


High-dimensional subspace expansion using classical shadowsGregory Boyd ^{1,2,*}, Bálint Koczor,^{2,1,3} and Zhenyu Cai ^{1,3,†}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 15 July 2024; revised 27 January 2025; accepted 29 January 2025; published 13 February 2025)

We introduce a postprocessing technique for classical shadow measurement data that enhances the precision of ground state estimation through high-dimensional subspace expansion; the dimensionality is only limited by the amount of classical postprocessing resources rather than by quantum resources. Crucial steps of our approach are the efficient identification of useful observables from shadow data, followed by our regularized subspace expansion that is designed to be numerically stable even when using noisy data. We analytically investigate noise propagation within our method, and upper bound the statistical fluctuations due to the limited number of snapshots in classical shadows. In numerical simulations, our method can achieve a reduction in the energy estimation errors in many cases, sometimes by more than an order of magnitude. We also demonstrate that our performance improvements are robust against both coherent errors (bad initial state) and gate noise in the state-preparation circuits. Furthermore, performance is guaranteed to be at least as good—and in many cases better—than direct energy estimation without using additional quantum resources, and the approach is thus a very natural alternative for estimating ground state energies directly from classical shadow data.

DOI: [10.1103/PhysRevA.111.022423](https://doi.org/10.1103/PhysRevA.111.022423)**I. INTRODUCTION**

Quantum computation often relies on the postprocessing of measurement results from a quantum device to produce meaningful results, either for reducing the load on quantum hardware by offloading computation that can be performed classically [1–3], for mitigating errors in noisy computations [4], or for both [5–8]. Classical shadow techniques [9] can output a rich set of classical data that is amenable toward various postprocessing techniques for efficiently probing many properties of the quantum state, offering the ability to measure first, ask questions later. Its versatility has been employed for a wide range of important tasks in quantum information [10–15].

Ground state energy estimation is one of the most fundamental tasks in quantum simulation [16,17] and is also more widely applicable in optimization problems [18]. One typically aims to prepare a good approximation to the ground state through a state-preparation circuit, from which one can directly estimate its energy as the expected value of the problem Hamiltonian. However, this often incurs an estimation error due to constraints in the size of the state preparation circuit and/or noise in the circuit. In this paper, we present a

postprocessing method using existing classical shadow measurement data to produce more accurate ground state energy estimation by performing subspace expansions [19–22] of unprecedented dimension, mitigating the constraints due to both the circuit size and the noise in the shadow data. Given that our approach is completely performed in postprocessing, it does not require any additional quantum resources and circuit runs, and it is guaranteed to output a result that is at least as good as the original value, with the possibility of achieving orders of magnitude performance improvement, especially in the cases with a highly mismatched starting state and/or a large amount of shadow data.

Our main contributions are a set of techniques for extracting relevant data from classical shadows and then performing numerically stable subspace expansions of high dimension (up to 10 000) using noisy data. We investigate the performance of such high dimensional subspace expansions, using simulated classical shadows of up to 14 qubits as well as simulations of the procedure with gate noise, demonstrating the effectiveness of our scheme as an error mitigation method. We also provide an analysis of noise propagation in the method as well as a derivation of the variances of the observables due to the finite number of snapshots from the classical shadow.

In Sec. II, we provide background on the two main components of our paper, subspace expansion, and classical shadows, Section III provides details on our shadow subspace expansion procedure and how we perform the numerical postprocessing to ensure a well-conditioned eigenvalue problem. In Sec. III B, we present our argument for the propagation of shot noise from the calculated matrix elements onto the rest of the computation, and in Sec. IV we present results of the numerical simulations of our method, investigating the effect

*Contact author: gregory.boyd@materials.ox.ac.uk†Contact author: cai.zhenyu.physics@gmail.com

of shot noise from classical shadows as well as simulated gate noise.

II. BACKGROUND

Here we first introduce the basic concepts that will form the core of our approach, such as subspace expansion using quantum computers and classical shadows for extracting a large number of expected values.

