decreasing the error of HCI and DMRG calculations below a desired threshold δ using a classical computer does not require an exponential cost in $1/\delta$, either in memory or in runtime. Therefore, the only way to achieve competitive results for SKQD or any other method with systematically improvable accuracy is to achieve a power-law behavior with a lower exponent and/or prefactor.

The results shown in Figs. 5 and 6 show that the problem becomes harder for DMRG and HCI

as the value of V is decreased, since the system becomes less dominated by the kinetic energy terms. We choose the lowest value of $V = 0.16$ to test the accuracy of SKQD on hardware experiments.

D. Compilation of time-evolution circuits

We use the Jordan-Wigner [44] encoding to map the fermionic degrees of freedom into the quantum processor. The second-order Trotter-Suzuki decomposition is used to realize each $|\psi_k\rangle$:

$$|\psi_k\rangle \approx \left[e^{-i\frac{\Delta t}{2}H_2} e^{-i\Delta t H_1} e^{-i\frac{\Delta t}{2}H_2} \right]^k |\psi_0\rangle, \quad (26)$$

where H_1 and H_2 correspond to the one- and two-body terms of the generic interacting-electron Hamiltonian:

$$H = \sum_{p,q} h_{pq} \hat{a}_{p\sigma}^\dagger \hat{a}_{q\sigma} + \sum_{\substack{p,q,r,s \\ \sigma\tau}} \frac{h_{pqrs}}{2} \hat{a}_{p\sigma}^\dagger \hat{a}_{q\tau}^\dagger \hat{a}_{s\tau} \hat{a}_{r\sigma}. \quad (27)$$

Note that any of the impurity-model Hamiltonians considered in this manuscript can be written in the above form for specific values of h_{pq} and h_{pqrs} , where the operators \hat{a}_p^\dagger label both impurity and bath modes.

Since $e^{-i\Delta t H_1} = \exp\left(-i\Delta t \sum_{pq,\sigma} h_{pq} \hat{a}_{p\sigma}^\dagger \hat{a}_{q\sigma}\right)$ is a fermionic Gaussian unitary, it can be realized exactly by a brickwork circuit of Givens rotations, applied to adjacent pairs of qubits, whose depth is equal to the number of fermionic modes in each spin-species, where the qubits are arranged in a one-dimensional chain [60–62] (see Fig. 2). As shown in Fig. 2, h_{pq} is close to diagonal both in the basis where the bath is diagonal, and in the basis of \mathbf{k} -adjacent natural orbitals. Additionally, we require $\Delta t \ll 1$, yielding a $e^{-i\Delta t H_1}$ unitary that is close to the identity. This observation is used to approximately compile $e^{-i\Delta t H_1}$ into a depth-3 for the SIAM, and depth-9 for the 4-impurity model brickwork circuit of Givens rotations with adjoint representation G , as shown in Fig. 2. The angles of the Givens rotations are obtained by maximizing the Hilbert-Schmidt norm of the $((K+1) \cdot L) \times ((K+1) \cdot L)$ matrix $G^\dagger \cdot \Lambda$, where Λ is the matrix exponential of $[-i(\Delta t)h]$, with h being the $((K+1) \cdot L) \times ((K+1) \cdot L)$ matrix with components h_{pq} . The optimization in the space of Gaussian unitaries is performed using gradient descent [34] with the *ADAM* [63] update rule. Gradients are computed using the automatic differentiation functionalities of the *Jax* package [64].

The compilation of $e^{-i\frac{\Delta t}{2}H_2}$ into a quantum circuit is simple when H_2 is diagonal, which is the case in the position basis, and the basis with the diagonal bath. In the Jordan-Wigner encoding, this is a

controlled-phase gate between the two qubits representing the spin-up and spin-down impurity degrees of freedom. This gate is applied via an auxiliary qubits (green in Fig. 3 (a) and (d)) that connect the spin-up and spin-down chains of qubits.

In the basis of \mathbf{k} -adjacent natural orbitals, H_2 is no longer diagonal. Consequently, its time evolution would naively require circuit depths growing as $\mathcal{O}((K+1) \cdot L)^4$. By construction, H_2 can be diagonalized by a single-particle basis transformation that mixes only 4 fermionic modes for the SIAM and 12 for the 4-impurity model. This basis transformation can be achieved by the application of Givens rotation gates acting on 4 (or 12) adjacent qubits that implement the Gaussian unitary of the change of basis (see Fig. 2). After this transformation has been applied, H_2 is diagonal and its time evolution is realized as described in the previous paragraph. This is depicted in Fig. 2.

For each spin-species, the initial state $|\psi_0\rangle$ is given by the superposition of all possible excitations of the three electrons closest to the Fermi level into the 4 closest empty modes starting from the state $|0000\dots 01\dots 1111\rangle$, and realized by the application of 7 Givens rotation gates of rotation angle $\pi/4$.

E. Experiment details

The experiments were run on IBM Quantum's *ibm_fez*, a Heron r2 processor with 156 fixed-frequency transmon qubits with tunable couplers on a heavy-hex lattice layout. $L = 1$ and $K = 29$ required 61 qubits to implement, while $L = 1$ and $K = 41$ required 85 qubits. Each Krylov dimension was sampled with 1×10^5 shots. We implemented $K = 29$ on both the basis with diagonal bath and the \mathbf{k} -adjacent basis. For the \mathbf{k} -adjacent basis, the average median error rates were: readout-error 1.33×10^{-2} , single-qubit error 2.40×10^{-4} and two-qubit gate-error 2.73×10^{-3} , with the latter two characterized by randomized benchmarking. The average relaxation and dephasing times were $138.5 \mu\text{s}$ and $101.25 \mu\text{s}$. Likewise, for the basis of diagonal-bath circuits had readout-error 1.48×10^{-2} , single-qubit error 2.60×10^{-4} and two-qubit gate-error 2.77×10^{-3} . $T_1 = 131 \mu\text{s}$ and $T_2 = 89.25 \mu\text{s}$ were slightly lower for these experiments than for the \mathbf{k} -adjacent basis ones. The number of two-qubit gates grew linearly with Trotter steps/Krylov dimension with a slope of 312 and 208 for the momentum and \mathbf{k} -adjacent basis respectively. The maximum number of two-qubit gates here were 5976 gates and 3281 gates respectively. For $K = 41$, we only considered the \mathbf{k} -adjacent basis. For these runs, we had average $T_1 = 131.5 \mu\text{s}$ and $T_2 = 96.25 \mu\text{s}$. The error rates

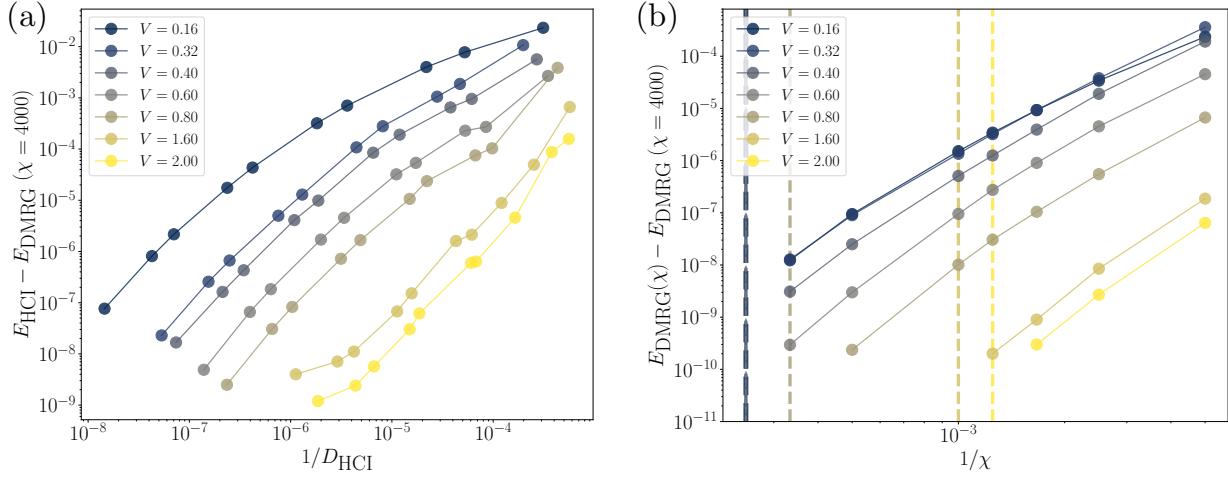


Figure 6. Convergence of HCl and DMRG energy estimations as a function of the corresponding cost control parameter for the 4-impurity model. (a) Scaling of the HCl energy error (relative to the best DMRG energy) versus size of the corresponding CI vector for different values of the hybridization amplitude V . (b) DMRG energy error versus bond dimension using $E_{\text{DMRG}}(\chi = 4000)$ as reference. The vertical dashed lines indicate the value of bond dimension for which the DMRG energy estimation converges.

were: readout-error 1.53×10^{-2} , single-qubit error 2.6×10^{-4} and two-qubit gate-error 2.80×10^{-3} . The two-qubit gates increased linearly at a rate of 337 gates per Trotter step and the maximum number of two-qubit gates was 6153.

For the 4-impurity Anderson model circuits, 28 bath sites were considered. Consequently, the number of qubits was 64. The number of two-qubit gates per Trotter step grew at the rate of 613 gates per step, signifying an increase in the density of entangling operations in comparison to the SIAM circuits. The average relaxation and dephasing times were $T_1 = 155.0 \mu\text{s}$ and $T_1 = 101.3 \mu\text{s}$ respectively. The error rates were: readout-error 8×10^{-3} , single-qubit gate error 2.2×10^{-4} and two-qubit gate error 3.8×10^{-3} . The maximum number of gates used was 5222.

