Algorithmic Shadow Spectroscopy

Hans Hon Sang Chan,[1,*] Richard Meister,[1] Matthew L. Goh,[1,2] and Bálint Koczor[1,2,3,†]
[1] *Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, United Kingdom*
[2] *Mathematical Institute, University of Oxford, Woodstock Road, Oxford OX2 6GG, United Kingdom*
[3] *Quantum Motion, 9 Sterling Way, London N7 9HJ, United Kingdom*

(Received 14 January 2024; revised 9 June 2024; accepted 9 January 2025; published 17 March 2025)

We present shadow spectroscopy as a simulator-agnostic quantum algorithm for estimating energy gaps using very few circuit repetitions (shots) and no extra resources (ancilla qubits) beyond performing time evolution and measurements. The approach builds on the fundamental feature that every observable property of a quantum system must evolve according to the same harmonic components: we can reveal them by postprocessing classical shadows of time-evolved quantum states to extract a large number of time-periodic signals $N_o \propto 10^8$, whose frequencies correspond to Hamiltonian energy differences with precision limited as $\epsilon \propto 1/T$ for simulation time T. We provide strong analytical guarantees that (a) quantum resources scale as $O(\log N_o)$, while the classical computational complexity is linear $O(N_o)$, (b) the signal-to-noise ratio increases with the number of processed signals as $\propto \sqrt{N_o}$, and (c) spectral peak positions are immune to reasonable levels of noise. We demonstrate our approach on model spin systems and the excited-state conical intersection of molecular CH_2 and verify that our method is indeed intuitively easy to use in practice, robust against gate noise, amiable to a new type of algorithmic-error mitigation technique, and uses relatively few shots given a reasonable initial state is supplied—we demonstrate that even 10 shots per time step can be sufficient. Finally, we measured a high-quality, experimental shadow spectrum of a spin chain on readily available IBM quantum computers, achieving the same precision as in noise-free simulations without using any advanced error mitigation, and verified scalability in tensor-network simulations of up to 100-qubit systems.

DOI: 10.1103/PRXQuantum.6.010352

I. INTRODUCTION

Quantum simulation is possibly the most natural use of quantum computers, with numerous potential applications in, e.g., understanding quantum field theory [1], quantum gravity [2], and may help us develop novel drugs and materials [3–6]. The hope is that future quantum hardware developments will lead to scalable, universal quantum computers that can simulate the time evolution of any quantum system. Many such Hamiltonian simulation algorithms are known, such as product formula or Trotterization [7], quantum signal processing [8], qubitization [9], or linear combination of unitaries [10] to name a few. The quantum circuit depth of these algorithms is only linear in the total simulation time, in stark contrast with the exponential depth of generic classical simulation algorithms known today—thus promising a *quantum advantage* in modelling systems of practical interest.

Unfortunately, highly nontrivial measures for quantum error correction will be necessary to run these Hamiltonian simulation circuits, and the technological overhead associated with even the simplest error-correction schemes means they remain prohibitive in the current era of noisy intermediate-scale quantum (NISQ) technology—a single logical qubit, for example, may need to be embedded into thousands of physical, noisy qubits. In response, many researchers are proposing sophisticated error-mitigation techniques and new quantum simulation algorithms [11–19] that seek to demonstrate practical utility with near-term quantum computers. In this paradigm, a parametrized shallow quantum circuit is typically used, and its parameters are updated by a classical computer such that the evolution of the parametrized quantum state closely approximates the trajectory of the true quantum evolution [20,21]. Alternatively, analog quantum simulators built from, e.g., ultracold atoms can mimic the dynamics of model systems, including variants of the continually elusive Fermi-Hubbard model [22–24].

---

[*] Contact author: hans.chan@materials.ox.ac.uk
[†] Contact author: koczor@maths.ox.ac.uk

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

And yet, even the ability to approximate time evolution with lower quantum resources is itself not useful enough in practice, because quantum mechanics poses fundamental limitations on how efficiently we can extract observable properties from a quantum simulation. This constraint is exemplified by the "holy grail" of quantum simulation; the computation of a Hamiltonian's energy spectrum, a task that has direct experimental relevance in, e.g., predicting electronic properties of solids to photochemical processes [25,26], but its accurate classical treatment generally suffers from prohibitive scaling. Powerful quantum phase-estimation protocols, involving Hamiltonian simulation, for example, are designed for such a task, but require extra ancillary qubit registers controlling the already challenging time-evolution circuits. The probability of a successful phase-estimation measurement also imposes strict demands on the overlap between initial states and the target eigenstates [27]. In NISQ-friendly ansatz-based methods that avoid additional qubits like the variational quantum eigensolver (VQE), the measurement challenge instead manifests as the need for many circuit-measurement repeats (shots) for the classical optimization of the fixed quantum circuit ansatz.

In this work, we propose and develop a novel means of estimating the energy-spectrum differences, or gaps, in a Hamiltonian, which we call "shadow spectroscopy." It requires no additional quantum resources such as ancilla qubits, and can, in principle, work on any quantum simulator platform (digital or analog) as well as any Hamiltonian simulation method of choice. The initial state only needs to have sufficient support for the set of eigenstates which we want to know the energy gaps for, instead of a high overlap with one target eigenstate. Although this condition relaxes the demands on the initial state, its preparation can remain an exponential bottleneck, if the considered physical model does not allow for efficient initialization, e.g., from prior knowledge of the system. Despite the general hardness of initial state preparation, for many practically relevant tasks including chemistry, solid-state physics, etc., a broad range of efficient initial-state preparation techniques have been developed using sophisticated classical approximations, such as matrix-product states, Hartree-Fock and post-Hartree-Fock methods—and we demonstrate that quantum heuristics, such as VQE, can be used successfully (see details in Appendix A).

One key innovation is to harness powerful classical shadow techniques [28–33] which enables us to capture a large number of observables of a quantum state despite using orders of magnitude fewer measurement shots than, e.g., VQE. Our strategy is to estimate classical shadows of time-evolved quantum states, and analyze how that information—as captured by the shadows—evolves in time. Due to the fundamental laws of quantum mechanics, oscillations of all observable properties in a quantum system contain exactly the same fundamental frequencies that precisely correspond to energy differences in the Hamiltonian. These can then be extracted via classical postprocessing techniques–illustrated in Fig. 1.

We tested our technique experimentally and numerically in a broad range of applications and in various model systems of up to 100 qubits, finding it very easy to use, in practice, and provides a promising alternative for computing excitation energies. We also confirm that the present approach is robust to gate errors due to our effective reconstruction of a large number of observables from randomized measurements, and that positions of the ideal spectral peaks are not altered by incoherent errors; indeed, artefacts in the spectrum introduced by gate errors

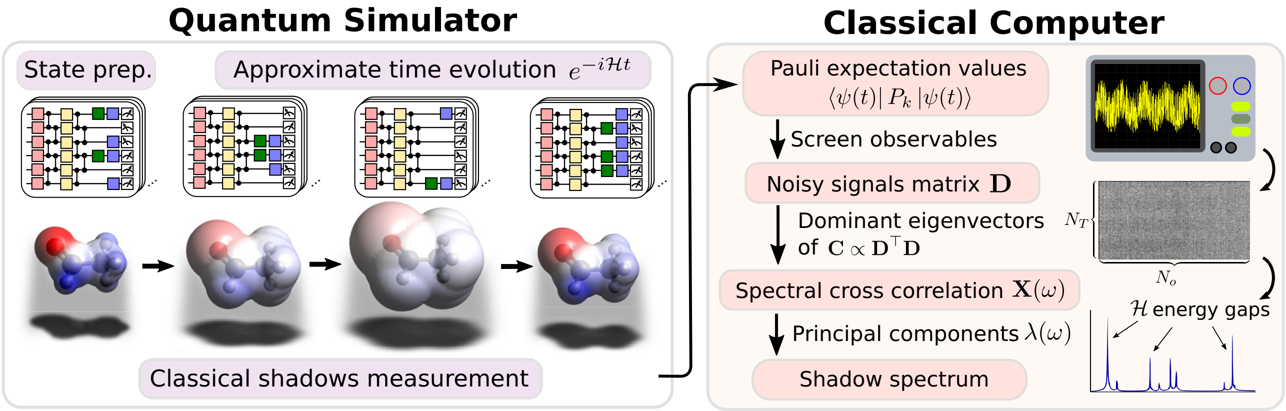

FIG. 1. Flowchart of algorithmic shadow spectroscopy: a plethora of well-established quantum simulation algorithms can be used to generate approximate time-evolved quantum states $|\psi(t_n)\rangle$—in this example, a mean-field many-electron wave function of the acetaldehyde molecule—from which we extract a series of classical shadows using randomized measurements. The shadows are stored classically as a list of binary numbers and measurement basis indexes. These are postprocessed using conventional HPC resources by computing a large noisy signal matrix $\mathbf{D}$, then the square matrix $\mathbf{C}$, and finally the shadow spectrum; these steps are elaborated upon in Fig. 2.

can even decrease as we increase the system size. Our approach is also amiable to a new type of quantum error-mitigation technique [34] that mitigates algorithmic errors incurred during the approximate time evolution. Another significant advantage is its robustness to finite circuit repetitions (shot noise) as illustrated in Fig. 2. In nearly all examples explored, only a few dozen measurements of the state were needed to resolve dominant peaks in the spectrum, whereas figures well beyond 10^6 shots are common for VQE. Specifically, we show that estimating and analyzing N_o signals only requires $\log N_o$ shots and a linear classical postprocessing time; having access to a large N_o, however, allows us to efficiently find "useful observables," to improve the spectral signal-to-noise ratio, and to suppress the effect of gate noise. In particular, postprocessing $N_o = 10^8$ observables can be performed in less than an hour with HPC resources and results in an improvement of the spectral intensity of $\propto 10^4$. Finally, our conclusions on the method's robustness against gate errors and shot noise are reinforced by the experimental demonstration of a shadow spectrum measured on freely accessible noisy IBM quantum computers—finding excellent agreement between numerical simulation and actual hardware results without using any advanced error-mitigation techniques.

The structure of this work is as follows. We first introduce the core ideas, and then discuss technical implementations in Sec. II. We then illustrate the utility of our approach on a broad set of examples in Sec. III. We finally discuss related questions in Sec. IV and conclude in Sec. V. While the method is robust and versatile, we discuss in detail its main limitations in Appendix A. We also report related techniques known in the literature in Appendix B.

II. SHADOW SPECTROSCOPY

In optical spectroscopy, an observable property of a quantum state, usually the expected value of some transition dipole moment (e.g., the magnetic dipole in nuclear magnetic resonance), is recorded as a function of time, and the Fourier-transformed signal reveals the desired transition energies. One could understand shadow spectroscopy as "substituting" measurement of time-dependent expectation values with classical shadows computed on a quantum device. As we demonstrate in numerical and hardware results, it offers great flexibility, and determining energy differences of low-lying single-electron excitation using only local Pauli shadows and classical postprocessing is straightforward. In this section, we present the main technical implementation details.

A. Preliminaries

Time evolution of any quantum state under Hamiltonian dynamics introduces periodic oscillations precisely determined by differences of eigenvalues in the Hamiltonian operator. The following statement summarizes the well-known property that expected values of observables oscillate with these frequencies under a closed Hamiltonian evolution [35,36].

Statement 1. Given any d-dimensional Hamiltonian $\mathcal{H}$ with eigenvectors $|\psi_k\rangle$, one can define any quantum state as $|\psi\rangle = \sum_{k=1}^{d} c_k\,|\psi_k\rangle$. Measuring the expected value of any observable O in the time-evolved quantum state $|\psi(t)\rangle := e^{-it\mathcal{H}}\,|\psi\rangle$ yields the signal $S(t)$ as

$$S(t) := \langle\psi(t)|O|\psi(t)\rangle = \sum_{k,l=1}^{d} I_{kl}\, e^{-it(E_l-E_k)}. \tag{1}$$

The above Fourier components have frequencies that correspond to differences of eigenvalues $E_l - E_k$ of the Hamiltonian. The intensity of the corresponding peak in the Fourier spectrum is $I_{kl} = c_k^* c_l \langle\psi_k|O|\psi_l\rangle$.

Proof. All d-dimensional Hamiltonians admit the spectral decomposition $\mathcal{H} = \sum_{k=1}^{d} E_k |\psi_k\rangle\langle\psi_k|$ with eigenvectors $|\psi_k\rangle$ spanning the full space. Thus, the evolution operator acts as

$$|\psi(t)\rangle = e^{-it\mathcal{H}}|\psi\rangle = \sum_{k=1}^{q} c_k e^{-itE_k}|\psi_k\rangle,$$

and the time-dependent $\langle\psi(t)|O|\psi(t)\rangle$ expected value is then obtained via a direct calculation. The generalization of this statement to density matrices as well as to infinite-dimensional but bounded Hamiltonians follows trivially. ■

An important consequence of the above statement is that time-dependent expected values contain the same harmonic components $e^{-it(E_l-E_k)}$ regardless of the observable measured. These frequencies precisely correspond to energy differences as $(E_l - E_k)$ in a Hamiltonian which—in principle—can be determined by naively picking a single observable O and measuring its corresponding time-dependent expectation value signal $S(t)$, then Fourier transforming $S(t)$.

In practice, a number of crucial factors determine the intensity of $S(t)$. First, it depends on the amplitudes of the state vector c_k; if the initial state is, e.g., a random state or a uniform superposition over all eigenstates, the Fourier spectrum may contain an exponentially large number of overlapping peaks and may thus prohibit us from learning anything from the spectrum. If we can instead initialize the quantum state such that only a handful of lower-energy eigenstates contribute, i.e., $|c_k| \gg |c_l|$ where all $k \in \mathcal{Q}$ and $l \notin \mathcal{Q}$, then we can isolate $\mathcal{O}(|\mathcal{Q}|^2)$ dominant peaks. This is the case, for example, when we can initialize in a superposition of only a few eigenstates as $|\psi\rangle = \sum_{k\in\mathcal{Q}} c_k\,|\psi_k\rangle$.

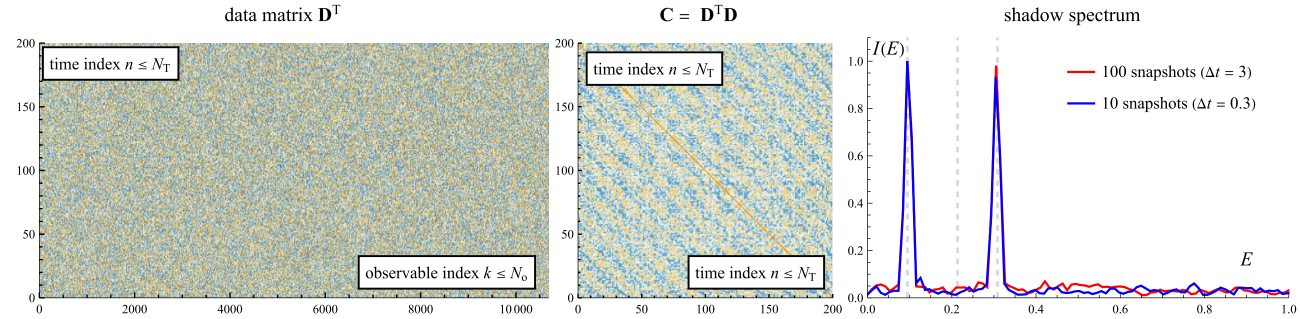

FIG. 2. A 14-qubit initial state $|\psi(0)\rangle \propto (1, \frac{1}{10}, \frac{1}{10}, 0, 0\dots)$ as a superposition of the three lowest-lying energy states of a spin problem similar to the one in Sec. III A 1. (Left) At each time increment ($N_T = 200$) we estimate every up to three-local Pauli string ($N_o = 10\,689$) from classical shadows of $N_s = 100$ snapshots. Overall 2000/10 689 signals are distinguishable from pure shot noise. (Middle) Time-periodic correlations are clearly visible in $\mathbf{C} \propto \mathbf{D}^T\mathbf{D}$. (Right) The shadow spectrum is obtained by effectively Fourier transforming the dominant eigenvectors of $\mathbf{C}$ (see details in main text). Peaks in the spectrum of the dominant eigenvectors of $\mathbf{C}$ reveal energy differences in the Hamiltonian as $E_0 - E_1$ and $E_0 - E_2$ (dashed lines are exact energies). A third peak does not appear as it is suppressed due to the initial state $\propto 0.1^2$, see text. Blue line: the signal-to-noise ratio is unchanged when reducing the number of snapshots to $N_s = 10$ and proportionally increasing the time steps $N_T = 2000$. The signal-to-noise ratio is increased as $\propto \sqrt{N_o}$ as we increase the number of observables N_o while the classical computational cost is linear $\mathcal{O}(N_o)$ and the required quantum resources (N_s) are only logarithmic $\mathcal{O}(\log N_o)$. The precision of energy-difference estimation exhibits ϵ^{-1} scaling as it is proportional to the total simulation time (linear in the used quantum resources).

In fact, we consider several known strategies in this work for preparing initial states such that $|c_1| \approx 1$ and $|c_k| \gg |c_l|$ for all $2 \le k \le q$ and $l > q$. Then we only have q dominant peaks of nonzero frequency from which we can learn the lowest q excitation energies as $E_2 - E_1, E_3 - E_1$, etc.

Second, the signal intensity depends on the matrix elements between eigenstates as $\langle\psi_k|O|\psi_l\rangle$, which can be arbitrarily small for any chosen observable O. One can identify problem-specific observables that yield dominant signals, such as excitation operators in quantum chemistry as we discuss in Sec. III B; in Sec. B 2 we also discuss how techniques from subspace expansion are similarly relevant [37]. However, even if we can guarantee that a certain observable yields a nonzero matrix element, in practice, it may still require a large number of measurements to sufficiently suppress shot noise, especially when a signal has lower intensity than the shot-noise background.

For this reason we employ classical shadows, which are designed to reconstruct a large set of observable expectation values and thus multiple $S_i(t)$. This allows us to search efficiently for local observables that yield intense signals. We discuss several other advantages of estimating many observables in Sec. IV, e.g., it can boost the signal-to-noise ratio, thereby overcoming limitations of resolving relatively low-intensity signals using only a logarithmic number of shots—and it also makes the approach noise robust.

B. Reconstructing many time-dependent signals using classical shadows

Classical shadow procedures [28] apply randomized measurements to several copies of an unknown quantum state, and enable us to estimate a very large number of properties with provable sample complexities. While essentially all variants of classical shadows are useful for our purposes [29–33], which we discuss below in Sec. IV, in this section we use specifically the near-term friendly variant that enables us to reconstruct expected values of local Pauli operators. Let us briefly recapitulate the main steps of the approach.

(1) We apply a random unitary U to rotate a copy of an unknown quantum state. For local Pauli strings, the unitaries are chosen randomly from single-qubit Clifford gates on each qubit and the procedure is thus equivalent to randomly selecting to measure in the X, Y, or Z bases—we measure each qubit to obtain N-bit measurement outcomes $|\hat{b}_i\rangle \in \{0,1\}^N$.

(2) We then generate the classical shadows by applying the inverse of the measurement channel $\mathcal{M}$, which can be done efficiently as the channel chosen is a distribution over Clifford circuits. The classical snapshots are generated as $\hat{\rho}_i = \mathcal{M}^{-1}(U^\dagger |\hat{b}_i\rangle \langle\hat{b}_i| U)$, the classical shadows are collections of these N_s snapshots $S(\rho; N_s) = [\hat{\rho}_1, \dots, \hat{\rho}_{N_s}]$.

(3) From these classical shadows we can construct K estimators of ρ from our N_s snapshots as $\hat{\rho}_{(k)} = 1/r \sum_{i=(k-1)r+1}^{kr} \hat{\rho}_i$ with $r = \lfloor N_s/K \rfloor$ and classically calculate estimators of the Pauli expectation values $\hat{o}_i(N_s, K) = \text{median}\{\text{Tr}(O_i\hat{\rho}_{(1)}), \dots, \text{Tr}(O_i\hat{\rho}_{(K)})\}$.

(4) The sample complexity of obtaining these estimators of N_o Pauli operators of locality q to error ϵ is $\mathcal{O}[3^q \log(N_o)/\epsilon^2]$.

In this context, the classical shadow approach is very NISQ friendly as it only requires applying random single-qubit rotations immediately prior to measurement in the standard basis. The approach also has a provable sample efficiency: recall that q-local Pauli strings are tensor products of single-qubit Pauli operators $P^{(q)} \in \{\mathrm{Id}, X, Y, Z\}^{\otimes N}$ such that $P^{(q)}$ acts nontrivially (with X, Y, or Z) to q qubits and trivially (with Id) on all other qubits. To reconstruct N_o at most q-local Pauli strings one needs to collect a number of samples $N_s = \mathcal{O}(\log N_o)$ that grows only logarithmically as we increase N_o. Furthermore, the classical computational resources are linear as $\mathcal{O}(N_o)$ and fast, optimized, and hardware agnostic code can, e.g., be found in the QuEST [38] family of quantum emulation software. By applying this procedure, we can determine N_o time-dependent signals with a similarly advantageous sample complexity.

Let us briefly summarize the sample complexity of our approach of estimating time-dependent signals using classical shadows.

Statement 2. Determining N_o-independent signals of N_T time increments as time-dependent expected values of at most q-local Pauli strings requires a number of circuit repetitions as

$$N_s \in \mathcal{O}[N_T 3^q \epsilon^{-2} \log(N_o)], \tag{2}$$

using the classical shadows technique.

Proof. Following Theorem 1 of Ref. [28], N_o Pauli strings O_i of locality q, can be estimated to precision parameters ϵ, δ via the number of batches $K = 2\log(2N_o/\delta)$ where the number of samples in the individual batches is $N_{\mathrm{batch}} = \frac{34}{\epsilon^2} \max_i \|O_i\|^2_{\mathrm{shadow}}$. This results in an overall number of samples $N_s = N_{\mathrm{batch}} K$ while the norm is given in Lemma 3 in Ref. [28] as $\|O_k\|^2_{\mathrm{shadow}} = 3^q$. As such, at each time increment we use this procedure once to estimate N_o Pauli strings of locality q, which has a sample complexity upper bounded by

$$\tfrac{68}{\epsilon^2} 3^q \log(2N_o/\delta). \tag{3}$$

The procedure is repeated for each time increment and is thus used overall N_T times from which the total sample complexity follows. ■

The above statement guarantees us that we can obtain a large number of signals with logarithmic efficiency. However each signal is still burdened with shot noise $S_i(t) + \mathcal{E}_i(t)$ such that the variance of the random fluctuations $|\mathcal{E}_i(t)| \leq \epsilon$ is globally bounded by the precision parameter ϵ inherited from classical shadows. One might feel tempted to average the signals as $\sum_i S_i(t)$. However, as we show, each signal is phase shifted by an amount that is specific to each observable as

$$S_i(t) = \mathrm{const} + \sum_{\substack{k,l=1 \\ k<l}}^{q} |I_{ikl}| \cos[t(E_l - E_k) + \phi_{ikl}]. \tag{4}$$

Here we have used that $I_{ikl} = |I_{ikl}| e^{-i\phi_{ikl}} = c_k^* c_l \langle k|O_i|l\rangle$. Since averaging over phase-shifted cosines may even cancel out the signal, and that individual phase shifts are not known in advance, our problem ultimately becomes that of finding the most intense common frequency components across a collection of noisy signals that share common harmonics.