A. Subspace expansion

In many applications of quantum computers (e.g., variational eigensolvers), the aim is to directly prepare a quantum state that minimizes a certain cost function. Let us consider the case where the cost function is the energy of the state with respect to a particular problem Hamiltonian H_{tot} , which means that we are trying to prepare the corresponding ground state ρ_0 . It is often the case that we cannot prepare the exact ground state ρ_0 due to limitations in qubit overhead, gate overhead, or noise, and the best approximation of the ground state is some—potentially noisy—state that we denote as ρ . Quantum subspace expansion is a way to effectively improve on the observable expectation values obtained from ρ by performing additional measurements and postprocessing.

In the case when ρ is a pure state $|\psi\rangle\langle\psi|$, we can improve upon the direct energy estimation $\langle\psi|H|\psi\rangle$ by constructing a subspace spanned by the set of states $\{G_i|\psi\rangle \mid 1 \leq i \leq N_G\}$, which are states obtained by applying a set of *expansion basis* operators $\{G_i\}$ to $|\psi\rangle$. We will then try to find the state within this subspace that minimizes the energy. Note that I is always an element of $\{G_i\}$ such that our starting state $|\psi\rangle$ is also a state within this subspace, thus the final state we obtain, by definition, has energy lower than or equal to that of $|\psi\rangle$. Any quantum state in this subspace can simply be written as

$$|\psi_{\bar{w}}\rangle = \frac{\sum_{i=1}^{N_G} w_i G_i |\psi\rangle}{\left\| \sum_{i=1}^{N_G} w_i G_i |\psi\rangle \right\|} = \frac{\Gamma_{\bar{w}} |\psi\rangle}{\|\Gamma_{\bar{w}} |\psi\rangle\|},$$

where $\Gamma_{\bar{w}} = \sum_{i=1}^{N_G} w_i G_i$ is the *expansion operator* which is parametrized by the weight vector \bar{w} . Different states $|\psi_{\bar{w}}\rangle$ in the subspace simply correspond to different weight vectors \bar{w} . For simplicity, in this paper, we will assume that the set of operators $\{G_i\}$ are Hermitian, and the weight vectors \bar{w} are real, which implies that $\Gamma_{\bar{w}}$ is also Hermitian.

The weight vector (and the corresponding state) of minimal energy within this subspace is given as

$$\bar{w}^* = \arg \min_{\bar{w}} \langle \psi_{\bar{w}} | H_{\text{tot}} | \psi_{\bar{w}} \rangle. \quad (1)$$

Applying the same argument to the case where our starting state ρ is a mixed state, the subspace-expanded state is then given as

$$\rho_{\bar{w}} = \frac{\Gamma_{\bar{w}} \rho \Gamma_{\bar{w}}}{\text{Tr}(\Gamma_{\bar{w}}^2 \rho)},$$

and the optimization problem becomes

$$\bar{w}^* = \arg \min_{\bar{w}} \text{Tr}(H_{\text{tot}} \rho_{\bar{w}}). \quad (2)$$

To obtain the weights \bar{w} that extremize the energy $\text{Tr}(H_{\text{tot}} \rho_{\bar{w}})$, we need to solve the generalized eigenvalue

problem:

$$\mathbf{H} \bar{w} = E_{\bar{w}} \mathbf{S} \bar{w}, \quad (3)$$

where

$$\mathbf{H}_{ij} = \text{Tr}(G_i H_{\text{tot}} G_j \rho), \quad (4)$$

$$\mathbf{S}_{ij} = \text{Tr}(G_i G_j \rho). \quad (5)$$

If it is the ground state energy that we are interested in, then we obtain an improved ground state energy estimate $E_{\bar{w}^*}$ as the eigenvalue when we solve the equation above. If we are interested in the expectation value of some other observable O with respect to the ground state, then we can measure the following matrix elements:

$$\mathbf{O}_{ij} = \text{Tr}(G_i O G_j \rho)$$

for different i, j and then reconstruct the improved observable expectation value using

$$\begin{aligned} \text{Tr}(O \rho_{\bar{w}^*}) &= \frac{1}{\text{Tr}(\Gamma_{\bar{w}^*}^2 \rho)} \sum_{i,j=1}^{N_G} w_i^* \text{Tr}(G_i O G_j \rho) w_j \\ &= \frac{\bar{w}^{*\dagger} \mathbf{O} \bar{w}^*}{\text{Tr}(\Gamma_{\bar{w}^*}^2 \rho)}. \end{aligned}$$