F. Configuration recovery with carry-over

In the self-consistent configuration recovery scheme introduced in Ref. [16], the only information that is passed in between self consistent iterations in the average fermionic mode occupancies in the approximated ground state found by SQD and SKQD. We realize that after finding a preliminary approximation to the ground state, additional information may be used to inform the subsequent diagonalization steps.

In particular, if a basis state $|b_j\rangle$ is identified as relevant in the description of the ground state at the previous recovery iteration ($|\Psi_j| > \tau$), then the

state $|b_j\rangle$ may be included as part of the definition subspaces in subsequent recovery iterations. τ is a threshold used to determine which basis states have an impact in the the description of the ground state. This strategy is only used in the 4-impurity experiments in this manuscript. We set the value of the threshold to $\tau = 10^{-8}$.

**SUPPLEMENTARY INFORMATION FOR: QUANTUM-CENTRIC ALGORITHM FOR
SAMPLE-BASED KRYLOV DIAGONALIZATION**

I. KRYLOV QUANTUM DIAGONALIZATION

Let \mathcal{H} denote an N -dimensional Hilbert space, where $N = 2^n$ and n denotes the number of qubits. Let H be a Hermitian operator on \mathcal{H} with eigenvalues $E_0 \leq E_1 \leq \dots \leq E_{N-1}$ and corresponding orthonormal eigenstates $|\phi_0\rangle, \dots, |\phi_{N-1}\rangle$. Our goal is to estimate E_0 . In general, the eigenvalue problem for an N -dimensional operator is computationally challenging, so we approximate E_0 by considering the corresponding eigenvalue problem in a subspace. In particular, given some initial state $|\psi_0\rangle$, consider the *Krylov subspace* \mathcal{S} spanned by

$$|\psi_k\rangle := e^{-ikH\Delta t} |\psi_0\rangle, \quad k \in \{0, 1, \dots, d-1\}, \quad (\text{S1})$$

for some chosen time step Δt and dimension d .

We then solve the following generalized eigenvalue problem

$$\tilde{\mathbf{H}}v = \tilde{E}\tilde{\mathbf{S}}v, \quad (\text{S2})$$

where v is a coordinate vector corresponding to the state $|\phi\rangle := \sum_{k=0}^{d-1} v_k |\psi_k\rangle \in \mathcal{S}$, and

$$\tilde{\mathbf{H}}_{jk} := \langle \psi_j | H | \psi_k \rangle, \quad \tilde{\mathbf{S}}_{jk} := \langle \psi_j | \psi_k \rangle. \quad (\text{S3})$$

Let

$$\Delta E_j = E_j - E_0 \quad (\text{S4})$$

for each $0 < j < N$ and let

$$|\psi_0\rangle = \sum_{k=0}^{N-1} \gamma_k |\phi_k\rangle \quad (\text{S5})$$

be the eigenstate decomposition of the initial state. For completeness, we recall the accuracy of estimating the ground state energy using the Krylov quantum diagonalization approach from [10].

Theorem 2 (Theorem 3.1 in [10]). *Let $\Delta t = \frac{\pi}{\Delta E_{N-1}}$ and d denote the Krylov dimension as defined in Eq. (S1). Assume that d is odd; for even d we can use the bound for $d-1$. Then the approximate ground state energy \tilde{E}_0 obtained from the method described in Eqs. (S1)–(S3) satisfies*

$$0 \leq \tilde{E}_0 - E_0 \leq 8\Delta E_{N-1} \left(\frac{1 - |\gamma_0|^2}{|\gamma_0|^2} \right) \left(1 + \frac{\pi \Delta E_1}{\Delta E_{N-1}} \right)^{-(d-1)}, \quad (\text{S6})$$

where γ_0 denotes the coefficient of $|\phi_0\rangle$ in Eq. (S5) and ΔE_j is defined in Eq. (S4).

Note that the accuracy in approximating E_0 improves exponentially with the dimension d of the Krylov subspace. Moreover, the accuracy is inversely proportional to the overlap ($|\gamma_0|^2$) of the initial state with the ground state of H . We summarize proof steps for Eq. (S6) below.

- First, note that if d is odd, then the projection of the Hamiltonian into the Krylov subspace as we have defined it is the same as the projection of the Hamiltonian into the shifted Krylov space spanned by

$$|\psi_k\rangle = e^{-ikH\Delta t} |\psi_0\rangle, \quad k \in \left\{ -\frac{d-1}{2}, -\frac{d-1}{2} + 1, \dots, \frac{d-1}{2} - 1, \frac{d-1}{2} \right\}. \quad (\text{S7})$$

This follows because the Gram matrices $\tilde{\mathbf{S}}$ of the two spaces are identical since time evolutions are unitary, and the Hamiltonian projection matrices $\tilde{\mathbf{H}}$ are identical because the time evolutions commute with the full Hamiltonian H .

- For any $0 < a < \pi$ and positive integer d , there exists a trigonometric polynomial p^* of degree $d_{\text{poly}} = \frac{d-1}{2}$ satisfying $p^*(0) = 1$ and

$$|p^*(\theta)| \leq 2(1+a)^{-d_{\text{poly}}} \quad \forall \theta \in (-\pi, \pi) \setminus (-a, a). \quad (\text{S8})$$

This polynomial is explicitly constructed as

$$p^*(\theta) = \frac{T_k\left(1 + 2\frac{\cos\theta - \cos a}{\cos a + 1}\right)}{T_k\left(1 + 2\frac{1 - \cos a}{\cos a + 1}\right)}, \quad (\text{S9})$$

where T_k is the k th-Chebyshev polynomial.

- Let the Fourier expansion of p^* be

$$p^*((E - E_0)\Delta t) = \sum_{k=-d_{\text{poly}}}^{d_{\text{poly}}} c_k e^{ikE\Delta t}. \quad (\text{S10})$$

Consider the following unnormalized state:

$$|\tilde{\phi}_K\rangle = \sum_{k=-d_{\text{poly}}}^{d_{\text{poly}}} c_k |\psi_k\rangle = \sum_{k=-d_{\text{poly}}}^{d_{\text{poly}}} \sum_{j=0}^{N-1} c_k \gamma_j e^{ikE_j \Delta t} |\phi_j\rangle. \quad (\text{S11})$$

Moreover, from Section 3.1 in [10] we recall that $\sum_k |c_k|^2 \leq 1$.

- The norm of $|\tilde{\phi}_K\rangle$ is given by

$$\left| \langle \tilde{\phi}_K | \tilde{\phi}_K \rangle \right|^2 = \sum_{j=0}^{N-1} |\gamma_j|^2 \left| \sum_{k=-d_{\text{poly}}}^{d_{\text{poly}}} c_k e^{ikE_j \Delta t} \right|^2 \geq |\gamma_0|^2 \left| \sum_{k=-d_{\text{poly}}}^{d_{\text{poly}}} c_k e^{ikE_0 \Delta t} \right|^2 = |\gamma_0|^2 |p^*(0)|^2 = |\gamma_0|^2. \quad (\text{S12})$$

- Then the energy error of $|\tilde{\phi}_K\rangle$ is given by

$$\frac{\langle \tilde{\phi}_K | (H - E_0) | \tilde{\phi}_K \rangle}{\langle \tilde{\phi}_K | \tilde{\phi}_K \rangle} \leq \frac{\sum_{j=0}^{N-1} (E_j - E_0) |\gamma_j|^2 \left| \sum_{k=-d_{\text{poly}}}^{d_{\text{poly}}} e^{ikE_j \Delta t} \right|^2}{\langle \tilde{\phi}_K | \tilde{\phi}_K \rangle} \leq \frac{\sum_{j=0}^{N-1} E_j |\gamma_j|^2 |p^*((E_j - E_0)\Delta t)|^2}{|\gamma_0|^2}. \quad (\text{S13})$$

Applying (S8) with $a = (E_1 - E_0)\Delta t = \frac{\pi \Delta E_1}{\Delta E_{N-1}}$ gives

$$p^*((E_j - E_0)\Delta t) \leq 2 \left(1 + \frac{\pi \Delta E_1}{\Delta E_{N-1}}\right)^{-d} \quad (\text{S14})$$

for each $1 \leq j \leq N-1$. The $j=0$ term cancels upon taking the difference with E_0 for the energy error.

- Since $|\tilde{\phi}_K\rangle$ is explicitly defined in Eq. (S11) as an element of the shifted Krylov space Eq. (S7), the lowest energy in the shifted Krylov space is upper bounded by the energy of $|\tilde{\phi}_K\rangle$. Finally, as we noted above, the projection of the Hamiltonian into the shifted Krylov space is identical to its projection into the original, unshifted Krylov space, so the lowest energies of the two projections are the same. Hence, the energy error from finding the lowest energy of the projected Hamiltonian in our Krylov subspace is upper bounded by the energy error Eq. (S13) of $|\tilde{\phi}_K\rangle$.

Thus, for an initial state with a nontrivial ground state overlap $|\gamma_0|^2 = \Theta(1)$ and for a Hamiltonian H that has a well-behaved spectrum, (i.e., ΔE_{N-1} not growing too quickly and ΔE_1 not too small), a constant error $\tilde{E}_0 - E_0 \leq \varepsilon$ can be achieved by taking $d = O(\log(1/\varepsilon))$.