C. Classical postprocessing

We first standardize the signals using the sample mean μ_k and the empirical standard deviation σ_k as

$$f_k(n) = \frac{\langle \psi(t_n)|P_k|\psi(t_n)\rangle - \mu_k}{\sigma_k}. \tag{5}$$

Above we have also introduced the discrete temporal index $n \leq N_T$ and the time variable $t_n = n\Delta t$. We first prescreen the data to filter out signals $f_k(n)$ that are not sufficiently different from statistical noise—an autocorrelation test, such as the Ljung-Box test [39,40], works very well in practice, requires linear computational time $\mathcal{O}(N_o)$ and can be completely parallelized requiring no communication between the classical processors, e.g., every CPU in a cluster receives a copy of the shadow data (a relatively small dataset) and processes it independently. We then define a data matrix $\mathbf{D} \in \mathbb{R}^{N_o \times N_T}$ that contains these signals $[\mathbf{D}]_{kn} = f_k(n)$ as row vectors and assume the column dimension is much larger than the row dimension $N_o \gg N_T$ due to our effective use of classical shadows as illustrated in Fig. 2.

We then perform a dimensionality reduction in linear time as $\mathcal{O}(N_o N_T^2)$. This involves computing a relatively small square matrix $\mathbf{C} \in \mathbb{R}^{N_T \times N_T}$ and finding its c dominant eigenvectors $v_1, v_2, \ldots v_c$ that maximize the average overlap with our experimentally estimated signals f_k. For example, the dominant eigenvector satisfies

$$v_1^T \mathbf{C} v_1 = \max_{\|v\|_2 = 1} \frac{1}{N_o} \sum_{k=1}^{N_o} |\langle f_k, v\rangle|^2. \tag{6}$$

Given c peaks in the spectrum, the matrix $\mathbf{C}$ has $2c$ dominant eigenvectors, which form the signal subspace (as in subspace methods in signal processing [41]), see further details in Appendix D 2. The shadow spectrum is effectively obtained by Fourier transforming the dominant eigenvectors $v_1, v_2, \ldots v_c$ that form the signal subspace. Specifically, we robustly estimate the spectral density among these vectors using further classical processing: we calculate the dominant singular value of a spectral cross-correlation matrix in time $\mathcal{O}(c^3 N_T)$ that is independent of N_o as we detail in Appendix D 1.

D. Robustness to shot noise

We build analytical models in Appendix E to understand how statistical shot noise due to finite circuit repetition, using N_s shots, is suppressed as we increase the number of observables: We find that the signal-to-noise ratio of a peak is proportional to $\propto N_T N_s \overline{I^2} \sqrt{N_o}$ where $\overline{I^2} := \frac{1}{N_o} \sum_{i=1}^{N_o} I_i^2$ is the average intensity in the group of N_o observables. Figure 2 (right) nicely illustrates that if we decrease the number of shots from 100 to 10 per time step but proportionally increase the number of time steps, the spectrum is qualitatively unchanged since the product $N_T N_s$ in the above expression for the signal-to-noise ratio is unchanged.

We can contrast this with a naive approach whereby one Pauli observable is picked randomly and we estimate its signal (i.e., randomly pick a single row of $\mathbf{D}$); by computing its spectrum one would obtain a peak intensity on average $N_T N_s \overline{I^2}$. Thus, shadow spectroscopy boosts the signal intensity by a factor $\sqrt{N_o}$. For example, we estimate that postprocessing $N_o = 10^8$ observables can be performed in less than an hour on a desktop PC and boosts the intensity by a factor 10^4 using only a logarithmic overhead in quantum resources.

We perform numerical simulation using both matrix-product-state (MPS) simulations (up to 100 qubits) and exact time evolution (up to 14 qubits) in Appendix J and confirm that the performance of our approach indeed improves as we increase the number of qubits. The reason is that an increasing system size gives rise to an increasing number of useful observable signals that our postprocessing approach can harness—hence in our MPS simulations we see an improved signal-to-noise ratio at a system of 100 qubits.

We now summarize the algorithm in Fig. 1 and state its computational complexity—refer to Appendix D for further details.

Statement 3. We first compute the data matrix $\mathbf{D} \in \mathbb{R}^{N_o \times N_T}$ of estimated signals with $N_o \gg N_T$ as defined via Eq. (5) and then compute the dominant eigenvectors $v_1, v_2, \ldots v_c$ of the square matrix $\mathbf{C} = \mathbf{D}^T \mathbf{D} / N_o$. We overall have a classical computational complexity that is linear in the number of observables $\mathcal{O}(N_T^2 N_o)$. The spectral density function $\lambda(\omega)$ at frequency ω is then obtained as the dominant singular value of the spectral cross-correlation matrix $\mathbf{X}(\omega)$, which we calculate from the dominant c eigenvectors $v_1, v_2, \ldots v_c$ of $\mathbf{C}$. This step has a complexity $\mathcal{O}(c^3 N_T)$.

E. Robustness to gate noise

It is crucial to consider how the ideal signal, as the time-dependent expected value of a Pauli string $S(t_n) := \langle \psi(t_n) | P_k | \psi(t_n) \rangle$, is altered when the time-evolved quantum states $|\psi(t_n)\rangle$ are prepared using a noisy quantum circuit. We present a simple but powerful proof in Theorem 1, which states that, under a broad class of typical gate-error models, the noisy spectrum $\mathcal{F}[S_k^{\text{noisy}}]$ decomposes into a sum of the ideal spectrum $\mathcal{F}[\eta S_k]$ and an additive $\mathcal{F}[(1-\eta) W_k]$ artefact. This artefact has a magnitude bounded by how close the noise model is to global depolarizing (white) noise and η is either a constant or a time-dependent exponential envelope depending on the simulation algorithm used.

Statement 4. The following are consequences of Theorem 1:

(1) **spectral peak centers are unchanged by incoherent noise**;

(2) as long as the circuit error rate is reasonable as $\xi \approx 1$ noise slightly shrinks the intensity of the ideal peaks but does not cause a broadening of their lineshape;

(3) additive artefacts might appear in the spectrum due to the Fourier transform $\mathcal{F}[(1-\eta) W_k]$, however, their intensity is generally bounded and the bounds decrease as we increase the system size (number of qubits);

(4) since the additive artefacts $\mathcal{F}[(1-\eta) W_k]$ are specific to each Pauli string P_k and assuming they are uncorrelated, i.e., their peaks appear at different frequencies in the spectrum, their effect is suppressed by a factor N_o in the shadow spectrum as we increase the number of observables.

Refer to Appendix C for further details. The above properties establish that shadow spectroscopy is generally expected to be highly robust to gate errors. Most importantly, a broad class of noise models are guaranteed to not change the peak centers in the shadows spectrum and thus allow one to estimate the frequency of a peak that would be produced by a fully error-free simulation circuit—as long as the circuit error rate is reasonable $\xi \approx 1$. We also argue in the Appendix that the approach is equally robust against errors that go beyond the assumed noise model—the most concrete evidence we give is our demonstration on current noisy quantum hardware in Sec. III A 1, which clearly confirms robustness. We additionally note the approach can likely be further improved by randomized compiling techniques [35] or by generalized Pauli twirling techniques [42–45] that convert local noise into probabilistic noise [11]. Furthermore, most error-mitigation techniques can also naturally be combined with classical shadows [11,46].

III. APPLICATIONS

Spectroscopy is one of the most employed experimental techniques for probing atomic and molecular physics, commonly used, e.g., to study the geometrical properties or the photochemical and electronic excited-state properties of matter. If a system's Hamiltonian can be efficiently represented on a quantum device, we can use shadow spectroscopy to discover its spectrum. Quantum technologies offer different platforms for simulating the time evolution of a quantum system and in the following we explore several applications targeting key evolutionary stages of quantum computing: We numerically simulate shadow spectroscopy on NISQ and early-fault-tolerant machines as well as on fully fault-tolerant quantum computers.

A. Shadow spectroscopy for early quantum advantage

1. NISQ—spin problems and local Hamiltonians

In the NISQ era, it is crucial to identify problems that are well suited to the limitations of NISQ hardware (such as qubit connectivity and relatively high gate error rates) yet still sufficiently complicated to yield genuine quantum advantage. The one-dimensional Heisenberg chain with random local magnetic fields and nearest-neighbor interactions as

$$\mathcal{H} = -J\sum_{j=1}^{N-1}\vec{\sigma}_j\cdot\vec{\sigma}_{j+1} + \sum_{j}^{N} h_j\sigma_j^z, \tag{7}$$

has been identified as a promising candidate for early quantum advantage [47,48]. Here $\vec{\sigma}_j = [\sigma_j^x, \sigma_j^y, \sigma_j^z]$ is a vector of Pauli x, y, and z matrices on the jth qubit, J is a coupling constant, and $h_j \in [-h, h]$ is sampled from a uniform distribution (for disorder strength h). This Hamiltonian is relevant in studies of many-body localization and self-thermalization [47,49], but its phase transitions remain poorly understood due to the difficulty of solving the model on a classical computer. Therefore, efficiently probing the spectral gap of this model with shadow spectroscopy on NISQ devices may help enhance understanding of key many-body quantum phenomena.

Motivated by the circuit-depth and gate-connectivity limitations of NISQ devices, we apply ansatz-based variational quantum simulation [20,21] for the dynamical simulation step of shadow spectroscopy. The state is approximated by an ansatz $|\psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})\,|\mathbf{0}\rangle$, where $\boldsymbol{\theta}$ is a vector of real classical parameters. Furthermore, we use a hardware-efficient ansatz circuit as $U(\boldsymbol{\theta})$ (see illustration in Fig. 9). As we detail in Appendix G 1, at each time step we estimate a small matrix A_{ij} and small vector C_i using the (simulated) quantum device. We then solve the corresponding linear system of equations using a classical computer to obtain an updated set of parameters $\boldsymbol{\theta}$ that best approximate the time-evolved state.

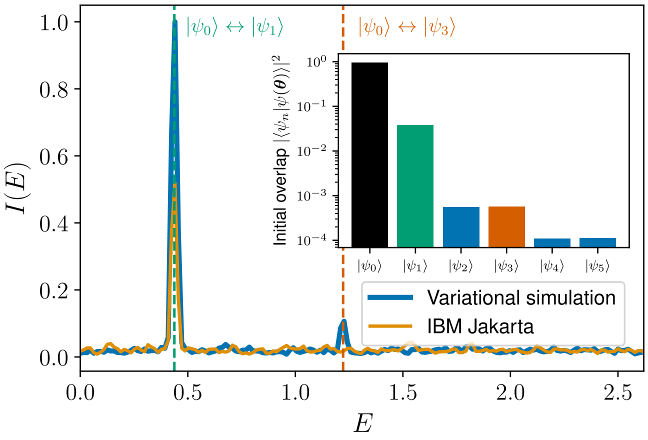

FIG. 3. Experimental shadow spectrum of a six-qubit disordered Heisenberg chain obtained using the IBM Jakarta QPU. Compared to noise-free simulation of the variational time evolution (blue), the noisy experimental spectrum (orange) correctly resolves exact energy gaps (dashed horizontal lines) between the ground state and first excited state—albeit with a lower peak intensity. The error in the gap estimate in the experiment is identical to that in ideal simulation as 2.8×10^{-3}. The peak corresponding to $E_3 - E_0$ has small support $< 10^{-3}$ and has low intensity due to the approximate time-evolution algorithm used. Inset: overlaps $|\langle\psi_n|\psi(\boldsymbol{\theta})\rangle|^2$ between the initial variational state $|\psi(\boldsymbol{\theta})\rangle$ and eigenstates $|\psi_n\rangle$.

We first apply variational imaginary-time evolution [50–52] to a set of random ansatz parameters and intentionally terminate the algorithm before it could fully reach the ground state to obtain a state that overlaps with the low-lying excited states. The inset in Fig. 3 shows the distribution of eigenstates. The advantage of near-converged initial states is that a majority overlap with the ground state guarantees resolvable signals that correspond to energy differences between the ground state and between the excited states (but not among the excited states). In principle, one could even assign peaks to specific low-lying excitations by repeating shadow spectroscopy on initial states of varying levels of convergence and comparing peak heights.

In Fig. 3 (blue line) we numerically simulate variational time-evolution circuits in 1000 time steps and perform shadow spectroscopy by estimating up to three-local Pauli operators (three batches of $N_s = 50$ snapshots used in each step). Indeed, the peaks corresponding to low-lying excitations are accurately resolved in a highly NISQ-friendly manner using shallow quantum circuits. Also note that expected values of only three-local Pauli strings were used: prior results suggest that local operators indeed yield dominant signal intensities, e.g., in translation invariant, gapped Hamiltonians of spin systems the intensity $|\langle\psi_k|O_i|\psi_l\rangle|$ is an exponentially decreasing function of the locality of O_i [53].

Next we provide evidence for the ability of our approach to extract accurate gap information from extant NISQ quantum hardware. Taking our classically precomputed time-evolution circuits, we measure classical shadows of 1000 time-evolved states on the seven-qubit IBM Jakarta quantum computer using the same budget of snapshots as above. The experimental shadow spectrum in Fig. 3 (orange line) clearly resolves the main peak and demonstrably achieves the same precision in predicting the energy gap as the pure-state simulator without using any error-mitigation techniques—this is thanks our quite general theoretical guarantees that only the peak intensity is shrunk by noise, which we analyze in the next section. The second, faint peak in simulation is not resolved in the experimental spectrum; but, as we detail in Appendix G 1, this peak is actually already suppressed by algorithmic errors from the approximate, variational time evolution, and here it is likely suppressed to the level of shot noise. We additionally report QASM simulations using the noise model of the IBM Jakarta chip in Appendix H finding the hardware experiment was actually closer to the ideal noiseless simulation than noisy simulation—this could be a result of out-of-date and inaccurate noise models for the quantum chip.

Moving forward, we expect that shadow spectroscopy can be a major enabler for early quantum advantage for the following reasons. First, variational circuits are substantially shallower and scale much better than, e.g., Trotter circuits; because the quantum state evolution is stored entirely in the classical parameters $\boldsymbol{\theta}$, the circuit depth is independent of T as opposed to linear in T for Trotter, allowing us, in principle, to increase the spectral resolution without increasing circuit depth. Secondly, although we note that optimization of the dynamical simulation parameters will likely still require a substantial measurement overhead, we require a significantly lower number of circuit repetitions than what is generally required for state-of-the-art VQE [17–19] where figures beyond 10^6 shots are quite typical. Thirdly, several behaviors of variational quantum eigensolvers that are typically considered to be deficiencies are actually advantageous for shadow spectroscopy. For example, one does not require an ansatz circuit that is sufficiently deep to precisely express the ground state. Instead, we need only be able to express a superposition of the low-lying eigenstates and approximately simulate time evolution using well-established variational techniques. Also, it is well documented that VQE methods typically get “stuck” in one of the exponentially many local minima around the ground state [14,54,55]. In contrast, an initial state that is only partially converged can be ideal for shadow spectroscopy, since substantial gradients for optimizing dynamics can be guaranteed [56] even when VQE is trapped in a false minimum (though it remains an active question when this corresponds to a good solution to the dynamics). Finally, although variational simulation unavoidably does not follow the exact time evolution and can thus change the eigenstate composition of the variational state, the appearance of additional spectral peaks nonetheless correspond to physical, albeit unintended, excitations of the system—which we observe in numerical simulations. We repeat simulations of spin models for up to 100 qubits in Appendix J using tensor-network techniques and confirm that an increasing system size improves the performance of the present approach.

2. Early fault tolerance—Fermi-Hubbard model

We now illustrate the utility of classical shadow spectroscopy in the next key evolutionary stage of quantum computers as early fault-tolerant devices. These machines will be able to implement significantly deeper circuits than in the NISQ-era enabling simple Hamiltonian simulation algorithms, such as Trotterization. However, due to the excessive resources required for, e.g., implementing T gates, gate operations will still incur a certain level of noise-limiting circuit depth thus limiting evolution times and the precision of the simulation (algorithmic errors).

We analyze in detail the resilience of shadow spectroscopy against both algorithmic and gate errors in another model of great potential for early quantum advantage. We consider the Fermi-Hubbard model as

$$H = -t \sum_{\langle i,j \rangle,\sigma} \left(c^\dagger_{i,\sigma} c_{j,\sigma} + c^\dagger_{j,\sigma} c_{i,\sigma} \right) + U \sum_i c^\dagger_{i,\uparrow} c_{i,\uparrow} c^\dagger_{i,\downarrow} c_{i,\downarrow},$$

where i and j number the lattice sites, $\langle \cdot, \cdot \rangle$ means pairs of neighboring sites, $\sigma \in \{\uparrow, \downarrow\}$ is the spin of the fermion, and $c^{(\dagger)}_{i,\sigma}$ are fermionic annihilation (creation) operators at site i with spin σ. We obtain the Hamiltonian of qubit Pauli operators $H = \sum_{\ell=1}^{L} H_\ell$ through the Jordan-Wigner (JW) transformation assuming a chain without periodic boundary conditions with parameters $t = 1$ and $U = 2$ [57]. We consider a small example system of 12 qubits (3×2 sites) and initialize our time evolution in an exact, equal superposition of the ground and first excited states as $|\psi(0)\rangle \propto |\psi_0\rangle + |\psi_1\rangle$. This guarantees a single peak in the shadow spectrum at the spectral gap $\Delta := E_0 - E_1$ enabling us to clearly analyze the robustness of shadow spectroscopy against increasing levels of errors.

3. Effects of algorithmic errors

We first examine how algorithmic errors alone influence the shadow spectrum. We assume that due to the limited resources (number of qubits and non-negligible gate errors) we need to resort to a first-order Trotterized time-evolution algorithm with limited circuit depth. Thus, in order to implement a sufficiently long simulation

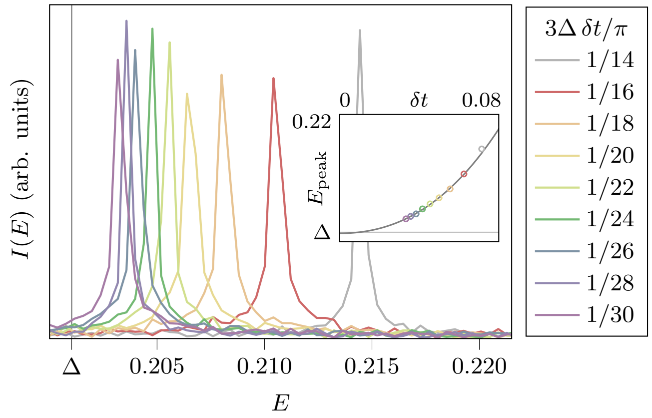

FIG. 4. Spectra obtained using Lie-Trotter-Suzuki time evolution for a Fermi-Hubbard model with exact spectral gap Δ. Spectra obtained for increasing time-step sizes δt—chosen as whole-numbered fractions of $\pi/3\Delta$—show peaks quite far away from Δ due to significant algorithmic errors. Classical shadows of only 50 snapshots were used to estimate expectation values of all up to three-local Pauli observables. Inset: positions of the peaks as a function of δt used; fitting a cubic polynomial (to all but the largest step size) and extrapolating to $\delta t \to 0$ allows us to estimate the exact gap Δ to high precision.

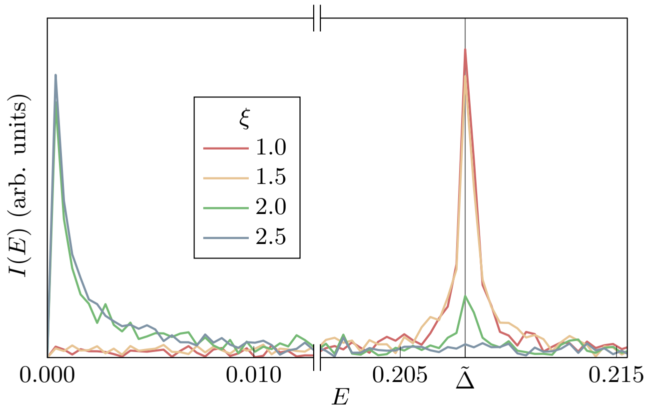

FIG. 5. Interesting regions (via interruption of the x axis) of the spectra obtained for the Hubbard model from simulations including gate noise for an increasing number of errors $\xi = \lambda N_{\mathrm{gates}}$ per a full circuit—classical shadows of 50 snapshots were used at every time step. The position of the peak without noise using the same Trotter step size as for the noisy simulations is marked as $\tilde{\Delta}$. The approach is very robust to errors as guaranteed via Statement 4: the peak position is unaffected as we increase ξ but the peak disappears as the buildup of errors escalates for large ξ.

$T = 1000\pi/\Delta$, we choose a relatively large δt, which, however, introduces a significant algorithmic error. Figure 4 shows the reconstructed spectra for different values of δt and confirms that all peaks are quite far from the exact energy gap due to the algorithmic error.

In Fig. 4 (inset) we plot the positions of the very accurately resolved peaks as a function of the step size δt. By fitting a cubic polynomial we extrapolate to $\delta t \to 0$ obtaining an excellent approximation of the spectral gap as $\Delta_{\mathrm{est}} = 0.2009 \pm 0.0003$, within error margin of the real gap $\Delta = 0.2010$. Rigorous theoretical guarantees on the efficacy of extrapolating Trotter errors are available in Ref. [58]. The present extrapolation approach is strictly more powerful than prior algorithmic error-mitigation techniques, including ones for extracting expected values of observables [34], as it allows an error-mitigated estimate of the spectral gap to be directly accessed without using, e.g., phase-estimation algorithms [58].

Furthermore, we expect that shadow spectroscopy is amiable to extrapolation-based error-mitigation techniques for a range of other time-evolution techniques whereby one has control of certain hyperparameters (such as δt in the present case) that influence the quality of the simulation, such as in qDRIFT [59] in stochastic evolution [60] and beyond. While in the present example we used a simple initial state to facilitate the analysis, care must be taken when multiple lines are present in the spectrum to accurately monitor changes between runs with different time-step sizes.

4. Effects of gate noise

In addition to algorithmic errors, we also investigate the influence of hardware noise. In particular, in an early fault-tolerant device we assume the dominant source of error is due to the imperfect implementation of T gates and describe an error model in Appendix G 2. In Fig. 5 we confirm that shadow spectroscopy is indeed immune to reasonable levels of shot noise as the relevant peaks at $E \approx 0.2$ are very pronounced and centered exactly where the noise-free peaks (dashed horizontal lines)

In particular, Fig. 5 shows three distinct cases around the error rate at which the result becomes unusable. At a circuit error rate of $\xi := \lambda N_{\mathrm{gates}} \approx 1.0$ and 1.5, the spectrum shows only one peak, the same as for the noise-free simulation. At $\xi \approx 2$ the peak drastically diminishes in height, and a new component close to $E \approx 0$ emerges. At $\xi = 2.5$ finally, the signal of the system at $E \approx 0.2$ disappears, and all the weight apart from the random noise has shifted over to $E \approx 0$. Indeed, this is an expected limitation and applies collectively to error-mitigation techniques as well: due the exponentially decreasing fidelity $F \approx e^{-\xi}$ the buildup of error escalates rapidly with increasing ξ and makes it impractical to extract useful information [12,61]. On the other hand, while in NISQ applications one needs to resort to error-mitigation techniques to extract useful information, the present approach is demonstrably robust against gate noise as the peak position is unaffected and only its intensity is affected by gate noise.