B. Shadow tomography

We see that the expansion procedure requires us to measure the matrix elements \mathbf{H}_{ij} , \mathbf{S}_{ij} , and \mathbf{O}_{ij} , which is $\mathcal{O}(N_G^2)$ observables that we need to measure. It is also often the case that the Hamiltonian cannot be directly measured and needs to be broken down to N_H subterms, which means that the total number of observables that we need to measure is $\mathcal{O}(N_G^2 N_H)$. Therefore, it can be costly to further increase the dimension of the expansion space N_G to achieve better subspace expansion results. One way to efficiently measure multiple observables is to use shadow tomography [9,23]. To construct a classical shadow of a quantum state, we repeat the following process: we sample a random unitary U_i from a distribution of unitaries \mathbb{U} , apply it to the quantum state, and then measure the resulting state in the computational basis. These steps combined can be described as the *process channel* \mathcal{M} . For certain choices of \mathbb{U} , we are able to efficiently invert the channel from the output bit string $|\mathbf{b}\rangle$ classically to obtain a classical snapshot $\rho_i = \tilde{\mathcal{M}}(U_i^\dagger |\mathbf{b}\rangle\langle\mathbf{b}| U_i)$, where $\tilde{\mathcal{M}}$ is the classical inversion of the process channel, and then use these snapshots to estimate observables in postprocessing.

In this paper, we primarily consider the case of Pauli shadows (although, given the Hamiltonian, it may be more convenient to express it in another basis, e.g., fermionic operators, where alternative classical shadows can be used instead [24]), where \mathbb{U} consists of all the tensor products of single-qubit Clifford operators. Using this Pauli shadow scheme, the sample complexity of obtaining estimators of M Pauli operators of weight l to error ϵ is $\mathcal{O}(3^l \log(M)/\epsilon^2)$ [9], meaning that, provided both the operators in set $\{G_i\}$ and the terms in the Hamiltonian are local, we can efficiently determine many expectation values to perform high-dimensional subspace expansion. Note that the $\log(M)$ factor above exists to make sure *all* M estimators reach the stated precision ϵ . In our case, our

end goal is not estimating these M observables individually; instead they are used for further postprocessing to predict one single observable of interest, thus this $\log(M)$ factor is not applicable to our case.

III. METHODS

We now set out the methodology for shadow subspace expansion (SSE) and some practical considerations in performing large subspace expansions in postprocessing using classical shadow data. While the present approach is immediately compatible with more advanced shadow protocols like shallow shadows [25–27] or fermionic shadows [24,28], for simplicity, we will demonstrate and explain our results on the example of local Pauli shadows detailed above—which is very efficient for simultaneously estimating low-weight Pauli observables.

For this reason, we construct the expansion operators $\Gamma_{\bar{w}^*}$ as linear combinations of local Pauli strings. The central assumption of subspace expansion is that the expansion basis will map the starting state ρ into another basis state in the subspace that has significant components orthogonal to the starting state, increasing the effective dimension of our subspace. For our special case in which the expansion basis is the set of local Pauli strings, this can be especially effective in many practical scenarios as discussed below and demonstrated later in our numerical experiments.

One such application scenario is the translation invariant, gapped spin problems, whereby elementary excitation operators are supported on local operators [29] (with an exponentially decreasing error as the locality increases). As such, given a jump operator that maps the ground state to the first excited state as $G|0\rangle = |1\rangle$ and vice versa, any low-energy eigenstate of the form $|\psi\rangle = \sqrt{1-\epsilon}|0\rangle + \sqrt{\epsilon}|1\rangle$ can be mapped to the ground state as $|0\rangle = \Gamma_{\bar{w}^*}|\psi\rangle$ with $\Gamma_{\bar{w}^*} = \sqrt{1-\epsilon}\mathbb{1} - \sqrt{\epsilon}G$ and some suitable normalization. Indeed, for more complex initial states, one potentially needs significantly more expansion terms, however, a low-energy initialization guarantees high overlap with the low-lying eigenstates.

Furthermore, in fermionic systems the ground state $|0\rangle$ is typically obtained as an expansion of the HF state $|\psi_{\text{HF}}\rangle$ into low-weight fermionic annihilation-creation operators. Reference [10] details that after encoding the fermionic problem as a qubit Hamiltonian via the Jordan-Wigner transform, this expansion has a significant overlap with local Pauli operators despite the fermionic operators being mapped to nonlocal Pauli operators.