II. PROOFS AND RELEVANT DETAILS FOR THEOREM 1

In this section we prove the performance guarantees for the Krylov diagonalization via quantum unitary sampling approach. As discussed in the main text, we prepare d different states on a quantum computer and collect M samples from each of them by measuring in the computational basis. In particular, let $|\psi_0\rangle$ denote the initial reference state. Then we get M samples from each Krylov basis state $|\psi_k\rangle$ as defined in Eq. (S1). In particular, we get a sequence of bitstrings $\{a_{km}\}_{k=0}^{M-1}$ for each k . Finally, we solve the eigenvalue problem in the subspace spanned by $\{a_{km} \mid k = 0, 1, \dots, d-1; m = 0, 1, \dots, M-1\}$.

We analyze our algorithm as follows. We first recall Definition 1 for convenience:
Definition 1 $((\alpha_L, \beta_L)$ -sparsity). For any state $|\psi\rangle$, let

$$|\psi\rangle = \sum_{j=1}^N g_j |b_j\rangle, \quad (\text{S15})$$

where $\{b_j\}$ is some ordering of length- N bitstrings such that $|g_1| \geq |g_2| \geq \dots \geq |g_N|$. We say that $|\psi\rangle$ exhibits (α_L, β_L) -sparsity on $|b_1\rangle$ through $|b_L\rangle$ if

$$\sum_{j=1}^L |g_j|^2 \geq \alpha_L \quad (\text{S16})$$

and

$$|g_1|^2, \dots, |g_L|^2 \geq \beta_L. \quad (\text{S17})$$

Since $|g_i|^2 \geq \beta_L, \forall i$, it implies that $\alpha_L \geq L\beta_L$, but α_L may be much larger in general. Here, we separate the two as α_L will govern the rate of convergence for successful runs while β_L will govern the probability of success.

Below, we restate Theorem 1 and prove it using several lemmas in Section II 1.

Theorem 1. *Let H be a Hamiltonian whose ground state $|\phi_0\rangle$ exhibits $(\alpha_L^{(0)}, \beta_L^{(0)})$ -sparsity. Let $|\tilde{\phi}\rangle$ be the lowest energy state supported on the L important bitstrings in $|\phi_0\rangle$. The error in estimating the ground state energy of H is bounded by*

$$\langle \tilde{\phi} | H | \tilde{\phi} \rangle - \langle \phi_0 | H | \phi_0 \rangle \leq \sqrt{8} \|H\| \left(1 - \sqrt{\alpha_L^{(0)}}\right)^{1/2},$$

provided all L important bitstrings are sampled. The success probability of sampling all L important bitstrings is at least $1 - \eta$ as long as the number of samples from each Krylov basis state exceeds $(d^2 \log(L/\eta)) / (|\gamma_0|^2(\beta_L^{(0)} - 2\sqrt{\tilde{\varepsilon}}))$, where

$$\tilde{\varepsilon} = 2 - 2\sqrt{1 - \varepsilon/\Delta E_1}$$

with ε defined in Eq. (5).

As discussed in Section I, we choose d according to (S6) so that the error for the Krylov quantum diagonalization approach is bounded by ε . If we were to run the Krylov method on the $\{|\psi_k\rangle\}$, we would obtain a state $|\psi\rangle \in \mathcal{S}$ with

$$0 \leq \langle \psi | H | \psi \rangle - \langle \phi_0 | H | \phi_0 \rangle = \tilde{E}_0 - E_0 < \varepsilon. \quad (\text{S18})$$

By the proof of Theorem 2, there exists $|\tilde{\phi}_K\rangle \in \mathcal{S}$ defined by (S11), which is not necessarily the lowest energy state in \mathcal{S} , but which satisfies (S6) and also satisfies the following inequality :

$$|\tilde{\phi}_K\rangle = \sum_{k=0}^{d-1} d_k |\psi_k\rangle, \quad |d_k| \leq \frac{1}{|\gamma_0|}, \quad (\text{S19})$$

where $|\widehat{\phi}_K\rangle$ is the normalized version of $|\widetilde{\phi}_K\rangle$ in Eq. (S11). Therefore, $d_k = c_k / |\langle \widetilde{\phi}_K | \widetilde{\phi}_K \rangle|$, which implies that $|d_k| = |c_k| / |\langle \widetilde{\phi}_K | \widetilde{\phi}_K \rangle| \leq 1 / |\gamma_0|$, where we invoked the condition that $\sum_k |c_k|^2 \leq 1$ (from Section 3.1 in [10]) and $|\langle \widetilde{\phi}_K | \widetilde{\phi}_K \rangle| \geq |\gamma_0|$.

Note that to prove Eq. (S18), we do not actually need to solve the generalized eigenvalue problem to obtain $|\psi\rangle$, but only to have a state satisfying (S6), so we will use $|\widetilde{\phi}_K\rangle$ for our analysis due to its additional structure as in Eq. (S19).

In Lemma 1, we prove that a state that has a low energy error with respect to the ground state energy is also close to the ground state in 2-norm distance. We then show in Lemma 2 that if $|\phi_0\rangle$ exhibits sparsity and if $|\psi\rangle$ is close to $|\phi_0\rangle$, then $|\psi\rangle$ also exhibits sparsity.

Lemma 1 (A state with low energy is close to the ground state). *Let H be a Hamiltonian with ground state $|\phi_0\rangle$ and spectral gap ΔE_1 . If $|\psi\rangle$ is a state such that $\langle\psi|H|\psi\rangle - \langle\phi_0|H|\phi_0\rangle < \varepsilon$ and $\langle\psi|\phi_0\rangle$ is real, then*

$$\|\psi\rangle - |\phi_0\rangle\|^2 < 2 \left(1 - \sqrt{1 - \frac{\varepsilon}{\Delta E_1}}\right) = O\left(\frac{\varepsilon}{\Delta E_1}\right). \quad (\text{S20})$$

Proof. Let $|\psi\rangle = \chi_0 |\phi_0\rangle + \chi^\perp |\phi^\perp\rangle$, where $|\phi^\perp\rangle$ is normalized and orthogonal to $|\phi_0\rangle$, such that $|\chi_0|^2 + |\chi^\perp|^2 = 1$. We can write $\langle\psi|H|\psi\rangle - \langle\phi_0|H|\phi_0\rangle < \varepsilon$ as

$$\varepsilon > |\chi_0|^2 E_0 + |\chi^\perp|^2 \langle\phi^\perp|H|\phi^\perp\rangle - E_0 = (|\chi_0|^2 - 1)E_0 + (1 - |\chi_0|^2) \langle\phi^\perp|H|\phi^\perp\rangle \quad (\text{S21})$$

$$\geq (1 - |\chi_0|^2) \Delta E_1. \quad (\text{S22})$$

Thus $|\chi_0|^2 > 1 - \frac{\varepsilon}{\Delta E_1}$, giving

$$\|\psi\rangle - |\phi_0\rangle\|^2 = |\chi_0 - 1|^2 + |\chi^\perp|^2 = 2 - 2\chi_0 < 2 - 2\sqrt{1 - \frac{\varepsilon}{\Delta E_1}}. \quad (\text{S23})$$

Expanding to the first order in $\frac{\varepsilon}{\Delta E_1}$ gives an error bound of $\frac{\varepsilon}{\Delta E_1} + O(\varepsilon^2)$. \square

Lemma 2. *If $|\phi_0\rangle$ exhibits $(\alpha_L^{(0)}, \beta_L^{(0)})$ -sparsity and $\|\psi\rangle - |\phi_0\rangle\|^2 < \varepsilon$, then $|\psi\rangle$ exhibits (α_L, β_L) -sparsity for*

$$\alpha_L = \alpha_L^{(0)} - 2\sqrt{\varepsilon} \quad \beta_L = \beta_L^{(0)} - 2\sqrt{\varepsilon}. \quad (\text{S24})$$

Proof. Expand $|\psi\rangle$ and $|\phi_0\rangle$ in the computational basis as $|\psi\rangle = \sum_{j=1}^N a_j |b_j\rangle$ and $|\phi_0\rangle = \sum_{j=1}^N c_j |b_j\rangle$, respectively. Then

$$\sum_{j=1}^L |a_j|^2 = \sum_{j=1}^L |c_j + (a_j - c_j)|^2 = \sum_{j=1}^L |c_j|^2 + \sum_{j=1}^L |a_j - c_j|^2 + 2 \sum_{j=1}^L \operatorname{Re}(c_j(a_j - c_j)) \quad (\text{S25})$$

$$\geq \alpha_L^{(0)} + 0 - 2 \sum_{j=1}^L |c_j| |a_j - c_j| \quad (\text{S26})$$

$$\geq \alpha_L^{(0)} - 2 \sqrt{\sum_j |c_j|^2 \sum_k |a_k - c_k|^2} \quad (\text{S27})$$

$$\geq \alpha_L^{(0)} - 2\sqrt{\varepsilon}, \quad (\text{S28})$$

showing that we may take $\alpha_L = \alpha_L^{(0)} - 2\sqrt{\varepsilon}$. In the last line we replaced $\sum_{j=1}^L |c_j|^2 \leq 1$ and used the fact that $\|\psi\rangle - |\phi_0\rangle\|^2 < \varepsilon$.