B. Shadow spectroscopy for mature quantum advantage

1. Fully fault-tolerant quantum chemistry

Finally, we explore the fully fault-tolerant quantum computing regime, where algorithmic errors can be suppressed exponentially via sophisticated post-Trotter time-evolution techniques: we describe computation of a molecule's low-lying electronic transitions using shadow spectroscopy. Typically the electronic structure Hamiltonian is specified in a second-quantized form as

$$\mathcal{H} = \sum_{pq} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{pqrs} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s, \tag{8}$$

and mapped onto multiqubit Pauli operators using the JW encoding. Mean-field or Hartree-Fock solutions (Slater determinants) can be computed classically efficiently and correspond to computational basis states.

While we discuss details in Appendix F, let us briefly summarize why quantum chemistry is highly amiable to shadow spectroscopy using local Pauli operators. Under the JW transform, excitation operators map to nonlocal Pauli strings, e.g., the single excitation $a_p^\dagger a_q \mapsto iA_p^\dagger A_q \prod_{k=p+1}^{q-1} Z_k$, where $A = (X + iY)/2$ is the qubit lowering operator. In spite of this, we find that in practice the signal intensities of observables like two-local Pauli strings, e.g., $X_p X_q$, $X_p Y_q$, or indeed any variant where we additionally append one or more single-qubit Z operators, yield dominant signals. The reason is that the intensity of these local observables is determined directly by the single-excitation coefficients in the configuration interaction expansion of the time-evolved states, which is the same as if the observable $a_p^\dagger a_q$ was measured instead.

2. Methylene spectra discovery

We now explore a specific quantum chemistry example: The bent methylene molecule CH_2 in the minimal STO-3G basis (14 qubits). Because of the challenges in isolating this molecular biradical experimentally, methylene was one of the first instances where classical computational quantum chemistry calculations predicted the properties of a molecule before it was then verified by photochemistry and spectroscopy experimentally [63–66]. It exhibits the following atypical features: First, its ground state (at equilibrium geometry) has an open-shelled triplet electron configuration instead of the usual closed-shell singlet. Second, the singlet and triplet states cross at a certain geometry, called a conical intersection in the excited-state chemistry literature, and the system becomes degenerate. We probe these features [67–69] by estimating shadow spectra at five evenly spaced geometries around the conical intersection, see Figs. 6(a)–6(e).

Recall that in the initial state we need only to have a reasonable support for eigenstates indexed k and l to observe the energy gap $E_k - E_l$: in Fig. 6 (top) we use an initial state as a superposition of the two basis states (Slater determinants) that correspond to the lowest-lying singlet and triplet states as

$$|\psi(0)\rangle = \frac{\sqrt{3}}{2} |11111111000000\rangle + \frac{1}{2} |11111110100000\rangle, \tag{9}$$

which can be prepared straightforwardly [70]. Here we time evolve these states exactly in $N_T = 500$ evenly spaced time intervals with $\Delta t = 10$, and at each time increment determine expected values of all three-local Pauli strings from 50 snapshots of classical shadows.

We plot the resulting shadow spectra for each geometry in Fig. 6 (top spectra). As expected, we find a dominant peak in each spectrum and their positions agree with the exact singlet-triplet gap (energy difference between the S_0 and T_1 eigenstates) to within error ± 0.4 mE$_h$ (0.25 kcal/mol), as predicted from the scaling $\propto \frac{2\pi}{N_T \Delta t}$—as singlet-triplet gaps are typically on the order of few kcal/mol, the low error of shadow spectroscopy is very promising. While physical optical spectroscopy only measures a single observable (e.g., electric dipole moment) where "spin-forbidden" transitions can prohibit resolution of the $S_0 \leftrightarrow T_1$ transition, shadow spectroscopy allows us to probe many different observables, some of which indeed lead to intense peaks as long as the initial state has support for them. At geometry (c), despite the near degenerate singlet-triplet intersection, the tiny expected frequency is still observed, confirming that indeed shadow spectroscopy is expected to be highly accurate when used in combination with sophisticated post-Trotter Hamiltonian simulation, even in the strong correlation regime where multiple eigenstates are often near degenerate.

We demonstrate another valuable feature as the flexibility to the initial state used: we initialize the time evolution in another superposition of basis states that approximate the two lowest-energy singlet states S_0 and S_1 as

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}} |11111111000000\rangle + \frac{1}{\sqrt{2}} |11111110010000\rangle. \tag{10}$$

In Fig. 6 (bottom spectra), we observe the same peak corresponding to the $T_1 \leftrightarrow S_0$ transition, suggesting the Slater determinants in Eq. (10) have sufficient fidelity with respect to the exact T_1 and S_0 eigenstates. We also observe new higher-energy peaks: their positions match the exact transition energies of $S_0 \leftrightarrow S_1$ and $T_1 \leftrightarrow S_1$ states within error ± 0.3 mE$_h$ (0.18 kcal/mol).

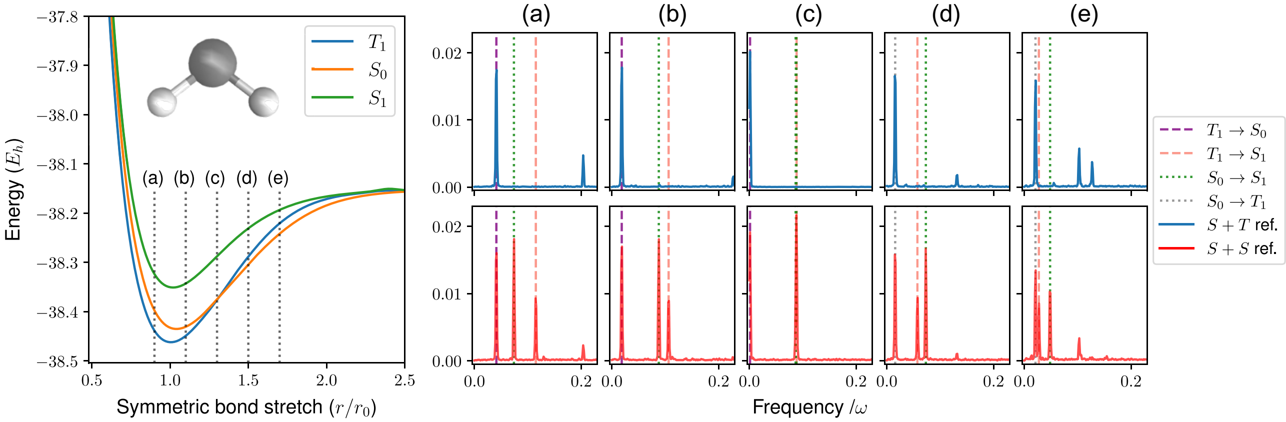

FIG. 6. Discovering the electronic spectrum of methylene with (classically emulated) shadow spectroscopy. We fix the bond angle to 101.89° and choose five symmetric C—H bond geometries between 0.5 and 3 times the equilibrium bond length of $r_0 = 1.1089$ Å [62]. The top row of spectra were generated using an initial state, which was a superposition of the lowest energy singlet and triplet determinants Eq. (9), and the bottom row of spectra a superposition of the lowest and first excited singlet determinants, Eq. (10). The dotted lines are the exact excitation energies from the singlet ground state, purple dashed lines correspond to those from the triplet ground state. From classical shadows of 50 snapshots, all up to three-local Pauli strings were determined with $N_o = 10689$, out of which 2000 signals passed a statistical test (they contain a deterministic signal with p values in the range $10^{-20} \le p \le 0.06$). Note that a number of high-energy peaks are observed too—assignment of these transitions can easily be achieved in practice using techniques detailed in the main text.

We conclude by noting that the low-lying excited-state spectrum of a more complex molecule will be more challenging to decipher. While in such complex scenarios, phase-estimation protocols are equally challenging (e.g., due to small overlaps with many eigenstates), the present approach may be helpful in reducing complexity, as different combinations of input states and Pauli strings measured only give rise to specific peaks. A reasonable strategy might be, as we have demonstrated in Fig. 6, to trial a few initial Slater determinants, linear combinations of them, different choice of Pauli operator observables, and multiple geometries. The difference between these scenarios should let the user assign peaks to energy transitions by elimination and obtain an excited-state landscape.

Another strategy is to make use of additional system knowledge: one could envisage a hybrid quantum-classical learning cycle that uses cheap, e.g., density-functional or perturbation-theory calculations to provide initial-state guesses, and from the shadow-spectroscopy results update the initial reference or geometry until a confident peak assignment is provided. One might also consider suppressing high-frequency satellite peaks by reducing support for high-energy states such as in the case of imaginary time evolution or variational quantum eigensolvers as we demonstrated in Sec. III A 1. If one instead desires, e.g., ionization energies or electron affinities, the present approach can also, in principle, be straightforwardly applied by initializing in a superposition of Slater determinants with different electron numbers (different Hamming weights). Further explorations in any of these directions may prove fruitful and offer a new pathway to quantum computing energy gaps in molecules without ever preparing eigenstates in a quantum computer and evaluating static electronic energies.

IV. FURTHER EXTENSIONS AND APPLICATIONS

Our postprocessing approach is the combination of relatively simple signal-processing techniques but they demonstrably work well in practice and guarantee a linear classical computational complexity. We note that we could use more advanced signal-processing techniques and the literature on this topic is rich [71–73]. We could also use compressed sensing to exploit sparsity in the shadow spectrum thereby reducing the overall number of time-point samplings (the number of time steps) [74,75]. Furthermore, while we primarily focused on well-established algorithmic time-evolution simulation, the presented techniques naturally extend to time evolutions performed with more advanced methods that correct the algorithmic errors incurred at the unitary level, as well as on analog quantum simulators combined with recent analog classical shadow techniques [76]

Beyond classical shadows, we considered problem-specific operator pools and constructed them explicitly for the case of quantum chemistry. This lends itself to a number of straightforward extensions. First, we could similarly construct problem-specific operator pools of nonlocal Pauli strings and use derandomization protocols [29] or shallow Clifford circuits [32,33] to estimate expected values efficiently. Second, the operator pools we constructed for fermionic problems can be grouped into large mutually commuting groups and we could measure these using well-established techniques [77–80]. Third, we could also construct an operator pool of fermionic operators and use fermionic shadows [30,31] to estimate expected values.

A number of further applications and generalizations are apparent. First, we can apply the present approach to an initial state that is relatively close to the ground state, e.g., a mean-field approximation for fermionic systems. By inspecting the shadow spectrum, one can semiquantitatively determine how many and which excited states contribute with dominant weights.

Second, we could also use the present approach to diagnose gaps in adiabatic quantum computing: as in adiabatic evolution, the quantum state closely follows the ground state, the shadow spectrum has peaks only at excitation energies between the ground and excited states. One can thus uniquely identify the spectral gap and determine whether it is sufficiently different from zero—effectively verifying the validity of the adiabatic theorem.

Third, while in the present work we primarily focused on determining energy gaps, it is straightforward to extend our approach to the direct estimation of energy levels: one usually proceeds by embedding the Hamiltonian into a larger system by adding an additional ancilla qubit. For example, in phase estimation one applies a controlled evolution while in NISQ-friendly phase estimation the Hamiltonian is appended with additional Pauli-Z operators [81]—finding operators that anticommute with the Hamiltonian are also equally effective [36]. These approaches, however, come with increased quantum resources and may require bespoke architectures.

V. DISCUSSION AND CONCLUSION

In this work, we developed a general technique that can be used to extract energy differences between discrete energy levels of a quantum system: information that is of key importance for many applications in chemistry, materials science, and beyond. Our approach can be viewed as the quantum digitization of spectroscopy: estimating observable properties of a time-evolved quantum system from its classical shadows allows us to learn fundamental frequencies in the time evolution as intense peaks in a shadow spectrum—these peaks precisely correspond to energy differences in the system's Hamiltonian. The primary limitation of the approach is that a sufficiently good initial state must be prepared that approximates the relevant eigenstates—this is unavoidably hard in general but is a common requirement for most classical and quantum computational methods, which we detail in Appendix A.

Classical shadows allow us to estimate a large number ($N_o \approx 10^8$ is realistic) of time-dependent observables, which is beneficial for at least three reasons. First, we can efficiently search for observables of high intensity, i.e., we estimate all q-local Pauli strings and select only the observables that give rise to intense signals using statistical correlation tests. Second, the effect of gate noise is potentially suppressed as $\propto N_o^{-1}$. Third, the signal-to-noise ratio (shot noise) in the spectrum is increased as $\propto \sqrt{N_o}$. While the classical computational complexity of our approach is linear and the required quantum resources are only logarithmic in N_o, we verified in tensor-network simulations for up to 100 qubits that indeed our approach is scalable and its performance improves for an increasing number of qubits. We also prove that the approach is immune to reasonable levels of gate errors as the ideal spectral peak positions are not altered by incoherent noise—which we verified in real quantum hardware demonstrations.

Furthermore, we provide a broad range of practically motivated demonstrative examples that illustrate how our approach indeed (a) requires no additional quantum resources beyond the ability to simulate time evolution and to extract classical snapshots (randomized measurements); (b) is highly robust to gate noise and shot noise, (c) admits a ϵ^{-1}-limited precision; (d) is amiable to a simple yet effective algorithmic error-mitigation technique. Shadow spectroscopy is also very versatile as it can be used for a large variety of quantum systems and Hamiltonians, and in principle with any quantum simulation platform. These include analog quantum simulators, NISQ machines, and early fault-tolerant quantum computers. For example, ion traps [82] can simulate time evolution of nontrivial spin models in an analog, non-gate-based fashion and can perform single-site Pauli measurements, making them an ideal platform for shadow spectroscopy.

The amount of absolute quantum resources required for the present approach can be very low; in both noisy hardware experiments and numerical simulations a few dozen circuit repetitions per time step were sufficient to measure shadow spectra. Thus we expect the present approach requires a moderately increased number of circuit repetitions compared to even mature phase-estimation protocols, which is, however, a worthwhile trade-off due to the above noted advantages—making the present approach a highly competitive alternative. Furthermore, shadow spectroscopy uses many orders of magnitude fewer shots than typical NISQ applications where figures well beyond 10^6 shots for a single eigenvalue estimation are very common—of course simulating time evolution with variational techniques may still require high repetition as in Sec. III A 1.

Shadow spectroscopy has numerous natural applications ranging from high-energy physics to materials science to modeling molecular systems. In fact, we identify quantum chemistry as a sweet spot for our approach: popular VQE algorithms typically suffer from grave performance deficiencies for quantum chemistry Hamiltonians such as astronomical sampling (circuit repetition) costs, barren plateaus, and relatively deep circuit depths required—extracting the spectrum additionally requires computing excited states, which exacerbates these drawbacks. In contrast, we prove analytically and demonstrate experimentally that local Pauli strings give rise to intense signals from which we can estimate energy differences as long as one can prepare good initial states and time evolve them; furthermore, the efficient fermionic shadows-measurement routine is very suitable for shadow spectroscopy for molecular systems [30,31].

Shadow spectroscopy may also be useful in entanglement spectroscopy [82,83], where the transition energies in a learned "entanglement Hamiltonian" correspond to the entanglement spectrum of a subsystem. Our method may provide a way of observing that spectrum experimentally with many fewer measurements.

We believe the present approach can be a major enabler for practical quantum advantage both on NISQ devices as well as on early fault-tolerant quantum computers due to its versatility, efficiency, and practicality. Furthermore, the present approach motivates deeper questions: can one find more direct ways to analyze time dependence from classical shadows circumventing the intermediate step of estimating expected values of Pauli operators? In the present work, we resorted to local Pauli strings and found that in all examples up to three-local Pauli strings gave rise to sufficiently informative signals about low-lying energy transitions in practice (which we verified up to 100 qubits).

ACKNOWLEDGMENTS

The authors thank Simon Benjamin for his support and help throughout all stages of this work. The authors thank Tyson Jones for developing efficient simulation code for classical shadows in QuESTlink. The authors are grateful to Zsuzsanna Koczor-Benda for helpful comments on quantum chemistry methods. The authors thank Ryan Babbush, Thomas O'Brien, Jarrod McClean, Daniel Marti-Dafcik, Sam McArdle, and Norbert Schuch for helpful comments. Render of the methylene and acetaldehyde molecules were supplied by Annina Lieberherr. The authors acknowledge the EPSRC Hub Grant under the Agreement No. EP/T001062/1 for hardware provision. The authors also acknowledge funding from the EPSRC projects Robust and Reliable Quantum Computing (RoaRQ, EP/W032635/1) and Software Enabling Early Quantum Advantage (SEEQA, EP/Y004655/1). H.H.S.C. acknowledges the support of the Croucher Foundation, Hong Kong. M.L.G. acknowledges the support of a Rhodes Scholarship. B.K. thanks the University of Oxford for a Glasstone Research Fellowship and Lady Margaret Hall, Oxford for a Research Fellowship. B.K. thanks UKRI for the Future Leaders Fellowship project titled Theory to Enable Practical Quantum Advantage (MR/Y015843/1). The numerical modeling involved in this study made use of the Quantum Exact Simulation Toolkit (QuEST), and the recent development QuESTlink [38], which permits the user to use Mathematica as the integrated front end, and pyQuEST [84], which allows access to QuEST from Python. We are grateful to those who have contributed to all of these valuable tools. The authors would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility [85] and the IBM Quantum services in carrying out this work. This research was funded in whole, or in part, by UKRI (MR/Y015843/1). For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

B.K. conceived the idea, produced proofs, and performed scaling analyses, H.H.S.C. performed the methylene simulations, spin-chain hardware experiment and developed quantum-chemistry-specific ideas, R.M. performed ideal and noisy simulations of the Hubbard model, M.L.G. performed ideal and noisy variational simulations of the spin chain. All authors contributed to writing the manuscript.

APPENDIX A: SUMMARIZING LIMITATIONS OF SHADOW SPECTROSCOPY

As mentioned in the Introduction, one main limiting factor of shadow spectroscopy is that it requires an initial state with a sufficient overlap with the eigenstates whose energy differences one aims to extract, i.e., typically the lowest-lying eigenstates in practice. Obtaining good initial states is, however, a challenging task, which can be a bottleneck for even advanced, fault-tolerant phase-estimation protocols [27]. This challenge can, indeed, be attributed to general, exponential hardness results for finding ground or eigenstates [86].

As such, applying quantum or classical heuristics becomes crucial for obtaining approximations of low-lying eigenstates and their state-preparation circuits, such as by using tensor-network methods [87]. Regarding quantum heuristics, in the present work we explored initialization through VQE in which case, hardness results manifest in the difficulty of circuit training: for deep, randomly initialized variational circuits one indeed suffers from exponentially increasing training costs due to barren plateaus [88] or due to exponentially many local traps [54,55] and thus the efficacy of obtaining good initialization for shadow spectroscopy boils down to finding well-motivated, problem-specific ansätze for which efficient training is possible.

The other main limiting factor of algorithmic shadow spectroscopy is the accuracy of Hamiltonian simulation: as seen from the Heisenberg chain and Fermi-Hubbard model examples, unitary-evolution simulations are only approximate, introducing algorithmic errors that shift the positions of the peaks. However, in this work we show that these errors can be mitigated (see Sec. III A 2) and indeed theoretical guarantees are available in the literature [58]. As mentioned, advanced time-evolution approximations that have lower algorithmic errors are also constantly being invented.

Furthermore, gate noise and shot noise only affect the signal-to-noise ratio (peak intensities), not peak-position accuracy as long as the noise model is probabilistic, e.g., Pauli noise, as we prove. However, our proofs do not apply to coherent gate errors or to error models that are not of the form of Eq. (C1), such as damping errors. We do not expect this to be a major drawback as in typical experiments one applies twirling operations to convert the aforementioned nonprobabilistic noise sources into a form that complies with Eq. (C1). Indeed, as we demonstrate numerically, even without twirling operations the approach is highly noise robust even for damping errors that violate our noise-model assumptions in Eq. (C1). In addition, these claims are in good agreement with our hardware experiment.

Our approach achieves a scaling $1/T$ in the sense that the evolution time required for an energy gap Δ to be resolved is $T \sim 1/\Delta$ (while often the terminology Heisenberg-limited scaling [89] is used in this context, this should be distinguished from the Heisenberg limit in quantum metrology); it follows that quantum resources as both circuit depth and the number of shots (assuming at least one shot taken per time step) required scales similarly with $1/\Delta$. This can be contrasted with the standard shot-noise scaling of energy estimation in VQE where the precision scales with the inverse square root of the number of shots. While achieving a finite threshold in amplitude precision is important in order to distinguish a peak from baseline white noise, the asymptotic resolution in the energy-gap estimate (i.e., frequency estimate) is determined asymptotically entirely by the peak width, which in the noiseless case is entirely determined by T^{-1} (whereas in the noisy case it is determined by the effective coherence time).

However, a limitation in this context is spectral gap closure, i.e., the smaller the spectral gap the deeper the necessary circuits. For example, in systems with "strong correlation," such as the Fermi-Hubbard model where $U/t \ll 1$ or at chemical bond dissociation limits in molecules, low-lying eigenstates may get exponentially close to each other, forming a dense band with degenerate or near-degenerate states. Note, however, that resolving individual eigenvalues becomes equally difficult with phase-estimation protocols in these cases. Other adversarial examples, which may give exponentially vanishing signal intensities include random quantum states due to the exponential concentration of eigenvalues of traceless Hamiltonians around $E = 0$. As such, the primary application of the present approach is to accurately resolve excitation energies in gapped quantum systems where the gap closes no faster than polynomially. Indeed, our numerical results in Sec. III B demonstrate that the technique is able to resolve the energy-gap regime near a conical intersection. While even answering the question whether a broad class of quantum systems is gapped in the thermodynamic limit is an undecidable problem [90], most practically relevant problems involve gapped systems, e.g., molecular systems.

Finally, we note that while shadow spectroscopy is formally similar to physical spectroscopy, it is strictly more general. The reason is that the intensities of the peaks in conventional spectroscopy are determined by the transition dipole moment, in the present case we have significant flexibility in choosing various other operators that we can estimate from classical shadows. Thus the primary aim of shadow spectroscopy is not necessarily to faithfully simulate conventional spectroscopy, but to rather estimate differences of eigenvalues making as frugal use of quantum resources as possible. Consequently, the approach can resolve peaks corresponding to spin-forbidden transitions, which will not be observed in optical spectroscopy. In contrast, if the aim was to faithfully simulate conventional spectroscopy, shadow spectroscopy can still in principle be used by estimating specifically the transition dipole moment from shadows.

APPENDIX B: COMPARISON TO THE LITERATURE

1. Simulated spectroscopy

A number of works have used similar ideas to track the time-dependent expected value of some specific observable to reveal transition energies in certain spin problems—in direct analogy with spectroscopy experiments. First, Ref. [91] proposed to simulate the evolution of a spin system under a diabatically ramped magnetic interaction, whereby estimating the time-dependent signal of an observable allows identification of transition energies. Second, Ref. [36] proposed to measure the time-dependent signal of a single observable O that anticommutes with the Hamiltonian $\{\mathcal{H}, O\} = 0$—for certain simple spin problems such an anticommuting observable can be constructed straightforwardly. Third, coupling the system to an ancilla qubit via an interaction term and measuring the time-dependent expected value of the Pauli-Z operator can also be used to identify spectroscopic transitions [92].