A. Shadow subspace expansion procedure

We are considering the setting in which we want to estimate some property O_{tot} of the ground state ρ_0 of some Hamiltonian H_{tot} , with their Pauli decompositions being $O_{\text{tot}} = \sum_{k=1}^{N_O} O_k$ and $H_{\text{tot}} = \sum_{k=1}^{N_H} H_k$, respectively. This can be done by preparing ρ_0 and obtaining a set of its classical shadows of size N_s for the estimations of various properties through postprocessing. However, due to noise in the circuit and/or imperfect ansatz circuit, the state we prepare and the shadow data we obtain are from some other state ρ instead.

We will now outline the way to obtain an improved estimate of $\text{Tr}(O_{\text{tot}}\rho_0)$ using the shadow data of ρ .

(1) Initial expansion basis: Select an initial expansion basis $\{G_i\}$ of size N_G . One way is to use the set of all Pauli strings up to a certain weight, where the weight is determined by the amount of shadow data (the more shadow data, the higher the weight we can reach).

(2) Filtering expansion operators: For each expansion basis operator G_i , we will perform a local subspace expansion of dimension 2 using the expansion basis $\{I, G_i\}$ and the corresponding improvement in energy is denoted as ΔE_i . This involves determining the expected values of the observables $\{G_i\}$, $\{H_k\}$, $\{G_i H_k\}$, and $\{G_i H_k G_i\}$ [this final group are just the H_k with an added phase as the G_i can be (anti)commuted through the H_k]. These observables have weights up to $M_G + M_H$, and overall the total number of expectation values to be determined from the measurement data is $N_G + N_H + N_G N_H$.

To filter out expansion basis operators that are unable to decrease the energy of the initial state, we will rank all operators in $\{G_i\}$ according to their improvement in energy ΔE_i and keep only the top K operators, which gives us the K -significant expansion basis. This process has a classical computational complexity linear in N_G of $O(N_G N_H)$; thus, N_G can be chosen as potentially very large, e.g., $N_G \gtrsim 10^6$.

(3) Equation for full expansion: We will perform subspace expansion using the K -significant expansion basis, which involves estimating the set of weight- $2M_G$ observables $\{G_i G_j\}$ to form the matrix S , and weight- $2M_G + M_H$ observables $\{G_i H_k G_j\}$ to form the matrix H , the matrix entries can be determined based on the observation that any product of local Pauli strings evaluates to a single Pauli string of higher locality [11]. They require the estimations of K^2 and $K^2 N_H$ observables, respectively.

(4) Solving the eigenvalue problem: To obtain the eigenspectrum in the subspace and the optimal weights \bar{w} . We detail our method in Sec. III C for solving the noisy generalized eigenvalue problem in Eq. (3). This step has a classical computational complexity of $O(K^p)$ with a polynomial power $2 \leq p \leq 3$ and is thus the time bottleneck. This reflects the importance of the screening step for choosing the best K observables.

Step 4 will output the energy of the state found via subspace expansion. The following steps are needed for estimating other observables on the output state:

(5) Use the shadows to estimate the set of weight- $2M_G + M_O$ observable $\{G_i O_k G_j\}$ of a size $K^2 N_O$.

(6) Use the optimal weight \bar{w}^* to reconstruct the optimal observable:

$$\begin{aligned} \text{Tr}(O_{\text{tot}}\rho_{\bar{w}^*}) &= \frac{\text{Tr}(O_{\text{tot}}\Gamma_{\bar{w}^*}\rho\Gamma_{\bar{w}^*})}{\text{Tr}(\Gamma_{\bar{w}^*}\rho\Gamma_{\bar{w}^*})} \\ &= \frac{\sum_{i,j} \beta_k w_i^* w_j^* \text{Tr}(G_i O_k G_j \rho)}{\text{Tr}(\Gamma_{\bar{w}^*}^2 \rho)}. \end{aligned}$$

We note that there are alternative methods for choosing the K -significant observables and we have described one such method based on shadow spectroscopy [10] in Appendix B.

B. Shot noise propagation

As explained in the previous section, we proceed by estimating matrix entries of \mathbf{S} and \mathbf{H} from classical shadows data. Due to the finite sample size, the matrix entries come with a certain amount of shot noise. Building on the main theorem of Ref. [30], we bound how this shot noise propagates into the solution of the generalized subspace expansion, which helps us pinpoint the hyperparameters that require particular attention in the postprocessing step.