We now prove a similar condition for β_L . Since $\|\psi\rangle - |\phi_0\rangle\|^2 < \varepsilon$, we must have $|a_j - c_j| < \sqrt{\varepsilon}$ for each $1 \leq j \leq N$. Then for each $1 \leq j \leq L$, we have

$$|a_j|^2 = |c_j + (a_j - c_j)|^2 = |c_j|^2 + |a_j - c_j|^2 + 2 \operatorname{Re}(c_j(a_j - c_j)) \quad (\text{S29})$$

$$\geq \beta_L^{(0)} + 0 - 2 |c_j| |a_j - c_j| \quad (\text{S30})$$

$$\geq \beta_L^{(0)} - 2\sqrt{\varepsilon}. \quad (\text{S31})$$

Thus we may take $\beta_L = \beta_L^{(0)} - 2\sqrt{\varepsilon}$. \square

We have shown in Lemma 2 that if the Krylov quantum diagonalization (KQD) approach has ε energy error and if the true ground state exhibits $(\alpha_L^{(0)}, \beta_L^{(0)})$ -sparsity, then a solution to the KQD approach exhibits (α_L, β_L) -sparsity for α_L, β_L given by (S24). Using this result, we now prove in Section II and Lemma 4 that with a high probability, each bitstring (among the L important bitstrings) corresponding to the ground state $|\phi_0\rangle$ appears nontrivially in at least one of the Krylov basis states. This is crucial for establishing a bound on the error corresponding the Krylov diagonalization via quantum unitary sampling approach.

Lemma 3 (Each relevant bitstring appears nontrivially in at least one Krylov basis state). *Let $|\hat{\phi}_K\rangle$ satisfy (α_L, β_L) -sparsity as in Definition 1 and $|\hat{\phi}_K\rangle = \sum_{k=0}^{d-1} d_k |\psi_k\rangle$ as in Eq. (S19). If each $|\psi_k\rangle$ is represented in the computational basis as $|\psi_k\rangle = \sum_{j=1}^N c_j^{(k)} |b_j\rangle$ for each $k = 0, 1, \dots, d-1$, then for each $j = 0, 1, \dots, L-1$ there exists some $c_j^{(k)}$ such that $|c_j^{(k)}|^2 \geq p$ with*

$$p = \frac{|\gamma_0|^2 \beta_L}{d^2}. \quad (\text{S32})$$

Here, $|\gamma_0|^2$ denotes the overlap of the KQD initial state $|\psi_0\rangle$ with the ground state, as defined in Eq. (S1).

Proof. Let $|\hat{\phi}_K\rangle = \sum_{j=1}^N a_j |b_j\rangle$ in the computational basis. For each $1 \leq j \leq L$, we have

$$\sum_{k=0}^{d-1} |c_j^{(k)}| \geq \sum_{k=0}^{d-1} |\gamma_0| |d_k| |c_j^{(k)}| \geq |\gamma_0| \left| \sum_{k=0}^{d-1} d_k c_j^{(k)} \right| = |\gamma_0| |a_j| \geq |\gamma_0| \sqrt{\beta_L}, \quad (\text{S33})$$

where the first inequality follows from $|d_k| |\gamma_0| \leq 1$, as in Eq. (S19), and the last inequality follows from Lemma 2. Therefore, Eq. (S33) implies that there must be some k for which $|c_j^{(k)}| \geq \frac{|\gamma_0| \sqrt{\beta_L}}{d}$, or equivalently $|c_j^{(k)}|^2 \geq \frac{|\gamma_0|^2 \beta_L}{d^2}$, as desired. \square

Lemma 4. *If we make M measurements from each $|\psi_k\rangle$, the probability of not obtaining all $|b_0\rangle, |b_1\rangle, \dots, |b_{L-1}\rangle$ among the sampled bitstrings is bounded by*

$$p_{\text{fail}} \leq L(1-p)^M \leq L e^{-Mp}. \quad (\text{S34})$$

Proof. The M measurements are independent, so the probability of not obtaining a particular bitstring $|b_j\rangle$ for $j \in \{0, 1, \dots, L-1\}$ is $(1-p)^M$. The probability of not obtaining at least one of the L bitstrings follows from union bound. \square

We now recall a result from [16] relating the energy of a state defined in a subspace to the ground state energy on the full n -qubit space.

Lemma 5 (Appendix B.1 from [16]). *Let $|\tilde{\phi}\rangle = \frac{1}{C} \sum_{j=0}^{L-1} c_j |b_j\rangle$, where $C = \sqrt{\sum_{j=0}^{L-1} |c_j|^2}$ is the normalization constant and $\sum_{j=0}^{N-1} c_j |b_j\rangle$ defines the ground state in the computational basis in decreasing order of coefficient magnitude, as in Definition 1. This state has energy close to the ground state energy, with difference bounded by*

$$\langle \tilde{\phi} | H | \tilde{\phi} \rangle - \langle \phi_0 | H | \phi_0 \rangle \leq 2\sqrt{2} \|H\| \left(1 - \sqrt{\alpha_L^{(0)}}\right)^{1/2}. \quad (\text{S35})$$

Proof. First, we rewrite the error in energy. Let $|\phi'\rangle = |\tilde{\phi}\rangle - |\phi_0\rangle$. Then we get

$$\langle \tilde{\phi} | H | \tilde{\phi} \rangle - \langle \phi_0 | H | \phi_0 \rangle = \langle \tilde{\phi} | H | \phi' \rangle + \langle \phi' | H | \phi_0 \rangle. \quad (\text{S36})$$

Then we have

$$\langle \tilde{\phi} | H | \tilde{\phi} \rangle - \langle \phi_0 | H | \phi_0 \rangle \leq |\langle \tilde{\phi} | H | \phi' \rangle| + |\langle \phi' | H | \phi_0 \rangle| \leq 2 \|\tilde{\phi}\| \|H|\phi'\| \leq 2 \cdot 1 \cdot \|H\| \|\phi'\|. \quad (\text{S37})$$

Finally, we calculate the norm difference as

$$\left\| |\tilde{\phi}\rangle - |\phi_0\rangle \right\|^2 = \sum_{j=0}^{L-1} \left(\frac{1}{C} - 1 \right)^2 |c_j|^2 + \sum_{L}^{N-1} |c_j|^2 = \left(1 - \frac{2}{C} + \frac{1}{C^2} \right) C^2 + (1 - C^2) = 2 - 2C, \quad (\text{S38})$$

and we have $C \geq \sqrt{\alpha_L^{(0)}}$ by (S16), so $\left\| |\tilde{\phi}\rangle - |\phi_0\rangle \right\| \leq \sqrt{2} \left(1 - \sqrt{\alpha_L^{(0)}} \right)^{1/2}$. Since $|\phi'\rangle$ is defined to be $|\tilde{\phi}\rangle - |\phi_0\rangle$, plugging the above into (S37) completes the proof. \square

1. Proof of Theorem 1

From Theorem 2, we have a state $|\psi\rangle$ with $\langle \psi | H | \psi \rangle - \langle \phi_0 | H | \phi_0 \rangle < \varepsilon$ with

$$\varepsilon = 8\Delta E_{N-1} \left(\frac{1 - |\gamma_0|^2}{|\gamma_0|^2} \right) \left(1 + \frac{\pi\Delta E_1}{\Delta E_{N-1}} \right)^{-(d-1)}. \quad (\text{S39})$$

Let

$$\varepsilon' = \sqrt{1 - \varepsilon/\Delta E_1}. \quad (\text{S40})$$

Then, from Lemma 1, we get

$$\left\| |\psi\rangle - |\phi_0\rangle \right\|^2 < 2 - 2\varepsilon'. \quad (\text{S41})$$

Given $|\phi_0\rangle$ exhibits $(\alpha_L^{(0)}, \beta_L^{(0)})$ -sparsity, by Lemma 2 $|\psi\rangle$ exhibits (α_L, β_L) -sparsity with parameters

$$\alpha_L = \alpha_L^{(0)} - 2\sqrt{2 - 2\varepsilon'} \quad \beta_L = \beta_L^{(0)} - 2\sqrt{2 - 2\varepsilon'} \quad (\text{S42})$$

Hence by Lemma 3, for each of the L important bitstrings b_j with $j = 0, 1, \dots, L-1$, we will be able to sample b_j with probability at least

$$p = \frac{|\gamma_0|^2}{d^2} \left(\beta_L^{(0)} - 2\sqrt{2 - 2\varepsilon'} \right) \quad (\text{S43})$$

from at least one of the $|\psi_k\rangle$. Given M measurements, the probability of failing to sample b_j is at most $(1-p)^M$. Repeating this for each bitstring in $\{b_j\}_{j=0}^{L-1}$, the probability of failing to sample all L bitstrings b_1 through b_L is

$$p_{\text{fail}} \leq L(1-p)^M \leq L e^{-Mp} = L \exp \left(-\frac{M|\gamma_0|^2}{d^2} \left(\beta_L^{(0)} - 2\sqrt{2 - 2\sqrt{1 - \frac{\varepsilon}{\Delta E_1}}} \right) \right) \quad (\text{S44})$$

by union bound and by using p from Eq. (S43) and ε' from Eq. (S40). If we succeed in sampling all L bitstrings, then the state $|\tilde{\phi}\rangle$ as defined in Lemma 5 exists in the sampled subspace, so the calculated energy will be bounded by

$$\langle \tilde{\phi} | H | \tilde{\phi} \rangle - \langle \phi_0 | H | \phi_0 \rangle \leq 2\sqrt{2} \|H\| \left(1 - \sqrt{\alpha_L^{(0)}} \right)^{1/2}, \quad (\text{S45})$$

which completes the proof of Theorem 1.