Indeed, problem-specific observables, as identified by the aforementioned works for spin problems, can be readily adopted for our approach. While the present work can indeed similarly be viewed as a variant of spectroscopy, it distinguishes itself as we estimate the signal of a very large number of observables using only very few circuit repetitions. Furthermore, the present approach can be applied, in principle, to any problem Hamiltonian or at least to ones that give a non-negligible signal upon measuring local observables—while noting that nonlocal operators may also be measured efficiently by other variants of classical shadows.

2. Quantum algorithms for excited states

The low-lying excited-state spectrum of a system can of course be directly calculated by evaluating the ground- and excited-state energy levels. A great plethora of variational methods, e.g., deflation [93,94], equations of motion [95,96], variational imaginary-time evolution [97], subspace search [98] have been proposed. Most relevant to our method are those which use a quantum computer to generate subspaces, which we discuss presently.

In subspace expansion [37], one can discover low-energy excited states from an estimated ground state $|\tilde{\psi}_G\rangle$ by applying low-weight operators as excitations to this initial state. Typically Pauli operators are used to produce a new set of states $|\psi_k\rangle = \mathcal{P}_k |\tilde{\psi}_G\rangle$ for calculating the overlaps $H_{kj} = \langle\psi_k| \mathcal{H} |\psi_j\rangle$ and $S_{kj} = \langle\psi_k|\psi_j\rangle$. Finally, diagonalizing H_{kj} then reveals better ground-state energies than that of $|\tilde{\psi}_G\rangle$. In our approach in Statement 1, the intensity of the signals is expressed as $I_{kl} = c_k^* c_l \langle\psi_k|O|\psi_l\rangle$. As such, we aim to apply operator pools that give non-negligible overlaps $\langle\psi_k|O|\psi_l\rangle$ and thus we can draw a loose analogy with subspace expansion where similar operators are required. However, the difference in efficiency between the present approach and subspace expansion is crucial: in subspace expansion one estimates the overlaps one-by-one resulting in a linear complexity $\mathcal{O}(N_o)$ whereas our effective use of classical shadows reduces the quantum resources to only logarithmic $\mathcal{O}(\log N_o)$. Nevertheless, the literature on subspace expansion will be relevant in identifying problem-specific observables O that give significant overlaps which we have considered in Sec. III B, e.g., excitation operators in quantum chemistry or certain local observables in spin problems. Furthermore, a significant advantage of the present approach is that explicit knowledge of which operators give intense signals is not necessary as we can efficiently reconstruct all q-local Pauli strings and identify the useful signals via statistical autocorrelation tests that we discussed in Sec. II C.

Methods that generate instead Krylov [99] and Krylov-like subspaces [100,101] are, in comparison, more loosely related to our approach. These algorithms generate a subspace by applying to a supplied initial state either (a) powers of the Hamiltonian or (b) real- or imaginary-time Hamiltonian evolution steps, which approximate powers of the Hamiltonian in the small time-step limit as per the Taylor expansion of the matrix exponential. Computing the subspace Hamiltonian and overlap matrix elements on a quantum computer and diagonalizing it classically gives approximations to eigenvalues of the modeled Hamiltonian. While our approach features time evolution of an initial state as well as classical signal postprocessing and matrix diagonalization, it does not have the burden of evaluating Krylov matrix elements, nor does it require an initial state with good ground-state overlap.

3. Variants of phase estimation

A number of works aim to estimate the expected value $\langle\psi|e^{-it\mathcal{H}}|\psi\rangle$ or a signal that is effectively equivalent to it. First, in NISQ to early-fault-tolerant iterative phase-estimation protocols [81], the autocorrelation signal $\langle\psi|e^{-it\mathcal{H}}|\psi\rangle$ is estimated by applying time-evolution steps controlled by a single ancilla and then measuring out the ancilla in shots; the probability of observing the $|0\rangle$ outcome is the autocorrelation signal. This signal contains frequencies precisely at eigenenergies and with amplitudes determined by the fidelity of the initial state with respect to the eigenstate—the approach thus requires a reasonably large initial overlap to the targeted eigenstates due to shot noise. A group of recent methods called statistical phase estimation builds on this single ancilla iterative phase-estimation protocol and uses sophisticated classical signal-processing techniques to extract as much information as possible from the single ancilla signal, thereby bringing down the sampling cost and circuit depth for generating said autocorrelation signal on a quantum computer—see e.g., Refs. [89,102,103]. Bespoke architectures have also been developed in Ref. [81] to reduce quantum resources required for a controlled evolution.

Problem-specific techniques have also been developed for quantum chemistry applications. For example, phase estimation for investigating elusive high-energy excitations [104] have been proposed. Refs. [105,106] use Hadamard-test circuits to estimate expected values of excitation operators—and ancilla-free variants are also discussed. Indeed, in Sec. III B we discuss such problem-specific operators and construct a large number of these that we can simultaneously measure with powerful classical shadow techniques.

Other approaches seek to generate the autocorrelation signal without the ancillary qubits. References [107,108] propose to create a superposition of two eigenstates and performs a Ramsey-type experiment by applying a time evolution. The approach alleviates the use of a controlled evolution but requires a Prep and a Prep† oracle that prepare a superposition of two eigenstates. Besides similarly requiring no controlled evolution, another significant advantage of our approach is that it requires no explicit knowledge of a state-preparation oracle. In fact, the present approach is very flexible with respect to the supplied initial state—as we demonstrated in numerical simulations—and can even work with input states that are a superposition of a large number of eigenstates.

Finally, the present approach can also be compared to Ref. [109], whereby one first applies a Haar-random circuit to the $|0\rangle$ state, applies a time-evolution operator, applies the inverse of the random circuit and finally measures the survival probability as probability of the $|0\rangle$ state. While we similarly apply randomized measurements, the advantage of our approach is that we need only use single-qubit rotations to obtain classical shadows.

In summary, these prior works estimate a single autocorrelation signal that may require cumbersome quantum resources such as controlled time-evolution or state-preparation oracles, and then perform classical postprocessing to extract eigenvales from the autocorrelation signal. In contrast, here we only use logarithmic quantum resources to cheaply estimate a large number of operator expectation value signals, which we then classically analyze in a nontrivial postprocessing step to reveal energy gaps.

APPENDIX C: NOISE ROBUSTNESS

1. Noise model assumptions

We recollect properties of a general class of noise models from Ref. [61]: most typical noise models used in practice, such as depolarizing, bitflip, or dephasing noise, admit the following probabilistic interpretation: a noisy quantum gate $\Phi(\rho)$ acts with probability $1-\epsilon$ as a noise-free operation U and with probability ϵ as an erroneous execution of the gate as

$$\Phi_k(\rho) = (1-\epsilon)U_k\rho U_k^\dagger + \epsilon\Phi_{\mathrm{err}}(U_k\rho U_k^\dagger). \tag{C1}$$

Here U_k is the kth ideal quantum gate in a quantum circuit and the CPTP map Φ_{err} happens with probability ϵ and represents all errors that happened during the execution of the gate.

A quantum circuit is then a composition of a series of ν such quantum gates and when this circuit is applied to any reference state it prepares the density matrix

$$\rho = \eta\rho_{id} + (1-\eta)\rho_{\mathrm{err}}. \tag{C2}$$

Here $\rho_{id} := |\psi\rangle\langle\psi|$ is the ideal noise-free state that the ideal circuit would prepare, ρ_{err} is an error density matrix that contains a mixture of all error events in the circuit and $\eta = (1-\epsilon)^\nu$ is the probability that none of the gates have undergone errors.

2. Robustness to gate noise

Theorem 1. We assume that a series of noisy quantum circuits are used to prepare the series of time-evolved quantum states $\rho(t_n)$ under the general noise model described in Appendix C 1. Any measured time-dependent signal, as the expected value of a Pauli string P_k, then decomposes as

$$S_k^{\mathrm{noisy}}(t_n) = \eta S_k(t_n) + (1-\eta)W_k(t_n).$$

The first term $S_k(t_n)$ is the ideal signal while the second term is a noise contribution whose amplitude is bounded $|W_k(t_n)| \le \|\rho_{\mathrm{err}}(t_n) - \mathrm{Id}/d\|_1$ via the distance from the maximally mixed state, i.e., the error term is zero for global depolarizing noise. Depending on the simulation algorithm used, the weight η is either a constant or a time-dependent exponential envelope and thus the Fourier transform $\mathcal{F}[\eta S_k](\omega)$ has peaks centered exactly at peaks of the ideal spectrum $\mathcal{F}[S_k](\omega)$.

Proof. We define the noisy time-dependent signal as

$$S_k^{\mathrm{noisy}}(t_n) = \mathrm{Tr}[\rho(t_n)P_k],$$

which is the expected value of a Pauli string P_k as estimated, e.g., via classical shadows. We assumed that the time-evolved quantum states $\rho(t_n)$ are prepared by a series of quantum circuits that are described by the error model in Appendix C 1 such that every quantum state decomposes as

$$\rho(t_n) = \eta(t_n)\rho_{id}(t_n) + [1-\eta(t_n)]\rho_{\mathrm{err}}(t_n).$$

Here the noise-free quantum state $\rho_{id}(t_n) = U(t_n)|\psi_0\rangle\langle\psi_0|U^\dagger(t_n)$ is prepared by the noise-free simulation circuit $U(t_n)$; it has a weight given by the probability that none of the gates in the circuit has undergone a noise event $\eta(t_n) = (1-\epsilon)^{\nu(t_n)}$; this weight depends on the per-gate error rate ϵ and on the number $\nu(t_n)$ of gates that are used to prepare the time-evolved quantum state $\rho(t_n)$. It immediately follows that the noisy signal decomposes as

$$S_k^{\mathrm{noisy}}(t_n) = \eta(t_n)S_k(t_n) + [1-\eta(t_n)]W_k(t_n),$$

where the noise signal $W_k(t_n) := \mathrm{Tr}[\rho_{\mathrm{err}}(t_n)P_k]$ can be bounded generally as

$$\begin{aligned}|\mathrm{Tr}[\rho_{\mathrm{err}}(t_n)P_k]| &= |\mathrm{Tr}[\rho_{\mathrm{err}}(t_n)P_k] - \mathrm{Tr}[\tfrac{\mathrm{Id}}{d}P_k]| \\ &\le \|\rho_{\mathrm{err}}(t_n) - \tfrac{\mathrm{Id}}{d}\|_1.\end{aligned}$$

Above we have used that Pauli strings are traceless and have unit operator norm. We have also used the general bound of Ref. [61], which can be applied to any observable O as

$$|\mathrm{Tr}[O\rho_1] - \mathrm{Tr}[O\rho_2]| \le \|O\|_\infty\|\rho_1 - \rho_2\|_1.$$

Finally, we derive the time-dependent weight $\eta(t_n)$: when the time evolution is simulated by a variational circuit, whose circuit structure is constant throughout the simulation as in Sec. III A 1, then the weight is a time-independent constant $\eta = (1-\epsilon)^\nu$. This constant is well approximated as [61,110,111]

$$\eta = (1-\epsilon)^\nu = e^{-\epsilon\nu} + \mathcal{O}(\epsilon^2/\nu). \tag{C3}$$

As such, the spectrum of the ideal signal $S_k(t_n)$ is only rescaled by a constant factor that depends exponentially on the expected number of errors in the full variational circuit as $\xi = \nu\epsilon$.

On the other hand, when the time evolution is simulated by product formulas as in Sec. III A 2 then the weight is time dependent as $\eta(t_n) = (1-\epsilon)^{nN_l}$ where N_l is the number of quantum gates in a single layer of the product-formula circuit. This is well approximated by a time-dependent exponential envelope function as

$$\eta(t_n) = (1-\epsilon)^{nN_l} = e^{-\epsilon N_l n} + \mathcal{O}\left(\frac{\epsilon^2}{nN_l}\right).$$

The exponential envelope in the ideal part of the signal $\eta(t_n)S_k(t_n)$ does not change the peak center in the Fourier transform but only broadens it for the following reason: we consider a time-continuous function $\eta(t)S_k(t)$ from which the signal $\eta(t_n)S_k(t_n)$ are samples evaluated at times $t_n = n\Delta t$. The Fourier transform of $\eta(t)S_k(t)$ can then be calculated analytically as

$$\mathcal{F}[e^{-\alpha t}S_k(t)](\omega) \propto \mathcal{F}[S_k](\omega) * \frac{\alpha}{\alpha^2+\omega^2}, \qquad \text{(C4)}$$

where $\alpha := \frac{\epsilon N_l}{\Delta t}$. The above is simply a convolution of the ideal spectrum $\mathcal{F}[S_k](\omega)$ with a Lorentzian function of width α—this convolution does not change the center of the peaks but only broadens their lineshape. ■

One could compare the above results to error-robustness guarantees of Ref. [35]. Specifically, Ref. [35] assumed an error model $\mathcal{UE}$ that factorizes into a product of the ideal channel $\mathcal{U}$ and a noise channel $\mathcal{E}$, and proved for small errors $\|\mathcal{E}\| \ll 1$, i.e., for small circuit error rates $\xi \ll 1$, that in first-order perturbation the phase evolution is unchanged by the noise and only the amplitude of the signal decreases. Theorem 1 presents considerably stronger error-robustness guarantees that apply to finite error rates ξ and to almost all typical error channels used in practice.

3. Approximate bounds on the error term

The above result establishes that the noisy signal $S_k^{\rm noisy}(t_n)$ is a superposition of two signals: the first component is the ideal signal $\eta S_k(t_n)$ up to a rescaling factor η, which however, does not change the center of peaks in the spectrum. The second component is a noise signal $(1-\eta)W_k(t_n)$ that may contribute artefacts to the spectrum. However, the magnitude of this error signal is bounded by the distance from white (global depolarising) noise $\|\rho_{\rm err}(t_n) - \frac{\rm Id}{d}\|_1$.

While it can be proved rigorously that in random circuits white noise is approached for an increasing circuit depth, the statement was recently analyzed for a broad class of practical shallow circuits [110,111]. In particular, the exact formula for random circuits was found to be a good approximation for practical circuits—albeit with a potentially large $a \approx 0.1$ problem-dependent prefactor. Let us now adapt bounds of Ref. [110,111] obtained for random circuits of ν gates as

$$(1-\eta(t_n))\|\rho_{\rm err}(t_n) - \tfrac{\rm Id}{d}\|_1 \approx \frac{a \times e^{-\xi}\xi}{\sqrt{\nu}}. \qquad \text{(C5)}$$

In the case when variational circuits of fixed depth are used to simulate time evolution, as in Sec. III A 1, the above bound is constant throughout the simulation. We thus find that the error component of the signal is globally bounded approximately as $|[1-\eta(t_n)]W_k(t_n)| \lessapprox \alpha\nu^{-1/2}$ assuming $\xi \leq 1$. Given the $\nu^{-1/2}$ dependence on the number of gates, we can expect that the artefacts in the spectrum due to noise diminishing as we increase the system size due to the increasing number of gates ν.

Let us now consider product formulas, as in Sec. III A 2, whereby the number of gates is time dependent as $\nu(t_n) = N_l n$ and thus yields the time-index-dependent upper bound in Eq. (C5) as $a\epsilon e^{-\epsilon N_l n}\sqrt{N_l n}$.

Let us assume a worst-case scenario whereby the error term saturates its above upper bound and thus

$$\begin{aligned} S_k^{\rm noisy}(t_n) &= \eta(t_n)S_k(t_n) + [1-\eta(t_n)]W_k(t_n) \\ &= e^{-\epsilon N_l n}[S_k(t_n) + a\epsilon\sqrt{N_l n}], \end{aligned}$$

whereby inside the square brackets we have a sum of the ideal signal and an additive worst-case noise. The noise component is a signal that grows proportionally with the square root of the time index n. We can compute the square of the l_2 norm of the error component of the signal as

$$\|\alpha\epsilon\sqrt{N_l n}\|_2^2 = a^2\epsilon^2 N_l \sum_{n=1}^{N_T} n \approx a^2\epsilon^2 N_l N_T^2.$$

Given the l_2 norm is invariant under the Fourier transform, we find that while the l_2 norm of the ideal signal $\|S_k(t_n)\|_2$ is proportional to the peak height in the spectrum, the total spectral contribution of the artefacts is given by the l_2 norm as $a\epsilon N_T\sqrt{N_l} = a\xi/\sqrt{N_l}$. As such, under the above assumptions, the contribution of the error term to the spectrum diminishes as we increase the system size due to the increasing number of gates in a layer N_l of a product formula but while assuming a constant ξ.

4. Lineshape broadening

In the case when the time evolution is simulated by product formulas then the ideal signal is multiplied by an exponential envelope, which leads to a broadening of the lineshapes in the shadow spectrum via Eq. (C4). However, as long as the total circuit error rate is reasonably small as $\xi \approx 1$ the broadening is not a limiting factor in practice. In particular, the Lorentzian width of the peak in Eq. (C4) is given by the factor $\alpha = \epsilon N_l/\Delta t$. Since the total circuit error rate is expressed as $\xi = \epsilon N_l N_T$ using the overall number of simulation steps N_T, the width of the Lorentzian peak is simplified as $\alpha = \xi/N_T\Delta t$. Similarly, the natural linewidth due to finite simulation time can be expressed as $\propto 1/(N_T\Delta t)$.

As such, the broadening due to gate noise relative to the natural linewidth is simply given by the circuit error rate ξ. This guarantees that in practice, i.e., when $\xi \leq 1$ the peaks are not broadened by gate noise. This is nicely confirmed in Fig. 5 where the peaks for different ξ have approximately the same width as given by the natural linewidth. Furthermore, all zero-frequency components in the ideal signal, which are ideally removed by the standardization of the data matrix **D**, are broadened due to the exponential envelope and thus appear in the shadow spectrum as an intense peak at the low-frequency end. Indeed, this is nicely confirmed in Fig. 5.

APPENDIX D: CLASSICAL POSTPROCESSING

In this section, we describe details of our classical postprocessing technique and analyze its computational complexity. We first summarize the algorithm as illustrated in Fig. 1:

1. measure N_o time-dependent signals with classical shadows, with sample compelxity $\mathcal{O}(\log N_o)$;
2. possibly reject signals that are not statistically significantly different from noise via a Ljung-Box test;
3. arrange the signals into a matrix **D** as row vectors;
4. calculate the square matrix $\mathbf{C} \propto \mathbf{D}^T\mathbf{D}$—this is the time-determining step with a complexity $\mathcal{O}(N_o)$;
5. find the dominant eigenvectors $v_1, v_2, \ldots v_c$ of **C**—we need to choose a cutoff, e.g., $c = 4$ in Fig. 2, based on the eigenvalue distribution of **C**;
6. finally, we effectively Fourier transform these dominant eigenvectors. More specifically, we obtain the spectral (Fourier) intensity at frequency ω by calculating the spectral cross-correlation matrix; $\mathbf{X}(\omega)$ from the dominant eigenvectors $v_1, v_2, \ldots v_c$
7. the spectrum $\lambda(\omega)$ is then obtained as the principal component of each matrix $\mathbf{X}(\omega)$.

Below we first detail the correlation analysis technique from steps (6) and (7): this technique is closely related to generalized coherence techniques in classical signal processing [71–73], however, it would be prohibitively expensive when directly applied to our large number of observables $N_o \gg N_T$.

We thus describe a dimensionality-reduction technique in steps (2)–(5), which is closely related to a standard principal component analysis but slightly differs as we standardize the time signals f_k (row-wise mean and standard deviation is fixed) as in Eq. (5). In fact, this algorithm is directly analogous to well-established subspace methods in signal processing, such as the MUSIC algorithm [41]. In these signal subspace methods, one estimates a signal autocorrelation matrix, which is formally equivalent to our square matrix **C** that we report in Eq. (D3). One then defines a signal subspace that is by definition spanned by the eigenvectors that correspond to the largest c eigenvalues of the matrix **C**. Then, the noise subspace is spanned by the complement of the signal subspace. In contrast to our method detailed in Appendix D 2, MUSIC deviates as it does not use the signal eigenvectors but rather works directly with the noise subspace: the MUSIC spectrum is defined as $\propto [v(\omega)^T\tilde{\mathbf{C}}v(\omega)]^{-1}$, where $v(\omega)$ are discretized Fourier components of frequency ω and $\tilde{\mathbf{C}}$ is a projection onto the noise subspace of **C**. Indeed, if a frequency $v(\omega)$ is contained in the signal subspace then $v(\omega)^T\tilde{\mathbf{C}}v(\omega) \to 0$ and thus we observe a peak in the spectrum.

1. Correlation analysis of multiple signals

We assume that we are given a set of vectors $v_1, v_2, \ldots v_c$ as time signals such that each signal, e.g., $v_1(n)$ is of total length N_T and is standardized. We calculate the spectral cross-correlation matrix between these signals as

$$[\mathbf{X}(\omega_n)]_{kl} = \mathcal{F}[X_{kl}](\omega_n), \tag{D1}$$

where we use the discrete Fourier transform, or the fast Fourier transform of the cross-correlation between the individual signals as

$$X_{kl}(m) = \sum_{n=1}^{N_T-m-1} v_k(n+m)v_l(n).$$

Here $X_{kl}(m)$ is a time signal with the temporal index m and it quantifies the cross-correlation between the signals $v_k(n)$ and $v_l(n)$ as the overlap between the signals after a time lag of m. If the signals were infinitely long than the matrix elements are simply products of the individual Fourier transforms as

$$[\mathbf{X}(\omega_n)]_{kl} \propto [\mathcal{F}[v_k](\omega_n)]^* \times \mathcal{F}[v_l](\omega_n) \tag{D2}$$

and the matrix $\mathbf{X}(\omega_n)$ would be rank 1 with only one nonzero singular value as $\lambda(\omega_n) = \sum_k |\mathcal{F}[v_k](\omega_n)|^2$, which is just the squared sum of the individual spectra from Eq. (E1).

On the other hand, if the data is noisy and finite then the matrix $\mathbf{X}(\omega_n)$ is not simply rank 1 and thus by calculating the dominant singular values at each frequency $\lambda(\omega_n)$ we aim to estimate the spectral density of the ideal noise-free component.

The computational complexity of this approach is as follows. Calculating the cross correlations between the signals takes $\mathcal{O}(c^2N_T)$ time while calculating the matrices $\mathbf{X}(\omega_n)$ through fast Fourier transforms takes $\mathcal{O}(c^2 N_T \log N_T)$. Finally, obtaining the dominant singular values takes in the worst-case scales as $\mathcal{O}(c^3 N_T)$. As such, the worst-case scaling of the entire procedure is $\mathcal{O}(c^3 N_T)$.

Note that we could directly apply this procedure to our full dataset by considering $v_k = f_k$ from Eq. (5). However, the computational complexity via $c = N_o$ would then be $\mathcal{O}(N_o^3 N_T)$, which is prohibitive in practice for a large number of observables $N_o \gg N_T$. For this reason we first apply a dimensionality reduction to our data and consider only a few dominant eigenvectors as $v_1, v_2, \dots v_c$.