We can recast the generalized eigenvalue equation in Eq. (3) into the usual eigenvalue equation $\mathbf{M}\bar{w} = E_{\bar{w}}\bar{w}$, where $\mathbf{M} = \mathbf{S}^{-1}\mathbf{H}$. We quantify the shot noise in the matrix \mathbf{M} as the total variance $\epsilon_M^2 = \sum_{kl} \text{Var}\{\mathbf{M}_{kl}\}$, which is equivalent to the expected Hilbert-Schmidt distance of the matrix due to shot noise. We obtain the following bound in Appendix C as

$$\epsilon_M^2 \leq N_s^{-1} 3^{2w'} K \left[\|H_{\text{tot}}\|_{\infty} \|\mathbf{S}^{-1}\|_2^4 + 3^w \|\mathbf{S}^{-1}\|_2^2 \right]. \quad (6)$$

While the above bound is based on worst-case arguments and is thus expected to be very pessimistic, it nevertheless informs us of the following. First, the total variance can be reduced inversely proportionally by increasing the number of shots N_s (standard shot noise scaling). Second, as we restrict ourselves to simple Pauli shadows [9], we have an exponential dependence on the largest weight w' among the expansion operators G_k and on the largest weight w among the Pauli terms in the total Hamiltonian H_{tot} —of course, more advanced shadow techniques [27,31,32] can help mitigate this dependence. Third, the effect of shot noise depends polynomially on the Hilbert-Schmidt norm of the inverse matrix $\|\mathbf{S}^{-1}\|_2$, and thus truncation to the most significant singular values, as we detail below in Sec. III C, is essential to ensure this matrix norm is well controlled. Finally, the variance is also influenced by the absolute largest eigenvalue $\|H_{\text{tot}}\|_{\infty}$ of the problem Hamiltonian H_{tot} , which is indeed expected to grow polynomially with the system size, i.e., the number of qubits, in typical scenarios.

As we detail in Appendix E, our conclusion is consistent with the analysis of Ref. [33]. In particular, Ref. [33] assumed that in a single experiment the deviation due to shot noise from the ideal matrices \mathbf{H} and \mathbf{S} are expressed as $\delta\mathbf{H}$ and $\delta\mathbf{S}$, respectively. Then, the deviation in the solution of the generalized eigenvalue equation can be bounded as $\delta E_{\bar{w}} \leq \|\mathbf{S}^{-1}\|_{\infty} (\|\delta\mathbf{H}\|_{\infty} + \|H_{\text{tot}}\|_{\infty} \|\delta\mathbf{S}\|_{\infty})$. Indeed, the same spectral properties appear in this bound, however, Eq. (6) expresses a statistical variance due to propagation as opposed to the noise propagation in a single sample—hence Eq. (6) explicitly depends on further parameters, such as the number of shots.

C. Regularization of the generalized eigenvalue problem

Classical shadows allow us to efficiently estimate a very large number of expansion operators, however, many states obtained through low-weight expansion operators have high overlaps. As a result, our generalized eigenvalue equation potentially involves highly singular matrices. As detailed in the previous section, shot noise propagation is potentially magnified by these small singular values of the overlap matrix. Therefore, our approach requires careful regularization.

To numerically solve the generalized eigenvalue problem in Eq. (3), we must have a well-conditioned \mathbf{S} . In our case, the physical conditions also require \mathbf{S} to be positive semidefinite. Therefore, we can regularize the problem by transforming the matrices \mathbf{H} , \mathbf{S} to a subspace in the state expansion space where \mathbf{S} has sufficiently large positive eigenvalues. This is analogous to the process of projecting onto the nearest (in terms of Frobenius norm) positive semidefinite matrix [34], except we also dynamically choose a threshold to remove some eigenvalues above 0 to ensure the eigenvalue problem is well-conditioned. More exactly, the regularization is performed by restricting the problem to the subspace spanned by the first \tilde{K} eigenvectors of \mathbf{S} ranked by eigenvalues, which is in a sense the effective dimension of the matrix containing information that is distinguishable above the noise. These eigenvectors produce a $K \times \tilde{K}$ transformation matrix \mathbf{Q} , which can be used to transform the original problem to a dimension- \tilde{K} generalized eigenvalue problem with the following matrices:

$$\tilde{\mathbf{S}} = \mathbf{Q}^{\dagger} \mathbf{S} \mathbf{Q} \quad \text{and} \quad \tilde{\mathbf{H}} = \mathbf{Q}^{\dagger} \mathbf{H} \mathbf{Q}, \quad (7)$$

which outputs the energy $E_0^{\tilde{K}}$.