III. SPARSITY IN THE ISING MODEL

In this section, we prove the sparsity of the ground state in the computational basis for a particular Hamiltonian. We consider the transverse field Ising model with periodic boundary conditions

$$H_n(h) = - \sum_{i=0}^{n-1} Z_i Z_{i+1} - h \sum_{i=0}^{n-1} X_i. \quad (\text{S46})$$

Theorem 3. If $h = O((k/n)^a)$ for any $a > 1/2$, then in the limit $n \rightarrow \infty$ the ground state of $H_n(h)$ is fully supported on the $O(n^k)$ Z-basis states with Hamming weight at most k .

Proof. We will ignore the degeneracy of the ground state for simplicity, but the result holds without this assumption by symmetry. Let $|\phi_n(h)\rangle$ be the ground state of $H_n(h)$. With this assumption, the ground state is $|\phi_n(0)\rangle = |00\dots 0\rangle$ and $\lim_{h \rightarrow \infty} |\phi_n(h)\rangle = |++\dots+\rangle$. Thus, intuitively, h controls the sparsity of the ground state in the computational basis.

Let $|x|$ be the Hamming weight of the bit string x , and define

$$M_n(h) = \frac{1}{n} \sum_{i=0}^{n-1} \langle \phi_n(h) | Z_i | \phi_n(h) \rangle \quad (\text{S47})$$

$$M(h) = \lim_{n \rightarrow \infty} M_n(h) \quad (\text{S48})$$

$$\bar{S}_n(k, h) = \sum_{\substack{x \in \{0,1\}^n \\ |x| \leq k}} |\langle x | \phi_n(h) \rangle|^2 \quad (\text{S49})$$

$$S_n(k, h) = 1 - \bar{S}_n(k, h). \quad (\text{S50})$$

$S_n(k, h)$ is a proxy for sparsity – it being small implies that there is very little weight on states outside of the d -dimensional subspace defined by Hamming weight less than k , where $d = \sum_{w=0}^k \binom{n}{w}$.

Define α_x by $|\phi_n(h)\rangle = \sum_{x \in \{0,1\}^n} \alpha_x |x\rangle$. Furthermore, define $\bar{P}_n(w, h) = \sum_{\substack{x \in \{0,1\}^n \\ |x|=w}} |\alpha_x|^2$, so that $\bar{S}_n(k, h) = \sum_{w=0}^k \bar{P}_n(w, h)$. Then

$$M_n(h) = \frac{1}{n} \sum_{i=0}^{n-1} \sum_{x,y \in \{0,1\}^n} \bar{\alpha}_x \alpha_y \langle x | Z_i | y \rangle \quad (\text{S51})$$

$$= \frac{1}{n} \sum_{i=0}^{n-1} \sum_{x \in \{0,1\}^n} |\alpha_x|^2 (-1)^{x_i} \quad (\text{S52})$$

$$= \frac{1}{n} \sum_{w=0}^n \sum_{\substack{x \in \{0,1\}^n \\ |x|=w}} |\alpha_x|^2 \sum_{i=0}^{n-1} (-1)^{x_i} \quad (\text{S53})$$

$$= \frac{1}{n} \sum_{w=0}^n \sum_{\substack{x \in \{0,1\}^n \\ |x|=w}} |\alpha_x|^2 (n - 2w) \quad (\text{S54})$$

$$= \sum_{w=0}^n \bar{P}_n(w, h) (1 - 2w/n) \quad (\text{S55})$$

$$= \sum_{w=0}^k \bar{P}_n(w, h) (1 - 2w/n) + \sum_{w=k+1}^n \bar{P}_n(w, h) (1 - 2w/n) \quad (\text{S56})$$

$$\leq \bar{S}_n(k, h) + \sum_{w=k+1}^n \bar{P}_n(w, h) (1 - 2w/n) \quad (\text{S57})$$

$$\leq \bar{S}_n(k, h) + (1 - 2(k+1)/n) \sum_{w=k+1}^n \bar{P}_n(w, h) \quad (\text{S58})$$

$$= \bar{S}_n(k, h) + (1 - 2(k+1)/n) S_n(k, h) \quad (\text{S59})$$

$$= 1 - \frac{2(k+1)}{n} S_n(k, h). \quad (\text{S60})$$

It follows that

$$S_n(k, h) \leq \min \left(\frac{n(1 - M_n(h))}{2k + 2}, 1 \right). \quad (\text{S61})$$

If $S_n(k, h) \rightarrow 0$ as $n \rightarrow \infty$, then the subspace of Hamming weight $\leq k$ bitstrings is fully capturing the ground state. The dimension of this space is $d = \sum_{w=0}^k \binom{n}{w}$, which is $\sim n^k$ for constant k .

Eq. (S61) implies that if $M_n(h) = 1 - O((k/n)^a)$ for any $a > 1$, the ground state is fully supported on $O(n^k)$ bitstrings (ie. if $M_n(h) = 1 - O((k/n)^a)$ for $a > 1$, then $S_n(k, h)$ decays to zero with increasing n). From the phase diagram of the transverse field Ising model, we know [65, Eq. 3.12] that $M(h) = (1 - h^2)^{1/8}$ for $0 \leq h \leq 1$, and $M_n(h) \rightarrow M(h)$ converges continuously with $n \rightarrow \infty$. It follows that $h^2 = O((k/n)^a)$ for $a > 1$ suffices. This completes the proof. \square

Corollary 1. *If $h = O((n/k)^a)$ for any $a > 1/2$, then in the limit $n \rightarrow \infty$ the ground state of $H_n(h)$ is fully supported on $O(n^k)$ X-basis states.*

Proof. Apply a Hadamard matrix on each qubit to get a transformed Hamiltonian $H'_n(h) = -\sum_{i=0}^{n-1} X_i X_{i+1} - h \sum_{i=0}^{n-1} Z_i$. Then, by Theorem 3, as $n \rightarrow \infty$ the ground state of $H'_n(h)$ is fully supported on $O(n^k)$ X-basis states if $1/h = O((k/n)^a)$ for any $a > 1/2$. \square

Theorem 3 proves that the ground state of the transverse field Ising model is sparse in a product state basis (the Z-basis) deep in the ordered phase. Meanwhile, Corollary 1 proves that the ground state is sparse in a product state basis (the X-basis) deep in the disordered phase.

IV. COMPARISON WITH ALTERNATIVE NOTIONS OF SPARSITY

Note that our definition of sparseness (or peakedness), as defined in Eq. (S16) and Eq. (S17) differs from those in [55, 56]. In [56], a unitary circuit C is defined as δ -peaked if there exists at least one bitstring $s \in \{0, 1\}^n$ such that $|\langle s | C | 0 \rangle|^2 \geq \delta$. The focus of [56] was on random peaked circuits and determining whether such circuits can be distinguished from fully random circuits in classical polynomial time. Similarly, [55] defines a circuit family $\{U_n\}$ as peaked if for some $a \in \mathbb{Z}_{\geq 0}$, each U_n is δ -peaked for $n \geq 0$ with $\delta = n^{-a}$. In this context, an n -qubit circuit is considered peaked if it has an output probability that is at least inverse-polynomial in n . In [55], the authors developed classical algorithms for sampling and estimating output probabilities of constant-depth peaked quantum circuits. In contrast, our definition of sparsity requires the weight to be concentrated on L bitstrings, rather than requiring at least one bitstring to have high probability.

V. COMPARISON BETWEEN THE SAMPLE COMPLEXITY IN SKQD AND KQD IN THE TRANSVERSE FIELD ISING MODEL

To establish an insight on the practical performance of SKQD, we first showcase numerical simulations on a lattice model. These simulations use the shifted Krylov space given by Eq. (S7) and the usual $\Delta t = \pi / \Delta E_{N-1}$ in the context of Theorem 2.

We consider a perturbed transverse field Ising model

$$H = -\sum_{j=1}^{n-1} Z_j Z_{j+1} - h_1 \sum_{j=1}^n X_j - h_2 Z_1. \quad (\text{S62})$$

When $h_1 = h_2 = 0$, the ground states are spanned by the bitstrings $|0^n\rangle$ and $|1^n\rangle$. A positive h_2 breaks the degeneracy in favor for $|0^n\rangle$. In Section III, for $h_1 = h$ and $h_2 = 0$, we show that if $h = O((k/n)^a)$ for any $a > 1/2$, then the ground state of $H_n(h)$ is fully supported on $O(n^k)$ bitstrings.

For our numerical simulations, we use the initial state $|\chi_0\rangle = |0^n\rangle$. In Figure S1, we compare the performance of SKQD approach with the standard KQD approach. To ensure a fair comparison, we add a Gaussian noise $\mathcal{N}(0, \frac{1}{\sqrt{M}})$ to each matrix element, as described in Eq. (11). Let $h_1 = h_2 = 0.1$. We run the SKQD approach for $d = 15$ different Krylov basis states and for varying numbers of samples from each basis state. We perform simulations with different numbers of qubits, as shown in Fig. S1. We set $M = 5000$ in the KQD approach while computing each matrix element. Moreover, we selected the best instance for the SKQD approach from 1000 trials. Fig. S1 demonstrates that our approach (SKQD) outperforms the standard KQD approach across different numbers of qubits. Thus, our numerical simulations extend beyond the analytical results, showing that the SKQD approach can outperform the standard KQD approach under the sparsity assumption on the ground states.