2. Dimensionality reduction

In this section we assume that we have performed a time evolution in N_T time increments and recorded classical shadows at each time increment, each containing N_s snapshots as summarized in Statement 2. The classical computational complexity of reconstructing N_o observables from the shadow data at each time increment is $\mathcal{O}(N_o N_s)$ using the algorithm described in Ref. [28].

a. Calculating the square matrix

We order the time-dependent expected values into the data matrix $\mathbf{D} \in \mathbb{R}^{N_o \times N_T}$ as defined in Eq. (5). We can derive the time complexity of computing the square matrix $\mathbf{C} \in \mathbb{R}^{N_T \times N_T}$, which we obtain via the matrix-matrix product $\mathbf{D}^T\mathbf{D}$ as

$$[\mathbf{C}]_{mn} = \frac{1}{N_o}\sum_{k=1}^{N_o}[\mathbf{D}]_{mk}[\mathbf{D}]_{kn} = \frac{1}{N_o}\sum_{k=1}^{N_o} f_k(m) f_k(n). \tag{D3}$$

As such, a naive implementation of calculating the matrix product has a complexity $\mathcal{O}(N_T^2 N_o)$.

b. Calculating eigenvectors of the square matrix

Imagine a time signal $v(n) \in \mathbb{R}^{N_T}$ for which we calculate the expression with respect to the square matrix $\mathbf{C}$ as

$$\begin{aligned} N_o\, v^T \mathbf{C} v &= \sum_{m,n=1}^{N_T} v(m) C_{mn} v(n) \\ &= \sum_{m,n=1}^{N_T}\sum_{k=1}^{N_o} f_k(m) f_k(n) v(m) v(n) \\ &= \sum_{k=1}^{N_o}\left[\sum_{m=1}^{N_T} f_k(m) v(m)\right]\left[\sum_{n=1}^{N_T} f_k(n) v(n)\right]. \end{aligned}$$

Let us introduce the notation $\langle f_k, v\rangle := \sum_{n=1}^{N_T} f_k(n) v_n$ for scalar products between the kth observable's signal f_k and the probe signal v. We finally obtain the expression

$$v^T \mathbf{C} v = \frac{1}{N_o}\sum_{k=1}^{N_o} |\langle f\,, v\rangle_k|^2, \tag{D4}$$

as the average squared overlap between the individual signals and the probe signal $v(n)$. As such, the average squared overlap is maximized by the dominant eigenvector v_1 of the real, symmetric matrix $\mathbf{C}$ as

$$v_1^T \mathbf{C} v_1 = \max_{\|v\|_2=1} v^T \mathbf{C} v = \max_{\|v\|_2=1} \frac{1}{N_o}\sum_{k=1}^{N_o} |\langle f\,, v\rangle_k|^2.$$

We calculate the dominant eigenvectors $v_1, v_2, \dots v_c$ of $\mathbf{C}$ where we introduce a cutoff based on eigenvalues of the matrix $\mathbf{C}$. The complexity of an exact diagonalization is $\mathcal{O}(N_T^3)$ where time is independent of the number of observables—it only depends on the number of time steps N_T. Finally, the complexity of dimensionality reduction is $\mathcal{O}(N_T^3 + N_T^2 N_o)$. In numerical simulations we estimate the absolute time required to perform the entire dimensionality reduction algorithm in Fig. 7(c) and find good agreement with the theoretical scaling.

APPENDIX E: SHOT-NOISE PROPAGATION

In this section we derive simple analytical models to understand how shot noise affects our shadow spectra. We then verify in numerical experiments using actual shadow spectra that indeed these analytical models very well explain the effect of shot noise on shadow spectra.

1. Average spectral densities

Let us consider a collection of N_o time-dependent signals that we determine using classical shadows as $f_i(n) = \langle\psi(t_n)|P_i|\psi(t_n)\rangle$ with the discrete temporal index $n \le N_T$ and the time variable $t_n = n\Delta t$. The simplest possible way to estimate common frequencies in the signals is by estimating the mean-squared spectrum (I_{MSS})

$$I_{\text{MSS}}[\omega_n] = \frac{1}{N_o}\sum_{i=1}^{N_o} |F_i(n)|^2, \tag{E1}$$

as the average of the square of the fast Fourier transform $F_i(n) := \mathcal{F}[f_i](\omega_n)$ of each signal.

For a simple analytical model we precisely state the signal-to-noise ratio: in Eq. (4) we derived a general form of the signals. For ease of notation we assume the system has one transition (e.g., $k = 1$ and $l = 2$) with a corresponding frequency $\nu = E_k - E_l$ and intensity $I_i = |I_{ikl}| = |I_{i,1,2}|$. We showed in Eq. (4) that the ith observable gives

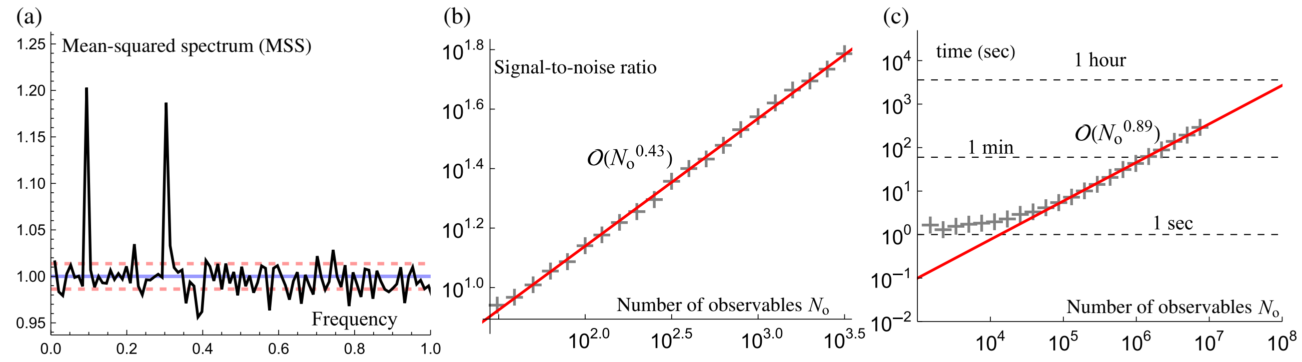

FIG. 7. (a) The same spectrum as in Fig. 2 but showing the result of a simple postprocessing technique whereby we calculate the sum of squares of the individual spectra as described in Appendix E 1. (Blue, solid) Theoretical expected value 1.00 of the baseline and (red, dashed) theoretical standard deviation of the baseline 0.0137 match well the empirical sample mean of the baseline 0.994 and standard deviation 0.0139. (b) The signal-to-noise ratio in the spectrum—the gap between the baseline and the peak relative to the standard deviation of the baseline—is improved as $\propto \sqrt{N_o}$ as described in Appendix E 1. (c) The computation time of the postprocessing step is asymptotically linear in the number of observables while deviation from linear is expected at low N_o where calculating the eigenvectors of the constant-sized matrix $\mathbf{C}$ dominates the time. The absolute time is quite reasonable for even a very large number of observables, i.e., less than an hour for $N_o \propto 10^8$. Computed on a desktop PC at a fixed $N_T = 1000$ and increasing N_o.

rise to a signal that is a phase-shifted sinusoidal function of time

$$f_i(n) = I_i \cos(\nu t_n + \phi_i), \tag{E2}$$

where both the intensity and the phase shift depends on the observable's index i. Ideally, when there is no shot noise present, the spectral density in Eq. (E1) yields a single peak centered at $\omega_{\text{peak}} = \nu$. The corresponding peak intensity $I_{\text{MSS}}[\omega_{\text{peak}}] \propto N_T\overline{I^2}$ is given by the average $\overline{I^2} := \frac{1}{N_o}\sum_{i=1}^{N_o} I_i^2$.

We now model the case when the signal has additive shot noise due to a finite number of classical snapshots. For a reasonably large number of shots the shot noise can be very well approximated by a normally distributed random variable. It is well known that the discrete Fourier transform of normally distributed noise is normally distributed white noise. Thus, to a good approximation, every point in the discrete Fourier transform is a sum $F_i(n) + \mathcal{E}_i(n)$ of the ideal value $F_i(n)$ and an additive random variable $\mathcal{E}_i(n)$—these random variables are independent and for ease of notation we assume they have an identical variance, which we denote as $\text{Var}[\mathcal{E}_i(n)] = \epsilon^2$ and the mean is computed as $\mathbb{E}[\mathcal{E}_i(n)] = 0$. The expected value of the mean-squared spectrum at the peak center can be calculated as

$$\begin{aligned}\mathbb{E}(I_{\text{MSS}}[\omega_{\text{peak}}]) &= \frac{1}{N_o}\sum_{i=1}^{N_o} \mathbb{E}[(N_T I_i + \mathcal{E}_i(\omega_{\text{peak}}))^2]\\ &= N_T\overline{I^2} + \epsilon^2.\end{aligned}$$

We can also calculate the mean value of the baseline of the spectrum by focusing on a point in the spectrum ω_{base} that is far away from the peak center ω_{peak}, where the ideal spectrum would be zero, thus

$$\mathbb{E}(I_{\text{MSS}}[\omega_{\text{base}}]) = \frac{1}{N_o}\sum_{i=1}^{N_o} \mathbb{E}(\mathcal{E}_i(\omega_{\text{base}})^2) = \epsilon^2.$$

As such, the baseline of the spectrum is not zero but is actually a constant that is determined by the level of shot noise ϵ. Since the baseline is constant, the signal-to-noise ratio is determined by the standard deviation of the baseline, which we can calculate via the variance as

$$\text{Var}(I_{\text{MSS}}[\omega_{\text{base}}]) = \frac{1}{N_o^2}\sum_{i=1}^{N_o} \text{Var}[\mathcal{E}_i(\omega_{\text{base}})^2] = \frac{2\epsilon^4}{N_o},$$

where we used that the variance of the square of a random variable of mean zero is $\text{Var}[\mathcal{E}_i(\nu')^2] = 2\epsilon^4$. It follows that the standard deviation of the baseline is $\epsilon^2\sqrt{2/N_o}$, and thus the signal-to-noise ratio can be calculated as

$$\text{SNR} = \frac{N_T\overline{I^2}\sqrt{N_o}}{\sqrt{2}\epsilon^2} \propto N_T N_s \overline{I^2}\sqrt{N_o}, \tag{E3}$$

where in the last equation we use that the statistical fluctuations are due to the finite number of circuit repetitions (shot noise) as $\epsilon^{-2} \propto N_s$. As such, in order to resolve a peak we need to set the total shot budget $N_T N_s$ as the amount of quantum resources such that $1/\overline{I^2} < N_T N_s\sqrt{N_o}$. Indeed, the signal-to-noise ratio can be improved just by increasing $\sqrt{N_o}$ at a fixed shot budged $N_T N_s$, i.e., without using more quantum resources. Furthermore, using the prescreening described in Sec. II C one can discard observables with negligible signal guaranteeing that $\overline{I^2} > I^2_{\text{min}} > 0$, while the actual value of $\overline{I^2}$ depends strongly on the considered system.

We find that this simple model quite accurately describes the effect of shot noise on actual shadow spectra: in Fig. 7(a) we calculate the average spectral density for the same simulation as in Fig. 2. As we describe in Sec. II C, we standardize our data matrix such that the standard deviation of each signal is unity and thus the variance in the Fourier transform is approximately $\mathrm{Var}[\mathcal{E}_i(n)] \approx \epsilon^2 \approx 1$. Indeed, Fig. 7(a) confirms that the expected baseline $\epsilon^2 = 1$ (blue solid line) in the spectrum matches well the empirical sample mean 0.994. Figure 7(a) also confirms that the theoretically expected standard deviation of this baseline $\sqrt{2/N_o}\epsilon^2 = \sqrt{2/10689} \approx 0.0137$ nicely matches the empirical sample standard deviation 0.0139.

In Fig. 7(b) we randomly select N_o observables from a pool of all at most three-local Pauli strings with $\tilde{N}_o = 10689$. The expectation value of the average signal intensity for such a random selection is $\mathbb{E}[\overline{I^2}] = \frac{1}{\tilde{N}_o}\sum_{i=1}^{\tilde{N}_o} I_i^2$ the average over the full set, which is a constant and is independent of the number of selected terms $N_o < \tilde{N}_o$. Figure 7(b) shows the average of the signal-to-noise ratio of 100 random selections of N_o observables for an increasing N_o. Indeed, Fig. 7(b) confirms that the signal-to-noise ratio increases as $\propto \sqrt{N_o}$. Note that such a random selection of observables can be considered as a worst-case scenario when no information about the observables is assumed. Of course, in the present work we explore various advanced techniques to improve upon randomly selecting observables, e.g., the prescreening step described in Sec. II C can be used to identify the N_o most intense signals in the pool of total size $\tilde{N}_o$ so that a well-informed decision can be made.

2. Dimensionality reduction of periodic signals

Here we build a simple analytical model to understand how shot noise affects the dimensionality reduction approach introduced in Sec. II C. First, let us write the phase-shifted cosine function in Eq. (E2) as a linear combination of cosine and sine functions as

$$f_i(n) = I_i \cos(\nu t_n + \phi_i) = c_i \cos(t_n) + s_i \sin(t_n),$$

with amplitudes $c_i = I_i \cos(\phi_i)$ and $s_i = -I_i \sin(\phi_i)$. Here we again used the discrete temporal index $n \le N_T$ and the time variable $t_n = n\Delta t$. Since discretized vectors of cosine and sine functions are mutually orthogonal, we can write the signal as a two-dimensional vector in Fourier basis with entries as $\tilde{f}_i(t) = \sqrt{N_T}(c_i, s_i)$ where $\sqrt{N_T}$ ensures normalization of the basis vectors. Thus, each observable's signal can be represented by a two-dimensional vector (in the Fourier basis) and arranging all N_o signals into a data matrix yields the effective $2 \times N_o$-dimensional matrix $\tilde{\mathbf{D}}$, where the tilde refers to the fact that we represent the matrix in a Fourier basis, as

$$\tilde{\mathbf{D}} = \sqrt{N_T}\begin{bmatrix} c_1 & s_1 \\ c_2 & s_2 \\ \vdots & \vdots \\ c_{N_o} & s_{N_o} \end{bmatrix}, \quad \tilde{\mathbf{C}} = \frac{N_T}{N_o}\begin{bmatrix} \sum_i c_i^2 & \sum_i c_i s_i \\ \sum_i c_i s_i & \sum_i s_i^2 \end{bmatrix}.$$

Above we have also calculated the 2×2 square matrix $\mathbf{C}$. We can readily evaluate the sum of the two eigenvalues of this matrix as $c_i^2 + s_i^2 = N_T\overline{I^2}$, which is identical to the peak height in the mean-squared spectrum in Eq. (E1) where $\overline{I^2}$ is again the average peak intensity as $\overline{I^2} := \frac{1}{N_o}\sum_{i=1}^{N_o} I_i^2$. Since the eigenvalues are invariant under a unitary transformation, it immediately follows that the matrix $\mathbf{C}$ (without tilde, not in Fourier basis) has the same two nonzero eigenvalues.

Similarly to Appendix E 1, we now assume that under the effect of shot noise each signal is a sum $f_i(n) + \mathcal{E}_i(n)$ of an ideal component $f_i(n)$ and an additive random variable $\mathcal{E}_i(n)$. For ease of notation, we assume the random variables are independent and have the same variance as $\mathrm{Var}[\mathcal{E}_i(n)] = \epsilon^2$ and mean $\mathbb{E}[\mathcal{E}_i(n)] = 0$. Given the data matrix entries as $d_{mi} := f_i(m) + \mathcal{E}_i(m)$, we can calculate the noisy matrix $\mathbf{C}_{\text{noisy}}$ as

$$[\mathbf{C}_{\text{noisy}}]_{mn} = \frac{1}{N_o}\sum_{i=1}^{N_o} d_{mi} d_{in}.$$

We can calculate the mean of the noisy matrix and find that only the diagonal entries are affected by shot noise as $\mathbb{E}[\mathbf{C}_{\text{noisy}}] = \mathbf{C} + \epsilon^2 \mathrm{Id}$, resulting in a sum of the ideal matrix and a matrix that is proportional to the identity. Thus the random noise has no effect on the expected value of the eigenvectors and it just shifts the ideal dominant eigenvalues by a constant factor as $N_T\overline{I^2} + \epsilon^2$ in direct analogy with Appendix E 1.

As such, in analogy with Appendix E 1 the signal-to-noise ratio is determined by the standard deviation of the matrix entries. We can straightforwardly evaluate the variance of the off-diagonal entries $m \neq n$ as

$$\begin{aligned} \mathrm{Var}[(\mathbf{C})_{mn}] &= \mathrm{Var}\left[\frac{1}{N_o}\sum_{i=1}^{N_o} d_{mi} d_{in}\right] \\ &= \frac{1}{N_o^2}\sum_{i=1}^{N_o} \mathrm{Var}[d_{mi} d_{in}] \\ &= \frac{1}{N_o^2}\sum_{i=1}^{N_o} \mathrm{Var}[d_{mi}]\mathrm{Var}[d_{in}] = \frac{\epsilon^4}{N_o}. \end{aligned}$$

The variance of the diagonal entries can be calculated similarly and yields the same expression ϵ^4/N_o.

In summary, we considered a system with only one frequency component as in Eq. (E2) and evaluated analytically that the matrix **C** has only two nonzero eigenvalues and the sum of eigenvalues is $N_T\overline{I^2}$. Furthermore, the standard deviation of the matrix entries due to shot noise is $\epsilon^2/\sqrt{N_o}$, which shifts the nonzero eigenvalues. Thus in order to tell apart the dominant eigenvalues from the noisy ones there must be a sufficient gap between $N_T\overline{I^2}$ and $\epsilon^2/\sqrt{N_o}$. This is effectively the same conclusion we found in Appendix E 1 where we used a simple mean-square spectrum estimate. As such, while the dimensionality reduction step is not expected to decrease noise in the signals, it allows us to make use of advanced spectral cross-correlation methods described in Sec. II C—where the latter can suppress noise by a constant factor—compare Fig. 7(a) with the spectrum in Fig. 2—but we do not expect it can fundamentally outperform the simple mean-squared spectrum estimate in Appendix E 1.

APPENDIX F: OBSERVABLES IN QUANTUM CHEMISTRY

1. Fermionic operators

We now analyze the example of a peak that corresponds to the gap between the ground state and the first excited state. The peak intensity with respect to an observable O (Statement 1) is calculated via an expression formally resembling the conventional transition dipole moment as

$$I_{g\to e} \propto \langle\psi_g|\, O\, |\psi_e\rangle\,. \tag{F1}$$

Here we write both the exact ground and excited states as a configuration interaction (CI) expansion, i.e., a sum of the HF Slater determinant and post-HF corrections:

$$\begin{aligned}|\psi_s\rangle = c_0^{(s)}\,|\psi_{\mathrm{HF}}\rangle &+ \sum_{pq} c_{pq}^{(s)} a_p^\dagger a_q\,|\psi_{\mathrm{HF}}\rangle\\ &+ \sum_{pqrs} c_{pqrs}^{(s)} a_p^\dagger a_q^\dagger a_r a_s\,|\psi_{\mathrm{HF}}\rangle + \ldots,\end{aligned}$$

where $|\psi_{\mathrm{HF}}\rangle$ is the mean-field Hartree-Fock state [112]. Thus Eq. (F1) can be expanded as

$$\begin{aligned}\langle\psi_g|\, O\, |\psi_e\rangle = c_0^{(g)*}c_0^{(e)}\,&\langle\psi_{\mathrm{HF}}|\, O\, |\psi_{\mathrm{HF}}\rangle\\ &+ c_0^{(g)*}\sum_{pq} c_{pq}^{(e)}\,\langle\psi_{\mathrm{HF}}|\, O a_p^\dagger a_q\, |\psi_{\mathrm{HF}}\rangle\\ &+ c_0^{(e)}\sum_{pq} c_{pq}^{(g)*}\,\langle\psi_{\mathrm{HF}}|\, a_q^\dagger a_p O\, |\psi_{\mathrm{HF}}\rangle + \ldots\end{aligned} \tag{F2}$$

Above we only expand up to the single-excitation CI terms from which it is clear that if O cancels out the excitation operators, we get $\langle\psi_{\mathrm{HF}}|\,\psi_{\mathrm{HF}}\rangle = 1$ and the corresponding CI coefficients would contribute to $\langle\psi_g|\, O\, |\psi_e\rangle$.

In particular, choosing the observable as the Hermitian operator that corresponds to the single excitation $O = a_m^\dagger a_n + a_n^\dagger a_m$ with m acting on occupied and n acting on virtual orbitals, the leading terms in Eq. (F2) can be evaluated analytically as

$$\begin{aligned}\langle\psi_{\mathrm{HF}}|\, O\, |\psi_{\mathrm{HF}}\rangle &= \langle\psi_{\mathrm{HF}}|\, a_m^\dagger a_n + a_n^\dagger a_m\, |\psi_{\mathrm{HF}}\rangle = 0,\\ \langle\psi_{\mathrm{HF}}|\, O a_p^\dagger a_q\, |\psi_{\mathrm{HF}}\rangle &= \langle\psi_{\mathrm{HF}}|\, a_m^\dagger a_n a_p^\dagger a_q\, |\psi_{\mathrm{HF}}\rangle + 0 = \delta_{np}\delta_{mq},\\ \langle\psi_{\mathrm{HF}}|\, a_q^\dagger a_p O\, |\psi_{\mathrm{HF}}\rangle &= 0 + \langle\psi_{\mathrm{HF}}|\, a_q^\dagger a_p a_n^\dagger a_m\, |\psi_{\mathrm{HF}}\rangle = \delta_{np}\delta_{mq}.\end{aligned} \tag{F3}$$

Thus, in leading order, the intensity of the observable $\langle\psi_g|\, a_m^\dagger a_n + a_n^\dagger a_m\, |\psi_e\rangle$ is determined by the corresponding single-excitation coefficients $c_0^{(g)*}c_{mn}^{(e)}$ and $c_0^{(e)}c_{mn}^{(g)*}$ as all other terms shown explicitly in Eq. (F2) give null contributions. Of course higher-order CI terms in Eq. (F2) might also contribute, but they might have lower contributions.

Furthermore, by choosing an observable of a higher excitation, the intensity of the corresponding signal will be determined in leading order by its higher-order CI coefficients in Eq. (F2); for example, by choosing an observable $O = a_p^\dagger a_q^\dagger a_r a_s + a_s^\dagger a_r^\dagger a_q a_p$, the intensity in leading order is determined by the corresponding expansion coefficient $c_{pqrs}^{(s)}$. Thus, by simultaneously estimating expected values of all q-local fermionic operators one obtains a large number of signals but the intense signals will correspond to excitations that have dominant expansion coefficients in Eq. (F2). Such fermionic observables can be estimated using near-term-friendly fermionic shadows [30,31].

2. Jordan-Wigner encoding

By inspecting the JW transform of fermionic excitation operators, it is straightforward to see how a pool of q-local Pauli operators, even for $q \sim 4$, is sufficient to capture single-electron transitions between meaningful eigenstates, which is ideal for our classical-shadow-based approach.

Statement 5. Choosing the Pauli bitflip observable $O = X_m X_n$, the signal intensity $\langle\psi_g|\, X_m X_n\, |\psi_e\rangle$ is determined in leading order by the single-excitation expansion coefficients $c_0^{(g)*}c_{mn}^{(e)}$ and $c_0^{(e)}c_{mn}^{(g)*}$ of the state from Eq. (F3). Similarly, if we choose any other operator $O(m,n)$ by replacing any of the two Pauli-X operators in O with Pauli-Y operators and we additionally append Pauli-Z operators on any qubit other than m or n then the intensity is still determined by the same expansion coefficient.