The optimal value of \tilde{K} to be used can be determined from a series of such transformations by calculating the ground state energy E_0^l for each and stopping when the energies become unstable. This is determined by tracking the variance of the energy estimates within a moving window and terminating when the moving variance reaches a minimum, indicating a series of gradually varying energies as the regularization reaches the optimal value before the eigenvalue problem becomes unstable, resulting in an increase in the moving variance.

For example, given a sequence of ground state energies E_0^l , we calculate the differences $\Delta_l = E_0^l - E_0^{l-1}$, and then the moving variance with a window size W :

$$V_l = \text{Var}[\Delta_l, \dots, \Delta_{l+W}]. \quad (8)$$

The minimum of this moving variance provides a good point to stop increasing l where a low energy has been found, but before the eigenvalue problem has become unstable. In our numerics in this paper, we use a window size $W = 7$.

IV. RESULTS

We will first examine the use of SSE for solving the ground state energy for a 14-qubit spin chain model. The spin chain Hamiltonian takes the form

$$\mathcal{H} = J \sum_{i=1}^N \bar{\sigma}_i \cdot \bar{\sigma}_{i+1} + c_i Z_i, \quad (9)$$

with $J = 0.1$ and $c_i \in [-1, 1]$ as uniformly random on-site energies. This Hamiltonian has been considered in studies of self-thermalization and many-body localization and has been identified as a promising target for quantum advantage [35].

For the sake of demonstration, we initialize using the variational quantum eigensolver (VQE) [2] for 300 steps to an initial state with energy error $\Delta E = 0.235$ from the ground state, and then perform SSE.