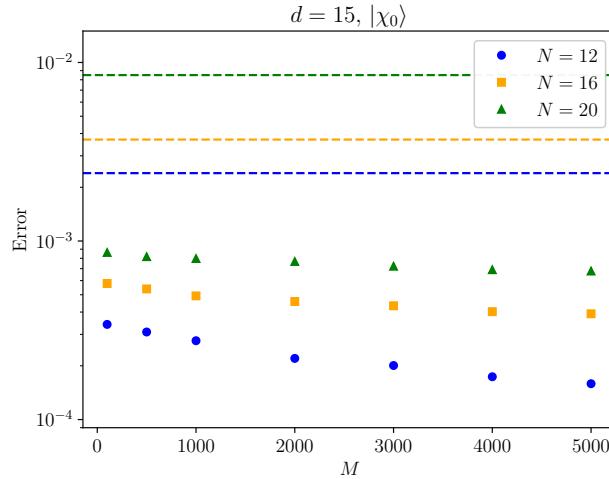


Figure S1. **Comparison of SKQD and KQD methods.** For a perturbed transverse field Ising model Hamiltonian H with equal transverse field and perturbation $h_1 = h_2 = 0.1$, and initial state $|\chi_0\rangle = |0^n\rangle$, the SKQD approach (markers) achieves lower error compared to the standard KQD method. We evaluate the SKQD approach with varying numbers of samples measured per Krylov state, denoted by M , and set the number of Krylov basis states to $d = 15$. In the KQD approach (dashed lines), we incorporate Gaussian noise with a standard deviation of $\frac{1}{\sqrt{M}}$, where $M = 5000$, while estimating matrix elements of H .

VI. EFFECT OF TROTTER ERROR IN SKQD

In this section, we briefly analyze the effect of Trotter error on the performance of SKQD. We assume that $U_k \equiv e^{-ik\Delta t H}$ can be approximated with γ Trotter error for all $k \in \{0, 1, \dots, d-1\}$. Let V_k denote the Trotter approximation for U_k . Then the following inequalities hold:

$$\|U_k - V_k\| \leq \gamma, \quad (\text{S63})$$

$$\implies \|U_k |\psi_0\rangle - V_k |\psi_0\rangle\|_1 \leq \gamma, \quad (\text{S64})$$

$$\implies \|p_k - \hat{p}_k\|_1 \leq \gamma, \quad (\text{S65})$$

$$\implies |p_k^{(j)} - \hat{p}_k^{(j)}| \leq \gamma, \quad (\text{S66})$$

$$\implies \hat{p}_k^{(j)} \geq p_k^{(j)} - \gamma, \quad (\text{S67})$$

where the third inequality follows from monotonicity of the trace distance. Here, $p_k(\hat{p}_k)$ is the distribution on computational basis state for the state $U_k |\psi_0\rangle$ ($V_k |\psi_0\rangle$). Moreover $q_k^{(j)}$ denotes the probability of bitstring j base on distribution q_k .

Then by combining Eq. (S67) with Lemma 3 and Lemma 4, we get that the number of shots (M) needed for sampling all relevant L bitstrings with a high probability (larger than $1 - \eta$) scales as

$$M \geq \frac{\log L/\eta}{(|\gamma_0|^2 \beta_L / d^2 - \gamma)}. \quad (\text{S68})$$

VII. ADDITIONAL SKQD EXPERIMENTS FOR THE SINGLE IMPURITY ANDERSON MODEL

The aim of this section is to study the accuracy of SKQD for in the SIAM for different system sizes than the one shown in the main text. In particular we consider the $K = 29$ bath-site model. As in the main text, we benchmark the accuracy of SKQD against DMRG. Each DMRG run performed 20 sweeps. The first four sweeps have a maximum bond dimension of 250, the next four sweeps have a maximum bond dimension of 400, and the remaining 12 sweeps a maximum bond dimension of 500. At each sweep we add noise of

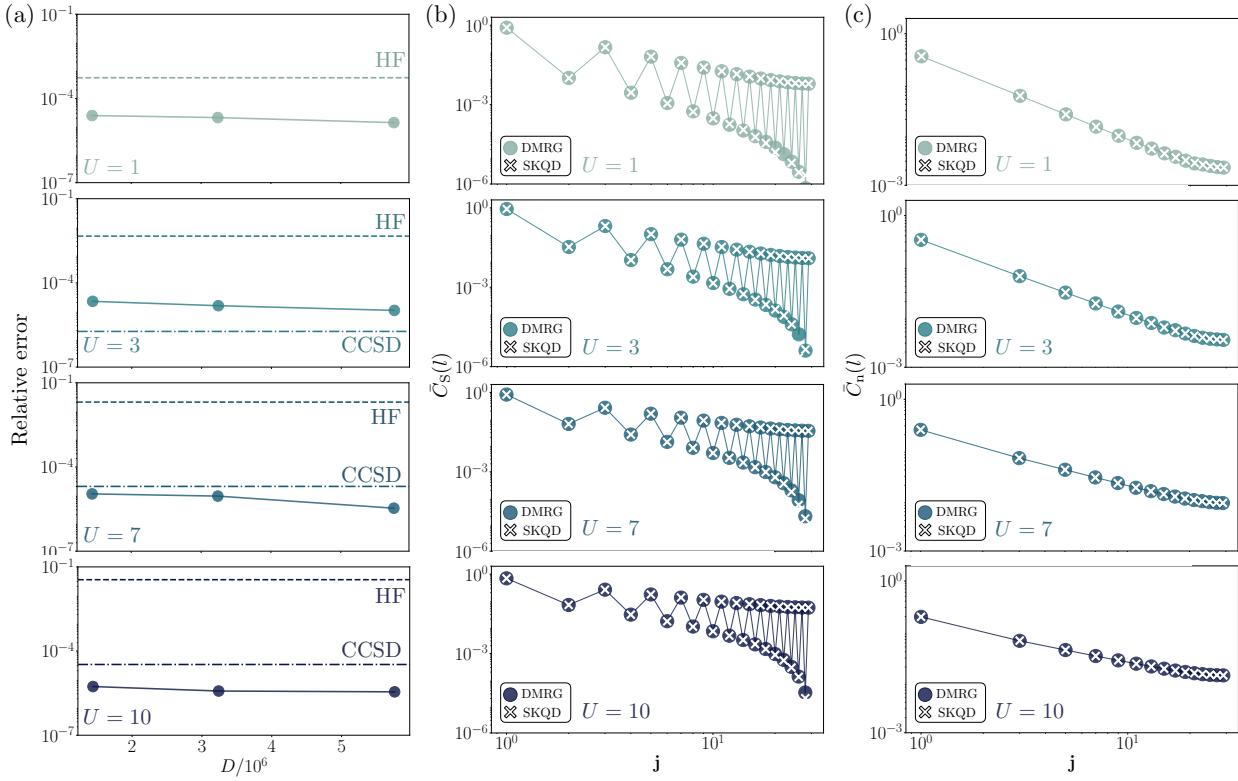


Figure S2. SKQD vs DMRG in the SIAM with 29 bath sites (61-qubit experiment). Different rows correspond to different values of the onsite repulsion U in the impurity. (a) Relative error in the ground state energy estimation using SKQD, as a function of the subspace dimension D . The DMRG estimation is taken as the ground truth. The Hartree-Fock (HF) and coupled cluster with single and double excitations (CCSD) are also included for reference. The dots correspond to the SKQD estimation in the \mathbf{k} -adjacent natural orbitals. (b) Comparison of the two-point spin correlation functions (see Eq. 22) obtained with DMRG and SKQD. (c) Comparison of the two-point density correlation functions (see Eq. 23) obtained with DMRG and SKQD.

amplitude 10^{-4} in the first four sweeps, 10^{-5} in the next four sweeps, 10^{-7} in the next four sweeps, and 0 in the remaining. We analyze the relative error in the ground state energy and the agreement in the prediction of the two-point spin and density correlation functions (see Eqs. 22 and 23).

Figure S2 (a) shows the relative (to DMRG) error in the SKQD ground-state energy estimation as a function of the subspace dimension on the SKQD eigenstate solver D . The Hartree-Fock (HF) and coupled cluster with single and double excitations (CCSD) errors are also shown for reference. The SKQD relative error decreases from values $\sim 10^{-5}$ to $\sim 10^{-6}$ as U increases from $U = 1$ to $U = 10$. This is a consequence of the increased ground-state sparsity for larger values of U .

Panel (b) of Fig. S2 compares the values of $\bar{C}_S(j)$ obtained from SKQD to those obtained with DMRG. The SKQD estimations are in excellent agreement with the DMRG values for most values of j , the distance between the impurity spin and the bath spin. There are small deviations for odd values of j at larger values of j , where the value of the correlation is negligible. Panel (c) in Fig. S2 compares the values of $\bar{C}_n(j)$ obtained from SKQD to those obtained with DMRG, for even values of j . The SKQD estimations are in excellent agreement with the DMRG values for all values of j .

The accuracy for the system size presented in this Appendix ($K = 29$), does not significantly differ from the accuracy on the larger system size ($K = 41$) shown in the main text. We conclude that the accuracy of SKQD does not significantly deteriorate with system size in the SIAM.