$X_3 X_4$ 0.95	$Y_3 Y_4$ 0.95	$X_3 X_4 Z_7$ 0.95	$X_3 X_4 Z_6$ 0.95	$X_3 X_4 Z_9$ 0.95	$X_3 X_4 Z_8$ 0.95	$Y_3 Y_4 Z_9$ 0.95	$Y_3 Y_4 Z_8$ 0.95	$Y_3 Y_4 Z_7$ 0.95	$Y_3 Y_4 Z_6$ 0.95	$X_3 X_4 Z_0$ 0.95	$Y_3 Y_4 Z_0$ 0.95	$X_3 X_4 Z_1$ 0.95	$Y_3 Y_4 Z_1$ 0.95	$X_3 X_4 Z_{10}$ 0.95
$Y_3 Y_4 Z_{10}$ 0.95	$X_4 Y_3 Z_2$ 0.95	$X_3 Y_4 Z_2$ 0.95	$X_4 Y_3 Z_5$ 0.95	$X_3 Y_4 Z_5$ 0.95	$X_3 X_4 Z_{11}$ 0.95	$Y_3 Y_4 Z_{11}$ 0.95	$X_3 X_4 Z_5$ 0.94	$Y_3 Y_4 Z_5$ 0.94	$X_3 Y_4 Z_{11}$ 0.94	$X_4 Y_3 Z_{11}$ 0.94	$X_3 Y_4$ 0.94	$X_4 Y_3$ 0.94	$X_3 Y_4 Z_0$ 0.94	$X_4 Y_3 Z_0$ 0.94
$X_4 Y_3 Z_9$ 0.94	$X_4 Y_3 Z_8$ 0.94	$X_3 Y_4 Z_7$ 0.94	$X_3 Y_4 Z_6$ 0.94	$X_3 Y_4 Z_9$ 0.94	$X_3 Y_4 Z_8$ 0.94	$X_4 Y_3 Z_7$ 0.94	$X_4 Y_3 Z_6$ 0.94	$X_3 Y_4 Z_1$ 0.94	$X_4 Y_3 Z_1$ 0.94	$X_3 Y_4 Z_{10}$ 0.94	$X_4 Y_3 Z_{10}$ 0.94	$X_3 X_4 Z_2$ 0.94	$Y_3 Y_4 Z_2$ 0.94	$X_2 Y_5 Z_4$ 0.19
$X_5 Y_2 Z_4$ 0.19	$Y_2 Y_5 Z_{10}$ 0.19	$X_2 X_5 Z_{10}$ 0.19	$X_2 Y_5 Z_3$ 0.18	$X_5 Y_2 Z_3$ 0.18	$Y_2 Y_5 Z_{11}$ 0.18	$X_2 X_5 Z_{11}$ 0.18	$X_2 X_5 Z_9$ 0.18	$X_2 X_5 Z_8$ 0.18	$X_2 X_5 Z_7$ 0.18	$X_2 X_5 Z_6$ 0.18	$Y_2 Y_5 Z_9$ 0.18	$Y_2 Y_5 Z_8$ 0.18	$Y_2 Y_5 Z_7$ 0.18	$Y_2 Y_5 Z_6$ 0.18
$Y_2 Y_5$ 0.18	$X_2 X_5$ 0.18	$Y_2 Y_5 Z_0$ 0.18	$X_2 X_5 Z_0$ 0.18	$Y_2 Y_5 Z_1$ 0.18	$X_2 X_5 Z_1$ 0.18	$X_3 Y_{10} Z_4$ 0.14	$X_{10} Y_3 Z_4$ 0.14	$X_3 X_{10} Z_2$ 0.14	$Y_3 Y_{10} Z_2$ 0.14	$X_2 X_5 Z_3$ 0.14	$Y_2 Y_5 Z_3$ 0.14	$X_5 Y_2 Z_{11}$ 0.14	$X_2 Y_5 Z_{11}$ 0.14	$X_5 Y_2 Z_0$ 0.14

FIG. 8. List of the 75 at most three-local Pauli strings, which give the most intense signals. The intensities are given by the matrix elements $|\langle\psi_g|O_i|\psi_e\rangle|$, where $|\psi_g\rangle$ and $|\psi_e\rangle$ are the exact ground and the first triplet excited state of a 12-qubit LiH Hamiltonian. The most intense observables are the ones that correspond to the HOMO-LUMO triplet transition of the basis states $|111100\ldots\rangle \to |111010\ldots\rangle$, which include, e.g., X_3X_4, X_3Y_4 or any Pauli string obtained by adding a Pauli-Z operator on any other qubit, e.g., $Y_3Y_4Z_7$. The next most intense ones also generate HOMO-LUMO transitions to the triplet excited state $|111100\ldots\rangle \to |110101\ldots\rangle$, including, e.g., X_2Y_5 or any other string that is related to this by adding a Pauli-Z operator to any other qubit.

Proof. In general, the excitation operators in the JW encoding are mapped to

$$a_p^\dagger \mapsto \alpha_p^\dagger := Z_{:p}A_p^\dagger, \qquad a_p \mapsto \alpha_p := Z_{:p}A_p$$
$$a_p^\dagger a_q \mapsto iA_p^\dagger Z_{q:p}A_q,$$

where $Z_{:p} := \prod_{k=1}^{p-1} Z_k$ and $A = (X + iY)/2$ is the qubit lowering operator and $Z_{q:p} = \prod_{k=p+1}^{q-1} Z_k$. Furthermore, the HF state becomes a computational basis state $|\psi_{\mathrm{HF}}\rangle \mapsto |11\ldots1100\ldots\rangle \equiv |b\rangle$ where b is simply a binary number.

Thus for any Pauli string $O = \bigotimes_{k=1}^{N} P_{j_k}$ the expected value in Eq. (F3) splits up as a product of single-qubit expected values as

$$\begin{aligned}\langle b|\, O(m,n)\alpha_p^\dagger\alpha_q\, |b\rangle &= \langle b|\, O(m,n)A_p^\dagger Z_{q:p}A_q\, |b\rangle\\ &= \langle b_p|\, P_{j_p}A^\dagger\, |b_p\rangle\, \langle b_q|\, P_{j_q}A\, |b_q\rangle\\ &\quad\times \prod_{k\neq p,k\neq q} \langle b_k|\, P_{j_k}Q_{pq}(k)\, |b_k\rangle\,,\end{aligned}$$

where p_p is the pth bit of the binary number b and $Q_{pk}(k) = Z$ if $q < k < p$ and the identity operator otherwise. We can evaluate the product of coefficients as

$$|\, \langle b_p|\, P_{j_p}A^\dagger\, |b_p\rangle\, \langle b_q|\, P_{j_q}A\, |b_q\rangle\, | = \delta_{np}\delta_{mq}/4,$$

assuming that $m < n$ without loss of generality and we only show the absolute value of the expression for ease of notation. Above the expression evaluates in absolute value to 1 only when the two Pauli-X (or Y) operators in the observable $O(m,n)$ act on the same two qubits to which A and $A^\dagger$ act, i.e., to qubits p and q. Furthermore, $\prod_{k\neq p,k\neq q} \langle b_k|\, P_{j_k}Q_{pq}(k)\, |b_k\rangle \in \{\pm1\}$ is only a sign factor that is determined by the bitstring b and by the number of Pauli-Z operators in P and in $Z_{p:q}$. Thus we conclude that the expected values from Eq. (F3) in the JW encoding evaluate to

$$\begin{aligned}\langle b|\, O(m,n)\, |b\rangle &= 0,\\ |\, \langle b|\, O(m,n)\alpha_p^\dagger\alpha_q\, |b\rangle\, | &= \delta_{np}\delta_{mq}/4,\\ |\, \langle b|\, \alpha_q^\dagger\alpha_p O(m,n)\, |b\rangle\, | &= \delta_{np}\delta_{mq}/4.\end{aligned} \tag{F4}$$

The signal intensity $\langle\psi_g|\, O(m,n)\, |\psi_e\rangle$ of the observable $O(m,n)$ is determined in leading order by the single-excitation expansion coefficients $c_0^{(g)*}c_{mn}^{(e)}$ and $c_0^{(e)}c_{mn}^{(g)*}$ of the state from Eq. (F3). ∎

The crucial observation we make in the above statement is that even though fermionic excitation operators $a_p^\dagger a_q \mapsto iA_p^\dagger Z_{q:p}A_q$ are mapped to non-local Pauli strings in the JW encoding, we can still resolve signals of the corresponding operators just by measuring two-local Pauli strings X_pX_q or any variant where we additionally append Pauli-Z operators. Figure 8 summarises the at most three-local Pauli string signals for a 12-qubit LiH Hamiltonian that give the largest $\langle\psi_g|\, O\, |\psi_e\rangle$ values, and operators like, e.g., X_3X_4 indeed give the most intense signals. The above statement also suggest that three-local Pauli strings like $X_3Y_4Z_7$, where the Z operator is placed on any site, can still give rise to strong signals.

The information above can be used to inform which specific Pauli-string observable to reconstruct. Furthermore, we could use, e.g., classical MP2 simulations to estimate which excitations will be dominant and thus construct the corresponding families of Pauli strings. If we know which operators to measure, they can in principle be ordered into a small number of commuting groups and we could thus measure them simultaneously using well-established techniques, e.g., Ref. [113]. Since we can efficiently estimate the signal of every q-local Pauli string using classical shadows (as the data matrix $\mathbf{D}$) and perform a simple statistical autocorrelation test on the signals to determine whether they contain oscillations, we need not perform any *a priori* prediction on expected operator contributions. We did

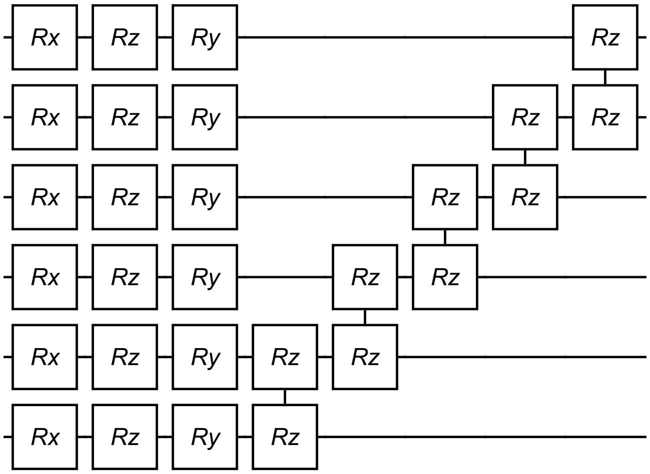

FIG. 9. A single layer of the hardware-efficient ansatz used for variational dynamics of the disordered Heisenberg chain.

this on the present LiH example and sorted the observable signals according to their statistical p values; a virtually identical list of Pauli strings as per the exactly calculated Fig. 8 was obtained.

APPENDIX G: ADDITIONAL DETAILS OF NUMERICAL SIMULATIONS

1. Ansatz-based variational quantum simulation

Our ansatz-based variational quantum simulations follow the methods introduced in Ref. [20]. At each time step, we calculate the quantum metric tensor

$$A_{ij} = \frac{\partial \langle \psi(\boldsymbol{\theta})|}{\partial \theta_i} \frac{\partial |\psi(\boldsymbol{\theta})\rangle}{\partial \theta_j} \tag{G1}$$

and gradient vector

$$C_i = \frac{\partial \langle \psi(\boldsymbol{\theta})|}{\partial \theta_i} \hat{H} |\psi(\boldsymbol{\theta})\rangle \tag{G2}$$

from the variational state. We then solve a corresponding linear system

$$\sum_j \mathrm{Re}(A_{ij})\dot{\theta}_j = \mathrm{Im}(C_i) \tag{G3}$$

to find the parameter time derivatives $\dot{\theta}_i$ such that we can compute the updated parameters as $\theta_i' = \theta_i + \dot{\theta}_i \delta t$. In practice, the metric tensor A may be ill conditioned due to degeneracies in parameter space, and so we apply Tikhonov regularization by instead finding $\dot{\boldsymbol{\theta}}$ that minimizes

$$\|\boldsymbol{C} - \boldsymbol{A}\dot{\boldsymbol{\theta}}\|^2 + \lambda \|\dot{\boldsymbol{\theta}}\|^2. \tag{G4}$$

We find that a regularization hyperparameter $\lambda = 10^{-4}$ and a time step $\delta t = 10^{-2}$ is sufficient to keep our fidelities with respect to the true dynamical state above 99%.

A Wick-rotated form of these dynamics can be used to variationally prepare ground states [50–52]. In variational imaginary-time evolution, the linear system at each time step is instead

$$\sum_j \mathrm{Re}(A_{ij})\dot{\theta}_j = -\mathrm{Re}(C_i). \tag{G5}$$

We apply this method to prepare our initial variational state, intentionally terminating before convergence to ensure non-negligible overlap with low-lying excited states.

Note that the given forms of the metric tensor [Eq. (G1)] and gradient vector [Eq. (G2)] are based on McLachlan's variational principle and do not implicitly account for a possible global phase mismatch with the target state. To ensure correct evolution, we include a global phase parameter in the ansatz. This yields the same linear system at each step as would be obtained by including explicit global phase-correction terms in the equations of motion [114]. Our numerics use a five-layer hardware-efficient ansatz of the form depicted in Fig. 9, with single-qubit and nearest-neighbor ZZ rotations only.

In Fig. 10, we provide additional simulations to highlight which factors affect secondary peak intensity and use a higher number of shots than in the hardware experiments, i.e., 4000 time steps and 3 × 50 shots at each time step. Specifically, we compare exact time evolution (blue line) to our variational approximation (orange line). Here we include gate noise comparable to typical state-of-the-art experimental devices in our variational observable calculations, applying a two-qubit depolarizing map described by

$$\Phi_{i,j}^{\lambda}(\rho) = \left(1 - \frac{16\lambda}{15}\right)\rho + \frac{\lambda}{15} \sum_{\nu_1,\nu_2 \in \{1,x,y,z\}} \sigma_i^{\nu_1}\sigma_j^{\nu_2}\rho\,\sigma_j^{\nu_2}\sigma_i^{\nu_1}, \tag{G6}$$

where σ_i^{ν} a Pauli-ν operator on qubit i and σ_i^1 is the identity. We choose error rates $\epsilon_2 = 10^{-3}$ after each two-qubit gate and the usual single-qubit depolarizing map with error rate $\epsilon_1 = 0.25\epsilon_2$. When computing the spectrum from exact diagonalization for comparison, the signal variances are *not* standardized, in order to preserve relative peak heights that would otherwise be enforced by a shot-noise floor.

Figure 10 highlights that the secondary peak can be resolved even under the effect of gate noise as more shots were used than in the experiment. Furthermore, Fig. 10 highlights the effect of approximate variational evolution on peaks with weak support. We observe that despite relatively small support in the $|\psi_3\rangle$ eigenstate, the associated

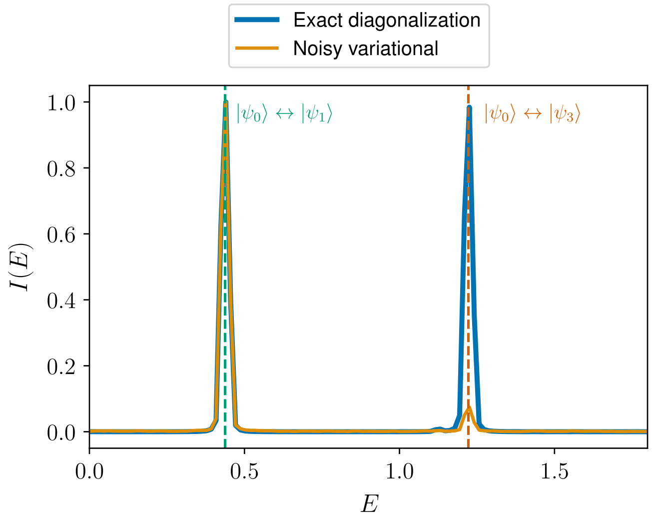

FIG. 10. Shadow spectra for the disordered Heisenberg chain obtained with dynamics via exact diagonalization (blue) and simulated noisy variational evolution (orange). The initial state has nontrivial support for $|\psi_3\rangle$, resulting in a peak corresponding to the $|\psi_0\rangle \leftrightarrow |\psi_3\rangle$ transition in the ideal spectrum. However, the algorithmic error associated with the variational evolution suppresses this peak.

ground-state transition is clearly resolved in the spectrum by our postprocessing method when computing observables using exact time evolution. However, the algorithmic error associated with approximate variational evolution strongly suppresses this peak. In Appendix H, we see that fewer shots used and stronger noise can completely eliminate these suppressed peaks.

2. Hubbard model

The time evolution between measurements is implemented using the first-order Lie-Trotter-Suzuki product formula

$$e^{-iHt} = \prod_{k=1}^{N_{\mathrm{Trott}}} \prod_{\ell=1}^{L} e^{-iH_\ell\,\delta t} + \mathcal{O}(\delta t^2).$$

We assume that in an early fault-tolerant scenario the dominant source of error are the applications of T gates and further assume that the dominant error mechanism is due to imperfect continuous rotations [115,116], which we describe via a depolarizing noise model. We thus assume a two-qubit depolarizing channel is applied for each hopping term in the Hamiltonian with probability λ as per Eq. (G6). This channel is applied after every multi-Pauli rotation of the form

$$\exp\left(-i\,\delta t\,\sigma_i^{\nu}\sigma_{i+1}^{z}\ldots\sigma_{j-1}^{z}\sigma_j^{\nu}\right), \tag{G7}$$

with $\nu \in \{x,y,z\}$. These types of rotations appear naturally from the Jordan-Wigner transformation of hopping terms. Although not directly employed in our calculations, the σ^z terms sandwiched between qubits i and j can be removed by introducing a network of Fermionic SWAP (FSWAP) gates, which only consists of local gates of depth $\mathcal{O}(N^{\frac{1}{2}})$ [117–120]. This motivates an error model whereby depolarizing noise acts on qubits i and j for each of these terms. Single-qubit rotations are burdened with the equivalent depolarizing channel for a single qubit at the same noise strength.

APPENDIX H: ADDITIONAL DETAILS OF HARDWARE DEMONSTRATION

The shadow spectrum of the six-site linear Heisenberg spin chain was quantum computed on the IBM Quantum platform, using the seven-qubit IBMQ Jakarta QPU (Falcon r5.11H processor) and the QASM simulator, both freely accessible in the Open Plan—no device reservation or queue priority privileges were used in these experiments. A set of 1000 time-evolved states at regular intervals were selected and prepared with the variational ansatz circuit and optimized classical parameters θ_t that uniquely define each time-evolved state (see details in Appendix G 1)—we do not perform variational optimization on actual hardware but use instead the parameters computed from earlier numerical simulations.

At each time step, 220 classical shadow measurements were taken; this was done by appending random basis rotations to 220 copies of an ansatz, and submitting them as one Qiskit Runtime Sampler primitive job with the option of performing only a single-shot measurement for each circuit. Other settings of the Sampler jobs were taken as defaults—specifically the resilience-level argument was set to 1 (minimal mitigation costs) and optimization level set to 3 (even heavier optimization). The 1000 Sampler primitives were further batched into ten Qiskit Sessions, where one every ten time steps were grouped into one Session (i.e., Sampler jobs indexed 0, 10, 20,…, 90 and 1, 11, 21…, 91, etc. would correspond to the first two Sessions). This took 9 days in total to complete. We highlight that while 220 shots were measured at each time step, for the spectrum in Fig. 3 we only actually used 3 × 50 randomly selected shots (split into three batches of 50 shots each for the median of means estimation from classical shadows).

Using these 1000 time-evolved classical shadows we determined expected values of all three-local Pauli strings and ordered them into a standardized data matrix $\mathbf{D}$ as detailed in Appendix D. We performed a Ljung-Box test and retained only the highest scoring 50 signals in $\mathbf{D}$. The correlation matrix $\mathbf{C}$ and its eigenvectors were then computed and the five eigenvectors with the highest eigenvalues were retained. The shadow spectrum was then computed from these five eigenvectors via the spectral cross-correlation method described in Appendix D (effectively computing Fourier transforms of the eigenvectors).

APPENDIX I: ANALYSIS OF THE EXPERIMENTAL DATA

In addition to the experimental spectrum in Fig. 3 using only 3×50 shots, here we analyze the performance of shadow spectroscopy using all 220 snapshots per time step collected from IBM Jakarta. These snapshot were again split into three batches and all up to three-local Pauli strings were estimated using classical shadows.

First, we analyze the noise model of the experimental device and compare results to noisy simulations using the theoretical noise model of IBM Jakarta by passing the FakeJakarta backend into the options of each Sampler Job run on IBM's QASM simulator (total compute time of < 0.5 hour). Recall that under global depolarizing noise at a circuit error rate ξ every Pauli string's expected value is shrunk by the same factor $\eta_{\text{exp}} \approx e^{-\xi}$ resulting in the global relation between experimental and ideal expected values $\langle P\rangle_{\text{exp}} = \eta_{\text{exp}}\langle P\rangle_{id}$. In Fig. 11 (left) we plot observable expected values and compare them to ideal and noisy simulations of IBM Jakarta: indeed the noisy observed expected values are shrunk compared to the ideal ones but not uniformly, i.e., some observables are shrunk more than others. This indeed confirms expectations that in the case of a Hamiltonian variational ansatz the global depolarizing model is only a crude approximation [111]. We estimate that the observables on average are shrunk by a factor $\eta_{\text{exp}} \approx 0.54$ in the experiment and by $\eta_{\text{sim}} \approx 0.55$ in simulation, but with a significant variability across the different observables via the standard deviations $\sigma_{\text{exp}} = 0.16$ and $\sigma_{\text{sim}} = 0.19$. The observed average shrink factor is comparable to a circuit error rate $\xi \approx -\ln \eta_{\text{exp}} \approx 0.62$, which is comparable but smaller than in our numerical simulations for the Hubbard model in Appendix K where the range $1.0 \le \xi \le 2.5$ was probed.

The observable expected values shown in Fig. 11 (left) were estimated the following way. First, an autocorrelation test (Ljung-Box test [39,40]) was performed on the time-dependent Pauli expected values and the worst-scoring 100 observables were selected, i.e., the ones that are most well described by random fluctuations around the mean. This way we can select observables whose expected values are (nearly) the same at each time step. We then estimated the time average of these constant signals. Since 1000 time steps were recorded with 200 shots at each time step, overall 2×10^4 shots were available to predict a time-constant expected value thus the level of shot noise is suppressed to a relatively low level (average standard deviation in Fig. 11 (left) is 0.01, hence no error bars are shown).

Second, we use all 220 shots per time step available from the experiment to estimate the shadow spectrum and compare to noisy and ideal simulations in Fig. 11 (right). These results clearly confirm our theoretical guarantees that reasonable levels of experimental noise do not affect the position of the peak. The intensity of the peak is, of course, affected by noise: the experimentally observed peak is shrunk by a factor 0.52, which broadly agrees with our estimated $\eta_{\text{exp}} \approx 0.54$. In contrast, the peak of the noisy simulation is shrunk by a factor 0.37, which is more significant than the estimated $\eta_{\text{sim}} \approx 0.55$ but still within range since the variability across observables was estimated $\sigma_{\text{exp}} = 0.19$, i.e., the observables that give rise to intense peaks are shrunk by a larger amount in the theoretical simulations.