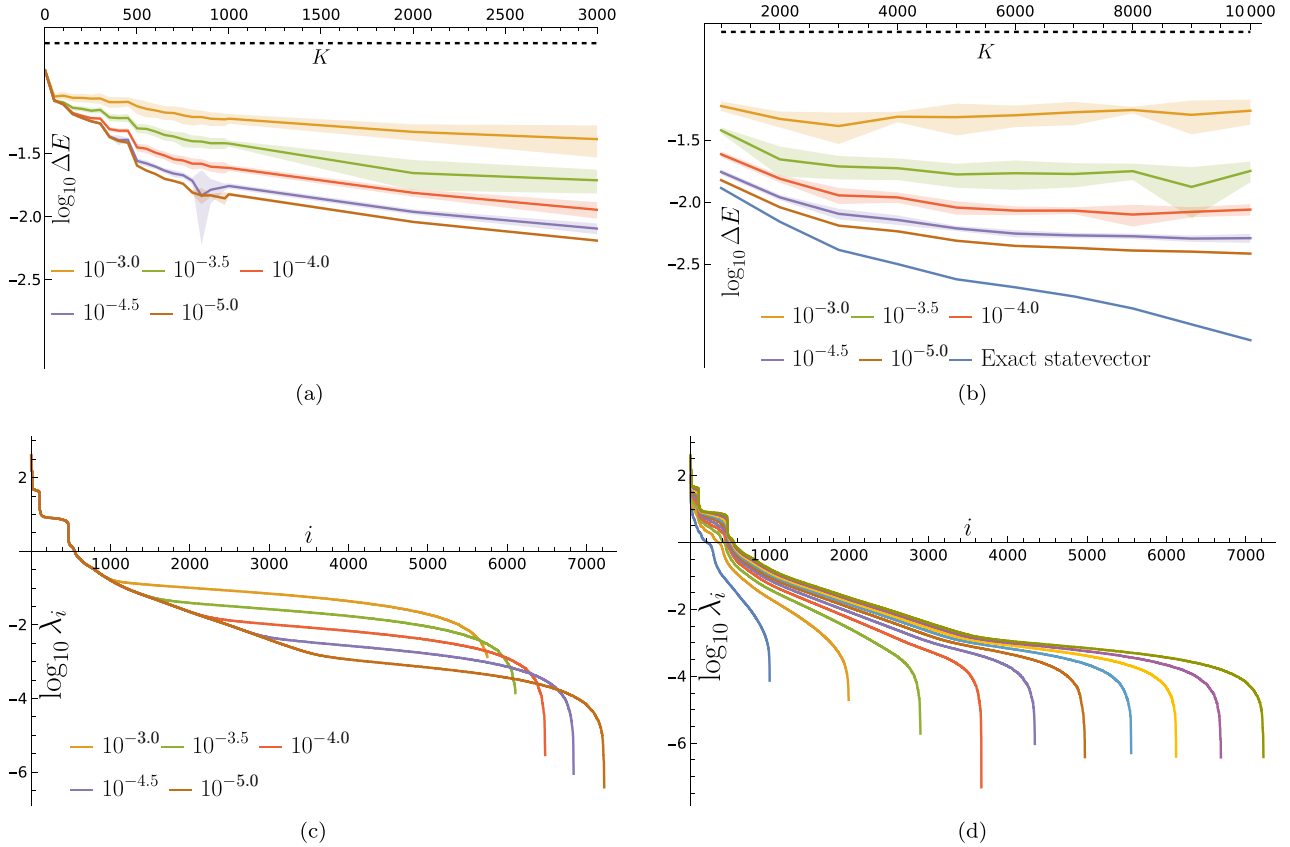


FIG. 1. (a) Error of SSE for the spin chain problem with subspace dimension K , with shot noise level ϵ per matrix element (denoted by the legend), with the energy of the input state denoted by the black dotted line. Faded regions indicate standard deviations of SSE result as determined by results from 20 independent runs. (b) Errors as K increases, demonstrating that increasing K beyond a threshold set by the noise level will not decrease the estimation error. (c) Impact of matrix element error on the eigenvalues of the matrix S , for $K = 10\,000$ on the spin chain problem (with unphysical negative eigenvalues discarded), indicating that the source of the decrease in accuracy in energy is the removal of information from the low magnitude eigenvalues. The truncation dimension required is typically found to be well above the noise-induced levelling off on this plot. (d) Impact of increasing K on the eigenvalues of the matrix S for $\epsilon = 10^{-5}$ for $K = 1000, 2000 \dots 10\,000$ increasing from left to right (with negative eigenvalues removed) indicating that the significant eigenvalues are increased as more information is included in the subspace expansion by increasing K , however, this does not substantially raise lower magnitude eigenvalues above the noise floor.

A. Performance with increasing subspace size

In this section, we will investigate the performance of SSE as the size of the subspace K increases. For simplicity, we will first assume the shot noise that is added to all the observables that form the entries of H and S is drawn from a complex Gaussian distribution of mean 0 and standard deviation ϵ , and then examine the effects of accurately modeled shot noise from classical shadow data in Secs. IV B and IV C.

Previously in the standard application of subspace expansion, the cost of measuring the expectation values grows quadratically with the subspace size, thus we are usually restricted to very small subspace size of $K \sim 10$. There is no such restriction for SSE. In Fig. 1(a), we apply SSE up to a subspace size of $K = 3000$ with different levels of shot noise. We see that increasing K beyond the $K \sim 10$ regime leads to a strong reduction in the errors in the estimated energy. We can achieve almost an order of magnitude reduction with a modest shot noise of $\epsilon \sim 10^{-3}$, and almost two orders of magnitude reduction when we have a low shot noise of $\epsilon \sim 10^{-5}$.

In Fig. 1(b), we look at even larger subspace sizes up to $K = 10^4$, which nears the maximum K available when using

up to 3-local Paulis as our expansion basis operators (10 690 for 14 qubits). In the noiseless case, the error in the estimated ground state energy decreases monotonically with the increase of K . With a finite level of shot noise per matrix element, the error in the ground state energy found by SSE hits an *error floor* as we increase the size of the subspace K . This error floor will decrease as we decrease the amount of shot noise, with a rough fit of $\Delta E_{\text{floor}} \propto \epsilon^{0.6}$ found in our numerics.