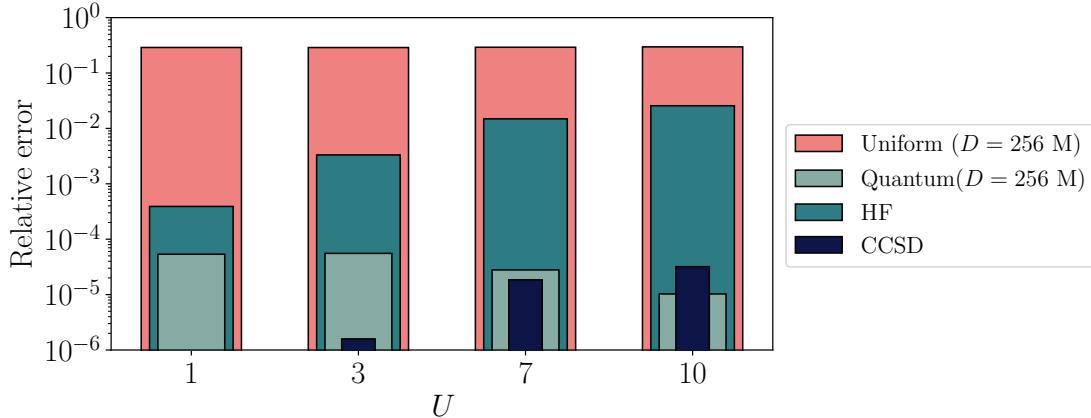


Figure S3. Signal in the quantum experiments. Relative error in the SKQD ground-state energy estimation for different values of U . The error is computed relative to the DMRG energy. SKQD is run both on samples coming from the quantum device and the uniform distribution. The HF and CCSD relative errors are included for reference.

VIII. SIGNAL IN THE QUANTUM EXPERIMENTS FOR THE SINGLE IMPURITY ANDERSON MODEL

Given the large circuit sizes of our experiments and the effect of noise, we investigate whether there is a useful signal coming out of the quantum circuits, comparing the outcome of SKQD (with configuration recovery) run on samples coming from the device and uniform random samples. The random samples are drawn from the uniform distribution in the space of bitstrings whose length is the same as the number of fermionic modes in the system.

This test is conducted on the SIAM with $K = 41$ bath sites and the same values of U as the ones shown in the main text: $U = 1, 3, 7, 10$. The subspace dimension chosen to project and diagonalize the Hamiltonian is $D = 2.56 \cdot 10^6$ electronic configurations, and the total number of sampled bitstrings is the same in both cases: $2.5 \cdot 10^6$.

Figure S3 shows the relative (to DMRG) error in the ground-state energy obtained from running SKQD on samples drawn from the device and samples drawn from the uniform distribution for different values of U . The relative error in the ground-state energy is orders of magnitude lower in the SKQD estimation run on samples coming from the quantum device. We therefore conclude that SKQD with configuration recovery is capable of extracting an useful signal from the device.

- [1] A Yu Kitaev. Quantum measurements and the Abelian stabilizer problem. *arXiv preprint quant-ph/9511026*, 1995.
- [2] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O’Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5(1):4213, 2014.
- [3] Dave Wecker, Matthew B Hastings, and Matthias Troyer. Progress towards practical quantum variational algorithms. *PRA*, 92(4):042303, 2015.
- [4] 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. *Nature Reviews Physics*, 3(9):625–644, 2021.
- [5] Martin Larocca, Supanut Thanansilp, Samson Wang, Kunal Sharma, Jacob Biamonte, Patrick J Coles, Lukasz Cincio, Jarrod R McClean, Zoë Holmes, and M Cerezo. A review of barren plateaus in variational quantum computing. *arXiv preprint arXiv:2405.00781*, 2024.
- [6] Jarrod R. McClean, Mollie E. Kimchi-Schwartz, Jonathan Carter, and Wibe A. de Jong. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. *Phys. Rev. A*, 95:042308, Apr 2017.

- [7] Robert M Parrish and Peter L McMahon. Quantum filter diagonalization: Quantum eigendecomposition without full quantum phase estimation. *arXiv preprint, arXiv:1909.08925*, 2019.
- [8] Mario Motta, Chong Sun, Adrian T. K. Tan, Matthew J. O'Rourke, Erika Ye, Austin J. Minnich, Fernando G. S. L. Brandão, and Garnet Kin-Lic Chan. Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution. *Nature Physics*, 16(2):205–210, 2020.
- [9] Katherine Klymko, Carlos Mejuto-Zaera, Stephen J. Cotton, Filip Wudarski, Miroslav Urbanek, Diptarka Hait, Martin Head-Gordon, K. Birgitta Whaley, Jonathan Moussa, Nathan Wiebe, Wibe A. de Jong, and Norm M. Tubman. Real-time evolution for ultracompact Hamiltonian eigenstates on quantum hardware. *PRX Quantum*, 3:020323, May 2022.
- [10] Ethan N. Epperly, Lin Lin, and Yuji Nakatsukasa. A theory of quantum subspace diagonalization. *SIAM Journal on Matrix Analysis and Applications*, 43(3):1263–1290, August 2022.
- [11] 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, July 2023.
- [12] Bo Yang, Nobuyuki Yoshioka, Hiroyuki Harada, Shigeo Hakkaku, Yuuki Tokunaga, Hideaki Hakoshima, Kaoru Yamamoto, and Suguru Endo. Dual-GSE: Resource-efficient generalized quantum subspace expansion. *arXiv preprint, arXiv:2309.14171*, 2023.
- [13] Ruyu Yang, Tianren Wang, Bing-Nan Lu, Ying Li, and Xiaosi Xu. Shadow-based quantum subspace algorithm for the nuclear shell model. *arXiv preprint, arXiv:2306.08885*, 2023.
- [14] Yasuhiro Ohkura, Suguru Endo, Takahiko Satoh, Rodney Van Meter, and Nobuyuki Yoshioka. Leveraging hardware-control imperfections for error mitigation via generalized quantum subspace. *arXiv preprint, arXiv:2303.07660*, 2023.
- [15] Keita Kanno, Masaya Kohda, Ryosuke Imai, Sho Koh, Kosuke Mitarai, Wataru Mizukami, and Yuya O Nakagawa. Quantum-selected configuration interaction: Classical diagonalization of Hamiltonians in subspaces selected by quantum computers. *arXiv preprint, arXiv:2302.11320*, 2023.
- [16] Javier Robledo-Moreno, Mario Motta, Holger Haas, Ali Javadi-Abhari, Petar Jurcevic, William Kirby, Simon Martiel, Kunal Sharma, Sandeep Sharma, Tomonori Shirakawa, Iskandar Sitzdikov, Rong-Yang Sun, Kevin J. Sung, Maika Takita, Minh C. Tran, Seiji Yunoki, and Antonio Mezzacapo. Chemistry beyond the scale of exact diagonalization on a quantum-centric supercomputer. *Science Advances*, 11(25):eau9991, 2025.
- [17] Nobuyuki Yoshioka, Mirko Amico, William Kirby, Petar Jurcevic, Arkopal Dutt, Bryce Fuller, Shelly Garion, Holger Haas, Ikko Hamamura, Alexander Ivrii, Ritajit Majumdar, Zlatko Minev, Mario Motta, Bibek Pokharel, Pedro Rivero, Kunal Sharma, Christopher J. Wood, Ali Javadi-Abhari, and Antonio Mezzacapo. Diagonalization of large many-body Hamiltonians on a quantum processor. *arXiv preprint, arXiv:2407.14431*, 2024.
- [18] Mario Motta, William Kirby, Ieva Liepuoniute, Kevin J Sung, Jeffrey Cohn, Antonio Mezzacapo, Katherine Klymko, Nam Nguyen, Nobuyuki Yoshioka, and Julia E Rice. Subspace methods for electronic structure simulations on quantum computers. *Electronic Structure*, 6(1):013001, 2024.
- [19] Oumarou Oumarou, Pauline J Ollitrault, Cristian L Cortes, Maximilian Scheurer, Robert M Parrish, and Christian Gogolin. Molecular properties from quantum krylov subspace diagonalization. *arXiv preprint arXiv:2501.05286*, 2025.
- [20] Youngseok Kim, Andrew Eddins, Sajant Anand, Ken Xuan Wei, Ewout Van Den Berg, Sami Rosenblatt, Hasan Nayfeh, Yantao Wu, Michael Zaletel, Kristan Temme, et al. Evidence for the utility of quantum computing before fault tolerance. *Nature*, 618(7965):500–505, 2023.
- [21] Kazuya Shinjo, Kazuhiro Seki, Tomonori Shirakawa, Rong-Yang Sun, and Seiji Yunoki. Unveiling clean two-dimensional discrete time quasicrystals on a digital quantum computer. *arXiv preprint arXiv:2403.16718*, 2024.
- [22] Danil Kaliakin, Akhil Shajan, Javier Robledo Moreno, Zhen Li, Abhishek Mitra, Mario Motta, Caleb Johnson, Abdullah Ash Saki, Susanta Das, Iskandar Sitzdikov, et al. Accurate quantum-centric simulations of supramolecular interactions. *arXiv preprint arXiv:2410.09209*, 2024.
- [23] Stefano Barison, Javier Robledo Moreno, and Mario Motta. Quantum-centric computation of molecular excited states with extended sample-based quantum diagonalization. *arXiv preprint arXiv:2411.00468*, 2024.
- [24] Ieva Liepuoniute, Kirstin D Doney, Javier Robledo-Moreno, Joshua A Job, Will S Friend, and Gavin O Jones. Quantum-centric study of methylene singlet and triplet states. *arXiv preprint arXiv:2411.04827*, 2024.
- [25] Akhil Shajan, Danil Kaliakin, Abhishek Mitra, Javier Robledo Moreno, Zhen Li, Mario Motta, Caleb Johnson, Abdullah Ash Saki, Susanta Das, Iskandar Sitzdikov, et al. Towards quantum-centric simulations of extended molecules: sample-based quantum diagonalization enhanced with density matrix embedding theory. *arXiv preprint arXiv:2411.09861*, 2024.
- [26] Yuri Alexeev, Maximilian Amsler, Paul Baity, Marco Antonio Barroca, Sanzio Bassini, Torey Battelle, Daan Camps, David Casanova, Frederic T Chong, Charles Chung, et al. Quantum-centric supercomputing for materials science: A perspective on challenges and future directions. *arXiv preprint arXiv:2312.09733*, 2023.
- [27] P. W. Anderson. Localized magnetic states in metals. *Phys. Rev.*, 124:41–53, Oct 1961.
- [28] Steven R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69:2863–2866, Nov 1992.
- [29] Steven R. White. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B*, 48:10345–10356,