Additionally, we observe a faint peak at around $E \approx 1.2$ in Fig. 11 (right) in our noise-free simulation. This faint peak is not visible in the noisy experiment (simulation) as it would require a higher number of shots to resolve—indeed this peak is analyzed in more detail in Appendix G 1 and it is confirmed that the peak can be resolved when more shots are available (4000 time steps vs 1000 time steps) even under gate noise.

APPENDIX J: SCALING ANALYSIS

1. Probing large system sizes using tensor-network techniques

Here we simulate large system sizes well beyond the capabilities of exact diagonalization to confirm that our approach is scalable as long as a sufficiently good initial state can be provided—and to verify that an increasing system size improves the performance of our approach. We consider the Heisenberg spin-chain problem with random site-dependent magnetic fields in Eq. (7) at a relatively small coupling $J = 0.01$ ($h = 1$), so that we can straightforwardly obtain a good initial state the following way. In particular, in the limit $J \to 0$ we can analytically solve the Hamiltonian and simply obtain the ground state as the bitstring $|b\rangle$. Here the jth bit b_j of the bitstring b is obtained as

$$b_j = \begin{cases} 0 & h_j < 0 \\ 1 & h_j \ge 0 \end{cases}. \tag{J1}$$

The first excited state can then be obtained by flipping the qth bit in this bitstring, where q corresponds to the absolute smallest on-site energy term $q = \arg\min_k |h_k|$. Thus, an equal superposition of the first and excited states is simply created by applying a Hadamard gate to the ground state as $H_q|b\rangle$. We note that we have intentionally chosen a system with weak coupling such that (a) a good approximate initial state can be chosen this way, (b) the system accumulates only modest bond dimension when evolving with tensor-network methods, remaining tractable and (c) the correctness of resulting peaks can be verified at system sizes too large for exact diagonalization by using perturbation theory. Given the bond dimension grows exponentially with JT, where T is the total simulation time, choosing a significantly larger coupling J

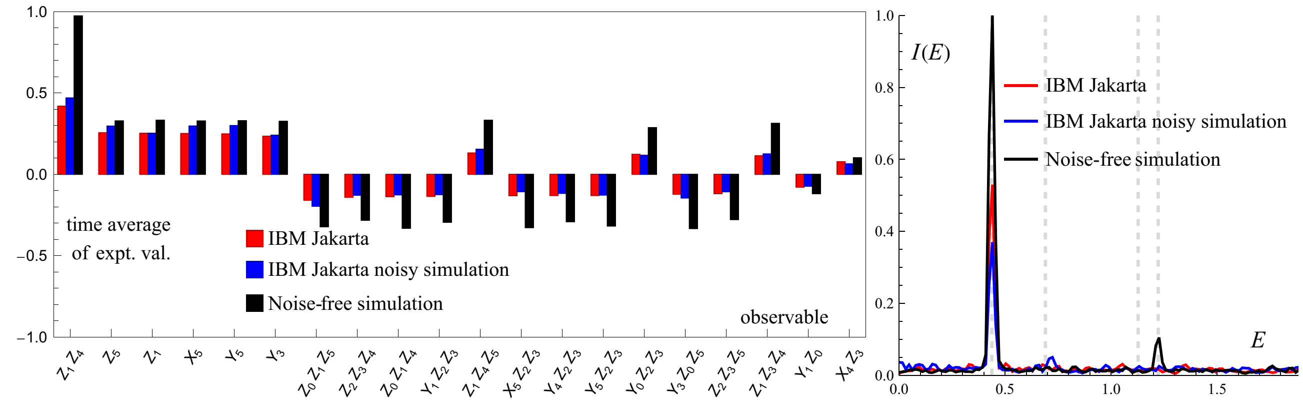

FIG. 11. (Left) Expected values of time-independent observables are estimated from classical shadows obtained from IBM Jakarta (red) and from noisy simulations of IBM Jakarta (blue). The level of shot noise in these estimates is low with an average standard deviation of 0.01 thus no error bars are shown. These expected values are shrunk compared to the ideal ones (black) by an average factor of $\eta_{\text{exp}} \approx 0.54$ in the experiment ($\eta_{\text{sim}} \approx 0.55$ in simulation), but with a significant variability across the different observables. The hardware noise is comparable to a circuit error rate $\xi \approx -\ln \eta_{\text{exp}} \approx 0.62$ which is comparable but smaller than in our simulations of the Hubbard model in Appendix K. (Right) Estimating the experimental shadow spectrum using all 220 shots per time step obtained from IBM Jakarta. The noisy simulation predicts slightly lower peak intensity but this deviation is well explained by the high variability of the shrinking of observables (see text). An additional, faint peak is visible in the noise-free simulation, which could be resolved by using more shots.

would get us into a regime where quantum computers can outperform classical computational techniques.

We simulate shadow spectroscopy by first representing the initial state $H_q|b\rangle$ as a matrix-product state (MPS) and evolve it using the time-evolving block decimation (TEBD) algorithm [121]. We perform this evolution using the TEBD implementation in the QUIMB Python package, using a Trotter step size

$$\Delta t = \left(\frac{\epsilon_{\text{Trotter}}}{T\|\mathcal{H}\|_F}\right)^{1/k}, \tag{J2}$$

where $\epsilon_{\text{Trotter}}$ is an error tolerance, T is the total evolution time, a kth-order Trotter decomposition is used, and $\|\mathcal{H}\|_F$ is the Frobenius norm of the Hamiltonian (equivalent to the sum of squares of its Pauli coefficients). We set a desired precision $\epsilon_{\text{Trotter}} = 10^{-3}$ and use $k = 4$. Furthermore, at each time step, we compress the state by performing a singular-value decomposition at each bond, and discard all contributions with low singular values, i.e., we discard singular values below the threshold $\lambda_j/\lambda_0 < \epsilon_{\text{SVD}}$, where ϵ_{SVD} is a cutoff hyperparameter and the singular values are indexed in descending order. We choose $\epsilon_{\text{SVD}} = 10^{-9}$ for our simulations.

This way we obtain a series of time-evolved MPS states as required for shadow spectroscopy, which we then use to compute time-dependent expectation values for all up-to-three-local Pauli strings through applying appropriate tensor contractions. We simulate the effects of shot noise by adding Gaussian random numbers of standard deviation $1/\sqrt{N_S}$ to all observables. We use $N_S = 100$ shots per time step, for which regime Gaussian noise is already a very good approximation of shot noise in practice. While at $N = 100$ qubits we obtain a large number of time-dependent signals, i.e., the number of up to three-local Pauli observables is 4 410 750, our postprocessing protocol still remains tractable on a laptop computer. We reject signals with $p > 0.01$ under a Ljung-Box test, and then perform the postprocessing detailed in Sec. II to generate shadow spectra. This procedure is performed for system sizes up to $n = 100$ qubits, generating the spectra depicted in Fig. 12.

We compare the observed peaks to results from exact diagonalization (only at $N = 10$) and to second-order Rayleigh-Schrödinger perturbation theory ($n \in \{25, 50, 100\}$). Apart from a small visible inaccuracy in the green dashed line (where perturbation theory is less accurate for the higher excited states) we generally find good agreement. We reiterate that the choice of a perturbative regime was deliberate to allow peak identification, i.e., for a weak coupling J, the actual values found by shadow spectroscopy are close to the perturbative prediction, but do not exactly agree. Finally, we confirm that the signal-to-noise ratio of the spectral peaks indeed *increase* with the number of qubits N, which clearly verifies the scalability of the present approach and is expected due to our theoretical results.

2. Exact time evolution

In this section we numerically analyze scaling properties of shadow spectroscopy by using a fixed input state and performing exact time evolution such that the numerically observed scaling is independent of the difficulty of

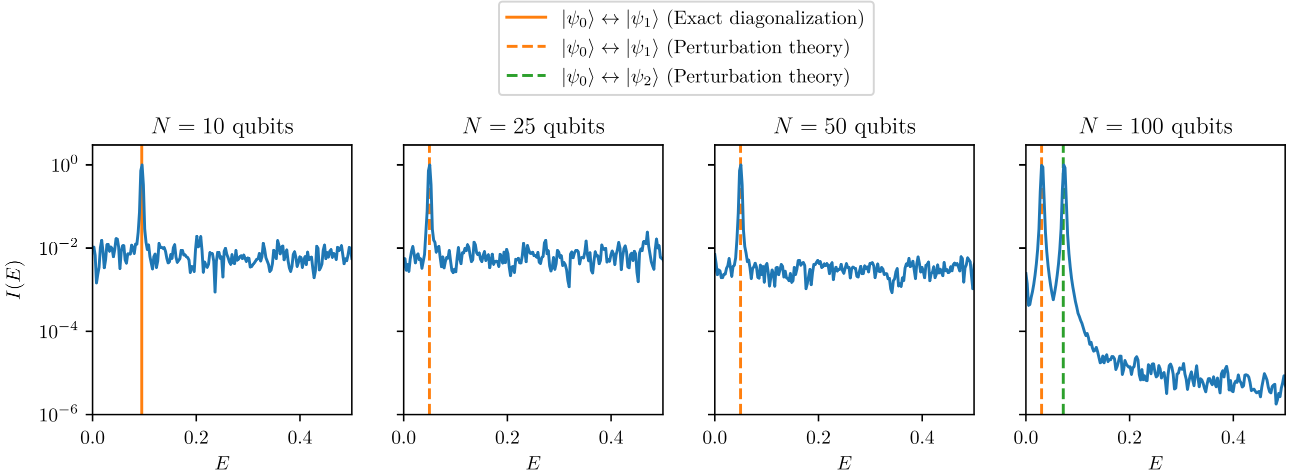

FIG. 12. Logarithmic intensity shadow spectra of a disordered Heisenberg spin chain ($J = 0.01$) up to $N = 100$ qubits verifying the scalability of our approach. A fixed input state is used, consisting of an equal-weight superposition of the ground state and first excited state of the $J = 0$ problem. Shadow spectra are computed for system sizes $n \in \{10, 25, 50, 100\}$ using synthetic shot noise equivalent to $N_S = 100$ snapshots per time step. At $N = 10$ qubits, the spectral peak matches the energy gap between the ground state and first excited state obtained from exact diagonalization (vertical solid orange line). At system sizes beyond the reach of exact diagonalization, we compare the peaks to second-order perturbation theory (vertical dashed lines), observing peaks corresponding to the gap between the ground and first excited states (orange), and at $N = 100$ an additional peak corresponding to the gap between ground and second excited states (green). Crucially, the standard deviation of the baseline decreases with an increasing number of qubits N (i.e., SNR increases with N) supporting the scalability of shadow spectroscopy—at a fixed number of snapshots N_S, peaks become *more* distinguishable from noise at larger system sizes due to the increased number of available signals.

state preparation and does not depend on algorithmic errors or gate errors either—which we analyzed separately. Here we choose the same spin problem as in our MPS simulations in the previous section (see Eq. (7)) but use an order of magnitude larger J coupling ($J = 0.1$) in order to probe regimes that are more difficult to approximate classically. We perform a series of simulations for an increasing number of qubits using an initial state with a fixed overlap with the first three eigenstates as $|\psi(0)\rangle \propto (1, \frac{1}{10}, \frac{1}{10}, 0, 0 \ldots)$, and time evolve in $N_T = 500$ time steps (using exact time evolution). We estimate local Pauli observables at each time step from classical shadows of 1000 snapshots and apply the Ljung-Box test to estimate the level of distinguishably of a signal from random noise, see Sec. II C.

In Fig. 13 (left) we plot for an increasing number N of qubits the number of statistically significant signals, i.e., signals whose Ljung-Box test returns statistical p values satisfying $p < 0.01$. Using only one-local Pauli observables results in a number of signals in Fig. 13 (left, red) that only grows as $\mathcal{O}(N^{0.4})$; this confirms that not all one-local Pauli operators yield meaningfully intense signals as expected from the entangled nature the spin ring's low-lying eigenstates. In contrast, we observe that measuring all up to three-local Pauli strings yields a significant growth of $\mathcal{O}(N^{2.5})$ of the number of statistically significant signals resulting in a large number $N_o = 1431$ of intense signals at $N = 14$ qubits. This is indeed advantageous in increasing the signal-to-noise ratio (SNR) of the spectrum but can also be used to identify a large number of observables to study physical properties of the relevant excitations.

In Fig. 13 (middle) we show the SNR of the shadow spectrum as a function of the number of observables used to produce the spectrum at $N = 14$ qubits by sorting the signals according to their statistical significance. To simplify the analysis, we use the postprocessing approach from Appendix E 1 and simply estimate the mean-squared spectrum from the most intense N_o signals—as opposed to randomly selecting observables as in Appendix E 1. This does not require additional quantum resources but only an increased classical postprocessing time: while in Sec. II C we selected randomly N_o observables from a pool and obtained a spectrum with a postprocessing complexity of $\mathcal{O}(N_o)$, here we perform a Ljung-Box test on every signal in the pool, then sort the signals according to statistical significance and calculate a plurality of spectra using an increasing number of signals.

We find that the SNR significantly improves as we increase the number of observables N_o up to an optimal point at $N_o = 407$. Beyond this point the SNR starts to decrease for an increasing N_o as the intensity of the signals with higher p values decreases. However, we note that typically one does not mind a decreased average SNR as long as the additional observables beyond the optimal point can resolve new, faint peaks. Furthermore, at the optimal point of $N_o = 407$, most of the signals are due to three-local Pauli observables, which demonstrates that it is

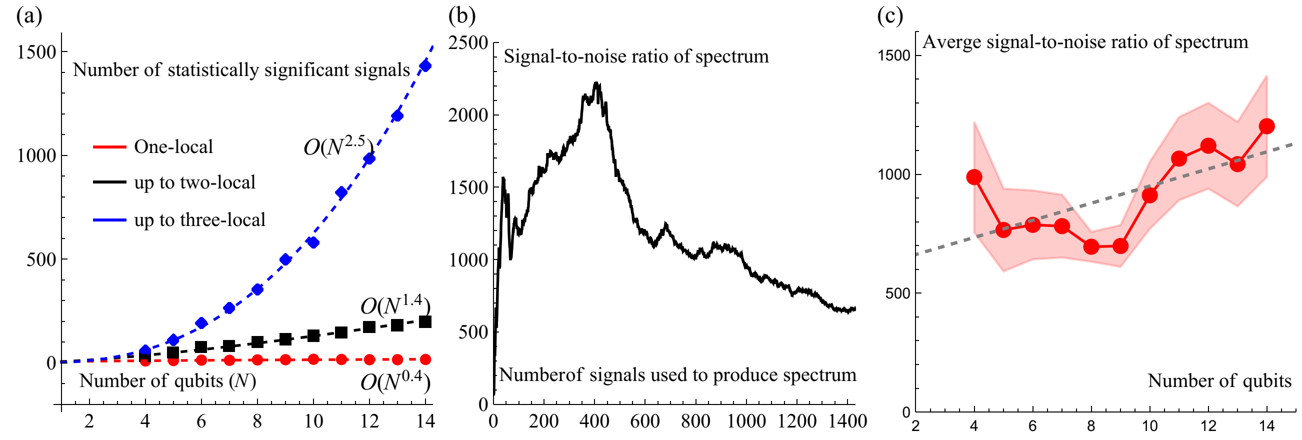

FIG. 13. Simulations of a spin problem similar to Eq. (7) using a fixed input state $|\psi(0)\rangle \propto (1, \frac{1}{10}, \frac{1}{10}, 0, 0 \ldots)$ and exact time evolution to analyze how the performance of shadow spectroscopy scales for an increasing number of qubits. (a) The number of signals with a Ljung-Box test statistical significance $p < 0.01$ using all one-local Pauli strings (red) or using up to all three-local Pauli strings (blue). (b) Using an increasing number of signals sorted according to their statistical significance at $N = 14$ qubits initially significantly increases the spectral SNR but beyond an optimal point at $N_o = 407$ the SNR decreases—note that in practice one would not mind a decreased SNR as long as the additional signals can resolve new, faint peaks. As most of the signals are due to three-local Pauli strings, it is indeed beneficial to go beyond one-local or two-local Pauli strings at the expense of only a polynomially increased classical postprocessing. (c) The average SNR (red dots) when using all statistically significant signals for an increasing qubit count. The slightly increasing SNR confirms that indeed shadow spectroscopy is efficient as long as sufficient overlap in the initial state is guaranteed—which is exponentially hard in general for arbitrary Hamiltonians.

indeed beneficial to use up to three-local Pauli observables rather than just using, e.g., two-local or one-local Pauli observables.

Finally, in Fig. 13(c) we calculate the SNR of the shadow spectrum for an increasing number of qubits. Specifically, in Fig. 13 (right, red dots) an average SNR is plotted, which is the mean of the curve in Fig. 13(b). Additionally shown is the standard deviation in Fig. 13(c, red shading), which is simply the standard deviation of the curve in Fig. 13(b). Given the constant overlap with the first three eigenstates of the spin-problem Hamiltonian—which would be exponentially difficult to guarantee for general Hamiltonians—the efficacy of shadow spectroscopy even appears to grow slightly for an increasing qubit count. This nicely verifies that shadow spectroscopy is indeed efficient given the support of the relevant excitation operators in the observables estimated from classical shadows and support of the low-lying eigenstates in the initial states.

APPENDIX K: NOISE MODEL ANALYSIS

We demonstrate our method's robustness to a variety of error models that resemble to dominant noise contributions in physical gates of near-term quantum devices. While our theoretical results in Appendix C generally guarantee robustness under the noise model in Eq. (C1), here we numerically probe error channels that go beyond Eq. (C1). We perform a set of calculations on the Fermi-Hubbard model presented in Sec. III A 2 using the following noise models.

(1) *Isotropic depolarizing noise*. Identical to the model used in the main text and described in Appendix G 2 whereby the depolarizing probability is uniform over the Bloch sphere. This model describes a scenario when different coherent and incoherent error sources are perfectly averaged out by twirling.

(2) *Anisotropic depolarizing noise*. In order to more realistically capture noise models of experimental devices that are dominated by T1 and T2 relaxation we consider an anisotropic depolarizing model whereby every single-qubit gate is followed by the noise channel of the form

$$\Phi_j^{\lambda}(\rho) = (1-\lambda)\rho + \frac{9\,\lambda}{10}\sigma_j^z\,\rho\,\sigma_j^z + \frac{\lambda}{20}\sigma_j^x\,\rho\,\sigma_j^x + \frac{\lambda}{20}\sigma_j^y\,\rho\,\sigma_j^y. \qquad \text{(K1)}$$

Above, Pauli-Z errors make up 90% of all error events, and the rest is split evenly between Pauli-X and Pauli-Y errors to reflect that the noise model is dominated by T2 relaxation. Multiqubit Pauli gadgets as in Eq. (G7) are followed by application of the single-qubit noise channel as above on each of the edge qubits i and j.

(3) *Anisotropic depolarizing and damping noise*. We now consider a noise model that violates our theoretical assumptions in Eq. (C1) and explicitly takes into account

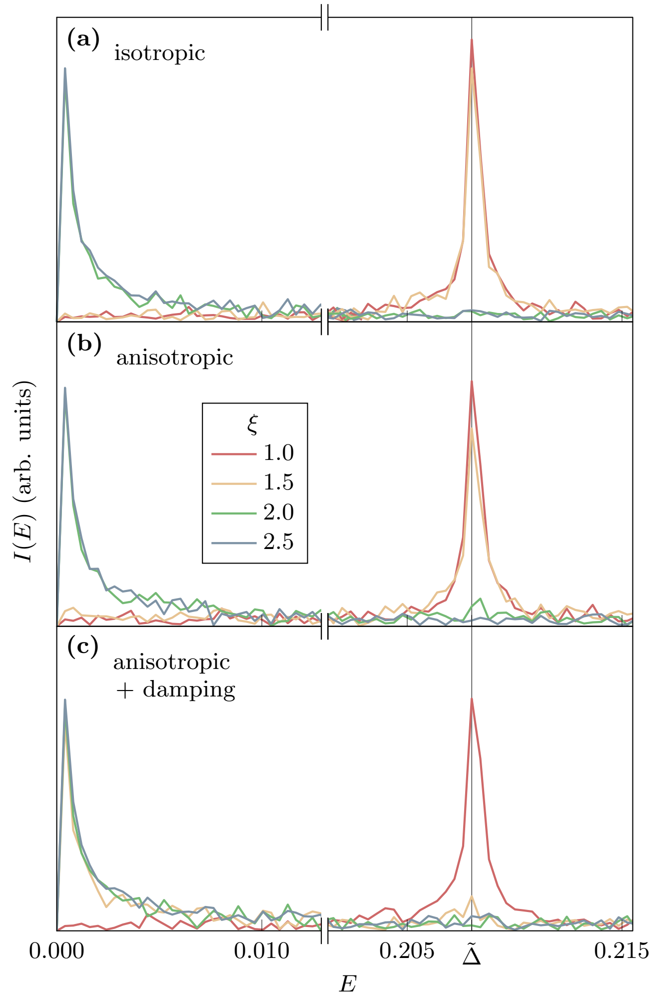

FIG. 14. Relevant regions of the energy spectra obtained for the Hubbard model using different noise models and various numbers of errors $\xi = \lambda N_{\text{gates}}$ per full circuit. Notice the interruption of the E axis. The position of the peak without noise using the same Trotter time step is marked as $\tilde{\Delta}$. The results demonstrate robustness to various types of noise. (a) Isotropic depolarizing noise, identical to that in Fig. 5. (b) Anisotropic depolarizing noise with $0.9\,\lambda$ probability of Pauli-Z errors, and $0.05\,\lambda$ probability each of a Pauli-X or a Pauli-Y error. Note that the peak at $\xi = 2$ (green line) seems to be slightly shifted to the right but this is likely just an artefact due to the peak height being comparable to additive shot noise (background noise). (c) Anisotropic depolarizing noise like in *(b)*, but each depolarizing channel is additionally followed by a damping channel with probability $\lambda/10$.

the effect of T1 relaxation: this noise model that is identical to the anisotropic depolarizing noise described above, but every application of $\Phi_j^\lambda(\rho)$ is followed by a damping channel

$$\mathcal{N}_j^\gamma(\rho) = K_0^{(j)} \rho K_0^{(j)\dagger} + K_1^{(j)} \rho K_1^{(j)\dagger}, \tag{K2}$$

with the usual damping Kraus operators

$$K_0^{(j)} = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}_j, \quad \text{and} \quad K_1^{(j)} = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}_j,$$

that act on qubit j. The probability of a damping event is 1/10 that of a depolarizing event via $\gamma = \lambda/10$ to reflect that the T1 relaxation time is typically longer than the T2 relaxation time.

Figure 14 shows the results of the simulations. All three investigated models have the main signal peak at the same position, which matches the theoretical value $\tilde{\Delta}$, indicating that shadow spectroscopy is very robust not only to isotropic and anisotropic depolarizing noise—as we expect from the arguments in Appendix C—but even to damping errors, to which Appendix C does not strictly apply.