- Oct 1993.
- [30] Steven R. White. Density matrix renormalization group algorithms with a single center site. *Phys. Rev. B*, 72:180403, Nov 2005.
- [31] Ang-Kun Wu, Benedikt Kloss, Wladislaw Krinitzin, Matthew T Fishman, JH Pixley, and EM Stoudenmire. Disentangling interacting systems with fermionic gaussian circuits: Application to quantum impurity models. *Physical Review B*, 111(3):035119, 2025.
- [32] Victor Barzykin and Ian Affleck. Screening cloud in the k -channel kondo model: Perturbative and large- k results. *Phys. Rev. B*, 57:432–448, Jan 1998.
- [33] Andreas Holzner, Ian P. McCulloch, Ulrich Schollwöck, Jan von Delft, and Fabian Heidrich-Meisner. Kondo screening cloud in the single-impurity anderson model: A density matrix renormalization group study. *Phys. Rev. B*, 80:205114, Nov 2009.
- [34] Javier Robledo Moreno, Jeffrey Cohn, Dries Sels, and Mario Motta. Enhancing the expressivity of variational neural, and hardware-efficient quantum states through orbital rotations. *arXiv preprint arXiv:2302.11588*, 2023.
- [35] Björn O Roos, Peter R Taylor, and Per EM Sigbahn. A complete active space scf method (casscf) using a density matrix formulated super-ci approach. 48(2):157–173, 1980.
- [36] Martin Head-Gordon and John A Pople. Optimization of wave function and geometry in the finite basis hartree-fock method. 92(11):3063–3069, 1988.
- [37] Hans-Joachim Werner and Peter J Knowles. A second order multiconfiguration scf procedure with optimum convergence. 82(11):5053–5063, 1985.
- [38] Jeppe Olsen. The casscf method: A perspective and commentary. 111(13):3267–3272, 2011.
- [39] Per Åke Malmqvist, Alistair Rendell, and Björn O Roos. The restricted active space self-consistent-field method, implemented with a split graph unitary group approach. 94(14):5477–5482, 1990.
- [40] Dominika Zgid and Marcel Nooijen. The density matrix renormalization group self-consistent field method: Orbital optimization with the density matrix renormalization group method in the active space. 128(14):144116, April 2008.
- [41] Debashree Ghosh, Johannes Hachmann, Takeshi Yanai, and Garnet Kin-Lic Chan. Orbital optimization in the density matrix renormalization group, with applications to polyenes and β -carotene. 128(14):144117, April 2008.
- [42] Sebastian Wouters and Dimitri Van Neck. The density matrix renormalization group for ab initio quantum chemistry. 68(9):272, September 2014.
- [43] Yingjin Ma, Stefan Knecht, Sebastian Keller, and Markus Reiher. Second-Order Self-Consistent-Field Density-Matrix Renormalization Group. 13(6):2533–2549, June 2017.
- [44] P. Jordan and E. Wigner. Über das paulische äquivalenzverbot. *Zeit. Phys.*, 47(9):631–651, Sep 1928.
- [45] Dian Wu, Riccardo Rossi, Filippo Vicentini, Nikita Astrakhantsev, Federico Becca, Xiaodong Cao, Juan Carrasquilla, Francesco Ferrari, Antoine Georges, Mohamed Hibat-Allah, Masatoshi Imada, Andreas M. Läuchli, Guglielmo Mazzola, Antonio Mezzacapo, Andrew Millis, Javier Robledo Moreno, Titus Neupert, Yusuke Nomura, Jannes Nys, Olivier Parcollet, Rico Pohle, Imelda Romero, Michael Schmid, J. Maxwell Silvester, Sandro Sorella, Luca F. Tocchio, Lei Wang, Steven R. White, Alexander Wietek, Qi Yang, Yiqi Yang, Shiwei Zhang, and Giuseppe Carleo. Variational benchmarks for quantum many-body problems. *Science*, 386(6719):296–301, 2024.
- [46] Kenji Sugisaki, Shu Kanno, Toshinari Itoko, Rei Sakuma, and Naoki Yamamoto. Hamiltonian simulation-based quantum-selected configuration interaction for large-scale electronic structure calculations with a quantum computer. *arXiv preprint arXiv:2412.07218*, 2024.
- [47] Mathias Mikkelsen and Yuya O Nakagawa. Quantum-selected configuration interaction with time-evolved state. *arXiv preprint arXiv:2412.13839*, 2024.
- [48] The ffsim developers. ffsim: Faster simulations of fermionic quantum circuits.
- [49] Caleb Johnson, Stefano Barison, Bryce Fuller, James R. Garrison, Jennifer R. Glick, Abdullah Ash Saki, Antonio Mezzacapo, Javier Robledo-Moreno, Max Rossmannek, Paul Schweigert, Iskandar Situdikov, and Kevin J. Sung. Qiskit addon: sample-based quantum diagonalization. <https://github.com/Qiskit/qiskit-addon-sqd>, 2024.
- [50] Ali Javadi-Abhari, Matthew Treinish, Kevin Krsulich, Christopher J. Wood, Jake Lishman, Julien Gacon, Simon Martiel, Paul D. Nation, Lev S. Bishop, Andrew W. Cross, Blake R. Johnson, and Jay M. Gambetta. Quantum computing with Qiskit, 2024.
- [51] Huanchen Zhai, Henrik R. Larsson, Seunghoon Lee, Zhi-Hao Cui, Tianyu Zhu, Chong Sun, Linqing Peng, Ruojing Peng, Ke Liao, Johannes Tölle, Junjie Yang, Shuo-xue Li, and Garnet Kin-Lic Chan. Block2: A comprehensive open source framework to develop and apply state-of-the-art dmrg algorithms in electronic structure and beyond. *The Journal of Chemical Physics*, 159(23):234801, 12 2023.
- [52] Qiming Sun, Timothy C Berkelbach, Nick S Blunt, George H Booth, Sheng Guo, Zhendong Li, Junzi Liu, James D McClain, Elvira R Sayfutyarova, Sandeep Sharma, et al. PySCF: the Python-based simulations of chemistry framework. *WIREs Comput. Mol. Sci.*, 8(1):e1340, 2018.
- [53] Qiming Sun, Xing Zhang, Samragni Banerjee, Peng Bao, Marc Barbry, Nick S Blunt, Nikolay A Bogdanov, George H Booth, Jia Chen, Zhi-Hao Cui, et al. Recent developments in the PySCF program package. *J. Chem. Phys.*, 153(2):024109, 2020.

- [54] William Kirby. Analysis of quantum Krylov algorithms with errors. *Quantum*, 8:1457, August 2024.
- [55] Sergey Bravyi, David Gosset, and Yinchen Liu. Classical simulation of peaked shallow quantum circuits. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 561–572, 2024.
- [56] Scott Aaronson and Yuxuan Zhang. On verifiable quantum advantage with peaked circuit sampling. *arXiv preprint, arXiv:2404.14493*, 2024.
- [57] László Borda. Kondo screening cloud in a one-dimensional wire: Numerical renormalization group study. *Phys. Rev. B*, 75:041307, Jan 2007.
- [58] Huanchen Zhai, Henrik R Larsson, Seunghoon Lee, Zhi-Hao Cui, Tianyu Zhu, Chong Sun, Linqing Peng, Ruojing Peng, Ke Liao, Johannes Tölle, et al. Block2: A comprehensive open source framework to develop and apply state-of-the-art dmrg algorithms in electronic structure and beyond. *The Journal of Chemical Physics*, 159(23), 2023.
- [59] Sandeep Sharma, Adam A Holmes, Guillaume Jeanmairet, Ali Alavi, and Cyrus J Umrigar. Semistochastic heat-bath configuration interaction method: Selected configuration interaction with semistochastic perturbation theory. *Journal of chemical theory and computation*, 13(4):1595–1604, 2017.
- [60] Dave Wecker, Matthew B. Hastings, Nathan Wiebe, Bryan K. Clark, Chetan Nayak, and Matthias Troyer. Solving strongly correlated electron models on a quantum computer. *Phys. Rev. A*, 92:062318, Dec 2015.
- [61] Ian D. Kivlichan, Jarrod McClean, Nathan Wiebe, Craig Gidney, Alán Aspuru-Guzik, Garnet Kin-Lic Chan, and Ryan Babbush. Quantum simulation of electronic structure with linear depth and connectivity. *Phys. Rev. Lett.*, 120:110501, Mar 2018.
- [62] Zhang Jiang, Kevin J. Sung, Kostyantyn Kechedzhi, Vadim N. Smelyanskiy, and Sergio Boixo. Quantum algorithms to simulate many-body physics of correlated fermions. *Phys. Rev. Applied*, 9:044036, Apr 2018.
- [63] Diederik P Kingma. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [64] James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy programs, 2018.
- [65] Pierre Pfeuty. The one-dimensional Ising model with a transverse field. *Annals of Physics*, 57(1):79–90, 1970.