Notice that due to the randomness of the classical shadow measurements, Fig. 14(a) is slightly different from Fig. 5 in the main text. Most notably, $\xi = 2$ produces a very faint peak in Fig. 5, but not in Fig. 14(a), illustrating that this error rate is right on the edge of the transition from recoverable signal to pure noise. The other, experimentally much more relevant noise levels, however, produce almost identical patterns. Note also that the signal in Fig. 14(c) loses intensity much earlier than (a) and (b)—already at $\xi = 1.5$, the peak is barely visible. We attribute this to the fact that there is an extra contribution of $\lambda/10$ to the noise in these simulations, which ξ does not account for. But, the important feature of unchanged peak position for the case where noise is present also holds in this case.

[1] C. Kokail, C. Maier, R. van Bijnen, T. Brydges, M. K. Joshi, P. Jurcevic, C. A. Muschik, P. Silvi, R. Blatt, C. F. Roos, and P. Zoller, Self-verifying variational quantum simulation of lattice models, Nature **569**, 355 (2019).
[2] D. Jafferis, A. Zlokapa, J. D. Lykken, D. K. Kolchmeyer, S. I. Davis, N. Lauk, H. Neven, and M. Spiropulu, Traversable wormhole dynamics on a quantum processor, Nature **612**, 51 (2022).
[3] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. Sawaya, *et al.*, Quantum chemistry in the age of quantum computing, Chem. Rev. **119**, 10856 (2019).
[4] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, Rev. Mod. Phys. **92**, 015003 (2020).
[5] B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan, Quantum algorithms for quantum chemistry and quantum materials science, Chem. Rev. **120**, 12685 (2020).
[6] M. Motta and J. E. Rice, Emerging quantum computing algorithms for quantum chemistry, WIREs Comput. Mol. Sci. **12**, e1580 (2022).
[7] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders, Efficient quantum algorithms for simulating sparse hamiltonians, Commun. Math. Phys. **270**, 359 (2007).

[8] G. H. Low and I. L. Chuang, Optimal hamiltonian simulation by quantum signal processing, Phys. Rev. Lett. **118**, 010501 (2017).

[9] G. H. Low and I. L. Chuang, Hamiltonian simulation by qubitization, Quantum **3**, 163 (2019).

[10] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, Simulating hamiltonian dynamics with a truncated Taylor series, Phys. Rev. Lett. **114**, 090502 (2015).

[11] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. Van Den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, *et al.*, Evidence for the utility of quantum computing before fault tolerance, Nature **618**, 500 (2023).

[12] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien, Quantum error mitigation, Rev. Mod. Phys. **95**, 045005 (2023).

[13] B. Koczor, Exponential error suppression for near-term quantum devices, Phys. Rev. X **11**, 031057 (2021).

[14] G. Boyd and B. Koczor, Training variational quantum circuits with covar: Covariance root finding with classical shadows, Phys. Rev. X **12**, 041022 (2022).

[15] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, arXiv:1411.4028.

[16] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 1 (2014).

[17] S. Endo, Z. Cai, S. C. Benjamin, and X. Yuan, Hybrid quantum-classical algorithms and quantum error mitigation, J. Phys. Soc. Jpn. **90**, 032001 (2021).

[18] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, Nat. Rev. Phys. **3**, 625 (2021).

[19] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, Noisy intermediate-scale quantum algorithms, Rev. Mod. Phys. **94**, 015004 (2022).

[20] Y. Li and S. C. Benjamin, Efficient variational quantum simulator incorporating active error minimization, Phys. Rev. X **7**, 021050 (2017).

[21] M.-C. Chen, M. Gong, X. Xu, X. Yuan, J.-W. Wang, C. Wang, C. Ying, J. Lin, Y. Xu, Y. Wu, S. Wang, H. Deng, F. Liang, C.-Z. Peng, S. C. Benjamin, X. Zhu, C.-Y. Lu, and J.-W. Pan, Demonstration of adiabatic variational quantum computing with a superconducting quantum coprocessor, Phys. Rev. Lett. **125**, 180501 (2020).

[22] L. Tarruell and L. Sanchez-Palencia, Quantum simulation of the Hubbard model with ultracold fermions in optical lattices, arXiv:1809.00571.

[23] R. A. Hart, P. M. Duarte, T.-L. Yang, X. Liu, T. Paiva, E. Khatami, R. T. Scalettar, N. Trivedi, D. A. Huse, and R. G. Hulet, Observation of antiferromagnetic correlations in the Hubbard model with ultracold atoms, Nature **519**, 211 (2015).

[24] U. Schneider, L. Hackermüller, J. P. Ronzheimer, S. Will, S. Braun, T. Best, I. Bloch, E. Demler, S. Mandt, D. Rasch, and A. Rosch, Fermionic transport and out-of-equilibrium dynamics in a homogeneous Hubbard model with ultracold atoms, Nat. Phys. **8**, 213 (2012).

[25] J. Singleton, *Band Theory and Electronic Properties of Solids*, Oxford master series in condensed matter physics (Oxford University Press, 2001). https://books.google.co.uk/books?id=xYZgngEACAAJ.

[26] L. González and R. Lindh, *Quantum Chemistry and Dynamics of Excited States: Methods and Applications* (Wiley, Chichester, UK, 2021). https://books.google.co.uk/books?id=TLxDnQAACAAJ.

[27] S. Lee, J. Lee, H. Zhai, Y. Tong, A. M. Dalzell, A. Kumar, P. Helms, J. Gray, Z.-H. Cui, W. Liu, M. Kastoryano, R. Babbush, J. Preskill, D. R. Reichman, E. T. Campbell, E. F. Valeev, L. Lin, and G. K.-L. Chan, Evaluating the evidence for exponential quantum advantage in ground-state quantum chemistry, Nat. Commun. **14**, 1952 (2023).

[28] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nat. Phys. **16**, 1050 (2020).

[29] H.-Y. Huang, R. Kueng, and J. Preskill, Efficient estimation of Pauli observables by derandomization, Phys. Rev. Lett. **127**, 030503 (2021).

[30] K. Wan, W. J. Huggins, J. Lee, and R. Babbush, Matchgate shadows for fermionic quantum simulation, Commun. Math. Phys. **404**, 629 (2023).

[31] A. Zhao, N. C. Rubin, and A. Miyake, Fermionic partial tomography via classical shadows, Phys. Rev. Lett. **127**, 110504 (2021).

[32] A. A. Akhtar, H.-Y. Hu, and Y.-Z. You, Scalable and flexible classical shadow tomography with tensor networks, Quantum **7**, 1026 (2023).

[33] C. Bertoni, J. Haferkamp, M. Hinsche, M. Ioannou, J. Eisert, and H. Pashayan, Shallow shadows: Expectation estimation using low-depth random Clifford circuits, Phys. Rev. Lett. **133**, 020602 (2024).

[34] S. Endo, Q. Zhao, Y. Li, S. Benjamin, and X. Yuan, Mitigating algorithmic errors in a hamiltonian simulation, Phys. Rev. A **99**, 012334 (2019).

[35] Y. Gu, Y. Ma, N. Forcellini, and D. E. Liu, Noise-resilient phase estimation with randomized compiling, Phys. Rev. Lett. **130**, 250601 (2023).

[36] K. P. Gnatenko, H. Laba, and V. Tkachuk, Energy levels estimation on a quantum computer by evolution of a physical quantity, Phys. Lett. A **424**, 127843 (2022).

[37] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, and W. A. de Jong, Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states, Phys. Rev. A **95**, 042308 (2017).

[38] T. Jones and S. Benjamin, Questlink–mathematica embiggened by a hardware-optimised quantum emulator, Quantum Sci. Technol. **5**, 034012 (2020).

[39] G. E. Box and D. A. Pierce, Distribution of residual autocorrelations in autoregressive-integrated moving average time series models, J. Am. Stat. Assoc. **65**, 1509 (1970).

[40] G. M. Ljung and G. E. Box, On a measure of lack of fit in time series models, Biometrika **65**, 297 (1978).

[41] M. H. Hayes, *Statistical Digital Signal Processing and Modeling* (John Wiley & Sons, New York, NY, US, 1996).

[42] M. Silva, E. Magesan, D. W. Kribs, and J. Emerson, Scalable protocol for identification of correctable codes, Phys. Rev. A **78**, 012347 (2008).

[43] E. Magesan, J. M. Gambetta, and J. Emerson, Characterizing quantum gates via randomized benchmarking, Phys. Rev. A **85**, 042311 (2012).

[44] Z. Cai and S. C. Benjamin, Constructing smaller Pauli twirling sets for arbitrary error channels, Sci. Rep. **9**, 1 (2019).

[45] Z. Cai, X. Xu, and S. C. Benjamin, Mitigating coherent noise using Pauli conjugation, npj Quantum Inf. **6**, 1 (2020).

[46] H. Jnane, J. Steinberg, Z. Cai, H. C. Nguyen, and B. Koczor, Quantum error mitigated classical shadows, PRX Quantum **5**, 010324 (2024).

[47] D. J. Luitz, N. Laflorencie, and F. Alet, Many-body localization edge in the random-field Heisenberg chain, Phys. Rev. B **91**, 081103 (2015).

[48] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, Toward the first quantum simulation with quantum speedup, Proc. Natl. Acad. Sci. **115**, 9456 (2018).

[49] R. Nandkishore and D. A. Huse, Many-body localization and thermalization in quantum statistical mechanics, Annu. Rev. Condens. Matter Phys. **6**, 15 (2015).

[50] S. McArdle, T. Jones, S. Endo, Y. Li, S. C. Benjamin, and X. Yuan, Variational ansatz-based quantum simulation of imaginary time evolution, npj Quantum Inf. **5**, 1 (2019).

[51] B. Koczor and S. C. Benjamin, Quantum natural gradient generalized to noisy and nonunitary circuits, Phys. Rev. A **106**, 062416 (2022).

[52] B. van Straaten and B. Koczor, Measurement cost of metric-aware variational quantum algorithms, PRX Quantum **2**, 030324 (2021).

[53] J. Haegeman, S. Michalakis, B. Nachtergaele, T. J. Osborne, N. Schuch, and F. Verstraete, Elementary excitations in gapped quantum spin systems, Phys. Rev. Lett. **111**, 080401 (2013).

[54] E. R. Anschuetz and B. T. Kiani, Quantum variational algorithms are swamped with traps, Nat. Commun. **13**, 7760 (2022).

[55] N. A. Nemkov, E. O. Kiktenko, and A. K. Fedorov, Barren plateaus are swamped with traps, arXiv:2405.05332.

[56] R. Puig, M. Drudis, S. Thanasilp, and Z. Holmes, Variational quantum simulation: A case study for understanding warm starts, PRX Quantum **6**, 010317 (2025).

[57] Hamiltonian generation and JW mapping were automatically performed using `openfermion` [122].

[58] G. Rendon, J. Watkins, and N. Wiebe, Improved accuracy for trotter simulations using Chebyshev interpolation, Quantum **8**, 1266 (2024).

[59] E. Campbell, Random compiler for fast hamiltonian simulation, Phys. Rev. Lett. **123**, 070503 (2019).

[60] Y. Ouyang, D. R. White, and E. T. Campbell, Compilation by stochastic Hamiltonian sparsification, Quantum **4**, 235 (2020).

[61] B. Koczor, The dominant eigenvector of a noisy quantum state, New J. Phys. **23**, 123047 (2021).

[62] I. Shavitt, Geometry and singlet-triplet energy gap in methylene: A critical review of experimental and theoretical determinations, Tetrahedron **41**, 1531 (1985).

[63] J. M. Foster and S. F. Boys, Quantum variational calculations for a range of ch_2 configurations, Rev. Mod. Phys. **32**, 305 (1960).

[64] H. F. Schaefer, Methylene: A paradigm for computational quantum chemistry, Science **231**, 1100 (1986). http://www.jstor.org/stable/1696782.

[65] J. F. Harrison, in *Advances in the Theory of Atomic and Molecular Systems: Conceptual and Computational Advances in Quantum Chemistry*, edited by P. Piecuch, J. Maruani, G. Delgado-Barrio, and S. Wilson (Springer Netherlands, Dordrecht, 2009), p. 33.

[66] C. D. Sherrill, M. L. Leininger, T. J. Van Huis, and H. F. Schaefer, Structures and vibrational frequencies in the full configuration interaction limit: Predictions for four electronic states of methylene using a triple-zeta plus double polarization (tz2p) basis, J. Chem. Phys. **108**, 1040 (1998).

[67] L. Veis and J. Pittner, Quantum computing applied to calculations of molecular energies: Ch2 benchmark, J. Chem. Phys. **133**, 194106 (2010).

[68] L. Veis and J. Pittner, Adiabatic state preparation study of methylene, J. Chem. Phys. **140**, 214111 (2014).

[69] K. Sugisaki, K. Toyota, K. Sato, D. Shiomi, and T. Takui, A quantum algorithm for spin chemistry: A Bayesian exchange coupling parameter calculator with broken-symmetry wave functions, Chem. Sci. **12**, 2121 (2021).

[70] Using single-qubit Pauli-X gates followed by a single application of a subspace-preserving two-qubit gate, such as the fSIM gate [123] (see, e.g., Ref. [124] for related strategies for preparing similar initial states).

[71] D. Ramirez, J. Via, and I. Santamaria, in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing* (IEEE, 2008), p. 3769.

[72] S. Malekpour, J. A. Gubner, and W. A. Sethares, Measures of generalized magnitude-squared coherence: Differences and similarities, J. Franklin. Inst. **355**, 2932 (2018).

[73] I. Santamaria and J. Via, in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, Vol. 3 (IEEE, 2007), p. III–985.

[74] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok, in *Compressed Sensing: Theory and Applications* (Cambridge University Press, Cambridge, UK, 2012), p. 1.

[75] K. M. Sherbert, N. Naimipour, H. Safavi, H. C. Shaw, and M. Soltanalian, Quantum compressive sensing: Mathematical machinery, quantum algorithms, and quantum circuitry, Appl. Sci. **12**, 7525 (2022).

[76] M. C. Tran, D. K. Mark, W. W. Ho, and S. Choi, Measuring arbitrary physical properties in analog quantum simulation, Phys. Rev. X **13**, 011049 (2023).

[77] O. Crawford, B. van Straaten, D. Wang, T. Parks, E. Campbell, and S. Brierley, Efficient quantum measurement of Pauli operators in the presence of finite sampling error, Quantum **5**, 385 (2021).

[78] T.-C. Yen, V. Verteletskyi, and A. F. Izmaylov, Measuring all compatible operators in one series of single-qubit measurements using unitary transformations, J. Chem. Theory Comput. **16**, 2400 (2020).

[79] A. Jena, S. Genin, and M. Mosca, Pauli partitioning with respect to gate sets, arXiv:1907.07859.

[80] P. Gokhale, O. Angiuli, Y. Ding, K. Gui, T. Tomesh, M. Suchara, M. Martonosi, and F. T. Chong, O(N3) measurement cost for variational quantum eigensolver on molecular hamiltonians, IEEE Trans. Quantum Eng. **1**, 1 (2020).

[81] L. Clinton, J. Bausch, J. Klassen, and T. Cubitt, Phase estimation of local Hamiltonians on NISQ hardware, New J. Phys. **25**, 033027 (2023).

[82] C. Kokail, R. van Bijnen, A. Elben, B. Vermersch, and P. Zoller, Entanglement hamiltonian tomography in quantum simulation, Nat. Phys. **17**, 936 (2021).

[83] T. V. Zache, C. Kokail, B. Sundar, and P. Zoller, Entanglement spectroscopy and probing the Li-Haldane conjecture in topological quantum matter, Quantum **6**, 702 (2022).

[84] R. Meister, pyQuEST - a Python interface for the Quantum Exact Simulation Toolkit, https://github.com/rrmeister/pyQuEST (2022).

[85] A. Richards, University of Oxford advanced research computing (2015).

[86] A. D. Bookatz, QMA-complete problems, arXiv:1212.6312.

[87] F. Jamet, C. Lenihan, L. P. Lindoy, A. Agarwal, E. Fontana, B. A. Martin, and I. Rungger, Anderson impurity solver integrating tensor network methods with quantum computing, arXiv:2304.06587.

[88] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, Nat. Commun. **9**, 1 (2018).

[89] L. Lin and Y. Tong, Heisenberg-limited ground-state energy estimation for early fault-tolerant quantum computers, PRX Quantum **3**, 010318 (2022).

[90] T. S. Cubitt, D. Perez-Garcia, and M. M. Wolf, Undecidability of the spectral gap, Nature **528**, 207 (2015).

[91] B. Yoshimura, W. Campbell, and J. Freericks, Diabatic-ramping spectroscopy of many-body excited states, Phys. Rev. A **90**, 062334 (2014).

[92] J. P. T. Stenger, G. Ben-Shach, D. Pekker, and N. T. Bronn, Simulating spectroscopy experiments with a superconducting quantum computer, Phys. Rev. Res. **4**, 043106 (2022).

[93] O. Higgott, D. Wang, and S. Brierley, Variational quantum computation of excited states, Quantum **3**, 156 (2019).

[94] H. H. S. Chan, N. Fitzpatrick, J. Segarra-Martí, M. J. Bearpark, and D. P. Tew, Molecular excited state calculations with adaptive wavefunctions on a quantum eigensolver emulation: Reducing circuit depth and separating spin states, Phys. Chem. Chem. Phys. **23**, 26438 (2021).

[95] P. J. Ollitrault, A. Kandala, C.-F. Chen, P. K. Barkoutsos, A. Mezzacapo, M. Pistoia, S. Sheldon, S. Woerner, J. M. Gambetta, and I. Tavernelli, Quantum equation of motion for computing molecular excitation energies on a noisy quantum processor, Phys. Rev. Res. **2**, 043140 (2020).

[96] A. Asthana, A. Kumar, V. Abraham, H. Grimsley, Y. Zhang, L. Cincio, S. Tretiak, P. A. Dub, S. E. Economou, E. Barnes, and N. J. Mayhall, Equation-of-motion variational quantum eigensolver method for computing molecular excitation energies, ionization potentials, and electron affinities, arXiv:2206.10502.

[97] T. Jones, S. Endo, S. McArdle, X. Yuan, and S. C. Benjamin, Variational quantum algorithms for discovering hamiltonian spectra, Phys. Rev. A **99**, 062304 (2019).

[98] K. M. Nakanishi, K. Mitarai, and K. Fujii, Subspace-search variational quantum eigensolver for excited states, Phys. Rev. Res. **1**, 033062 (2019).

[99] W. Kirby, M. Motta, and A. Mezzacapo, Exact and efficient Lanczos method on a quantum computer, Quantum **7**, 1018 (2023).

[100] M. Motta, C. Sun, A. T. Tan, M. J. O'Rourke, E. Ye, A. J. Minnich, F. G. Brandão, and G. K. Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, Nat. Phys. **16**, 205 (2020).

[101] K. Klymko, C. Mejuto-Zaera, S. J. Cotton, F. Wudarski, M. Urbanek, D. Hait, M. Head-Gordon, K. B. Whaley, J. Moussa, N. Wiebe, W. A. de Jong, and N. M. Tubman, Real-time evolution for ultracompact hamiltonian eigenstates on quantum hardware, PRX Quantum **3**, 020323 (2022).

[102] G. Wang, D. S. França, R. Zhang, S. Zhu, and P. D. Johnson, Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision, Quantum **7**, 1167 (2023).

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

[104] N. P. Bauman, H. Liu, E. J. Bylaska, S. Krishnamoorthy, G. H. Low, C. E. Granade, N. Wiebe, N. A. Baker, B. Peng, M. Roetteler, M. Troyer, and K. Kowalski, Toward quantum computing for high-energy excited states in molecular systems: Quantum phase estimations of core-level states, J. Chem. Theory Comput. **17**, 201 (2021).

[105] K. Sugisaki, C. Sakai, K. Toyota, K. Sato, D. Shiomi, and T. Takui, Bayesian phase difference estimation: A general quantum algorithm for the direct calculation of energy gaps, Phys. Chem. Chem. Phys. **23**, 20152 (2021).

[106] K. Sugisaki, C. Sakai, K. Toyota, K. Sato, D. Shiomi, and T. Takui, Quantum algorithm for full configuration interaction calculations without controlled time evolutions, J. Phys. Chem. Lett. **12**, 11085 (2021).

[107] A. E. Russo, K. M. Rudinger, B. C. Morrison, and A. D. Baczewski, Evaluating energy differences on a quantum computer with robust phase estimation, Phys. Rev. Lett. **126**, 210501 (2021).

[108] Y. Matsuzaki, H. Hakoshima, K. Sugisaki, Y. Seki, and S. Kawabata, Direct estimation of the energy gap between the ground state and excited state with quantum annealing, Jpn. J. Appl. Phys. **60**, SBBI02 (2021).

[109] I. Zintchenko and N. Wiebe, Randomized gap and amplitude estimation, Phys. Rev. A **93**, 062306 (2016).

[110] A. M. Dalzell, N. Hunter-Jones, and F. G. Brandão, Random quantum circuits transform local noise into global white noise, arXiv:2111.14907.

[111] J. Foldager and B. Koczor, Can shallow quantum circuits scramble local noise into global white noise? J. Phys. A: Math. Theor. **57**, 015306 (2024).

[112] Typically the HF solution has the dominant weight as $|c_0^{(g)}| > |c_{pq}^{(g)}|$ (even if the HF Slater determinant is not a good approximation) and thus in Eq. (F1) the first term in the expansion is the dominant term.

[113] V. Verteletskyi, T.-C. Yen, and A. F. Izmaylov, Measurement optimization in the variational quantum eigensolver using a minimum clique cover, J. Chem. Phys. **152,** 124114 (2020).

[114] X. Yuan, S. Endo, Q. Zhao, Y. Li, and S. C. Benjamin, Theory of variational quantum simulation, Quantum **3,** 191 (2019).

[115] B. Koczor, Sparse probabilistic synthesis of quantum operations, PRX Quantum **5,** 040352 (2024).

[116] B. Koczor, J. J. L. Morton, and S. C. Benjamin, Probabilistic interpolation of quantum rotation angles, Phys. Rev. Lett. **132,** 130602 (2024).

[117] Z. Cai, Resource estimation for quantum variational simulations of the Hubbard model, Phys. Rev. Appl. **14,** 014059 (2020).

[118] H. Jnane, B. Undseth, Z. Cai, S. C. Benjamin, and B. Koczor, Multicore quantum computing, Phys. Rev. Appl. **18,** 044064 (2022).

[119] F. Verstraete, J. I. Cirac, and J. I. Latorre, Quantum circuits for strongly correlated quantum systems, Phys. Rev. A **79,** 032316 (2009).

[120] I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K.-L. Chan, and R. Babbush, Quantum simulation of electronic structure with linear depth and connectivity, Phys. Rev. Lett. **120,** 110501 (2018).

[121] G. Vidal, Efficient classical simulation of slightly entangled quantum computations, Phys. Rev. Lett. **91,** 147902 (2003).

[122] J. R. McClean, *et al.*, Openfermion: The electronic structure package for quantum computers, arXiv:1710.07629.

[123] B. Foxen, *et al.*, (Google AI Quantum), Demonstrating a continuous set of two-qubit gates for near-term quantum algorithms, Phys. Rev. Lett. **125,** 120504 (2020).

[124] K. Sugisaki, S. Nakazawa, K. Toyota, K. Sato, D. Shiomi, and T. Takui, Quantum chemistry on quantum computers: A method for preparation of multiconfigurational wave functions on quantum computers without performing post-Hartree–Fock calculations, ACS Cent. Sci. **5,** 167 (2019).