The phenomenon of the error floor shown in Fig. 1(b), whereby the energy error with a given level of noise saturates as K is increased can be explained by looking at the eigenvalues of S . In Fig. 1(b), we show the eigenvalue spectrum of S for $K = 10^4$ at varying levels of shot noise, demonstrating that the magnitude of some eigenvalues will be so small such that they are buried under the shot noise as the noise increases. The information about the subspace carried by these eigenvalues and their associated eigenvectors needs to be excluded in the subspace expansion to keep the generalized eigenvalue problem well-conditioned. Hence, for a given shot noise level, there is a point where increasing K will only add low-magnitude eigenvalues buried under the noise and will

not assist in improving the ground state energy, reaching the error floor mentioned.

We also examine the effect of increasing K on the eigenvalues of S . In Fig. 1(c), we plot these eigenvalues for different K at the noise level of $\epsilon = 10^{-5}$. For low K , the fraction of eigenvalues that are above the level of the noise is greater, whereas for higher K we are including an increasing number of operators that add little information to the subspace expansion, and so have a growing number of eigenvalues that only contain information about the noise. One thing to note is that as K increases, the eigenvalues of the matrix grow roughly linearly (see Appendix F) but the eigenvalues fall off exponentially with index i (at least in the intermediate region between the large eigenvalues and those which consist only of noise)—the linear increase in signal-to-noise ratio as K increases is not able to raise many more eigenvalues above the noise threshold.

B. Simulations using shadow variances

Until this point, we have been considering shot noise producing uniform Gaussian errors on the matrix elements of S, H , analogous to shot noise from measuring each of the elements independently. However, in the realistic setting where the expectation values are estimated from a fixed number of classical shadow snapshots, we do not expect the level of noise on each matrix element to be the same. In particular, when using local Clifford shadows, the variance of our estimators will have an extra factor of 3^w , where w is the weight of the observable (see Appendix D), meaning that instead of a single value of ϵ shared between all matrix elements as in the numerics in Fig. 1, each matrix element corresponding to an observable of weight w has a new error $\epsilon \rightarrow \sqrt{3^w}\epsilon$. We therefore perform additional numerics where we accurately simulate the shot noise produced by the finite size of the classical shadow to examine its effect on the subspace expansion. In Fig. 2, we perform SSE on a series of states obtained from different steps along the VQE optimization of the spin chain model, with the numbers of snapshots between 10^6 and 10^8 , and compare the results to the energy of the initial state as determined from direct estimation with 10^6 snapshots. We find substantial improvements for all numbers of snapshots for high-energy initial states, but for states closer to the ground state, below a certain number of snapshots, the effect of shot noise on SSE means that we can get energy estimates with higher variance than direct estimation. This effect can be mitigated by increasing the number of snapshots. In our experiment, with 10^8 snapshots we will always obtain a better result compared to direct estimation.

It is worth noting that the original state (corresponding to the identity expansion operator) is in the full K -dimensional subspace. However, when we try to regularize the subspace via transformation and truncation as outlined in Sec. III C, the resultant \bar{K} -dimensional subspace may not fully contain the original state anymore. This is the reason we can reach an energy above direct estimation. This can be potentially solved by creating another regularization procedure, explicitly keeping the original state in the output subspace, e.g., by regularizing only the subspace outside the original state. In the present case, since the direct estimated energy, being an

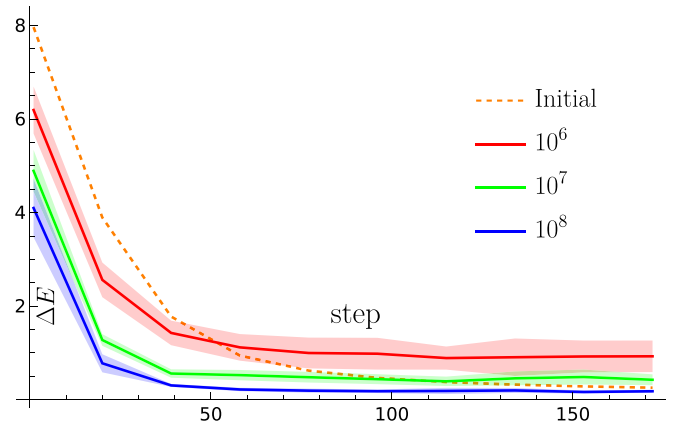


FIG. 2. Energies of states over the course of a VQE optimization of the spin chain model using simple gradient descent with line search, obtained using 10^6 snapshots. The orange dashed line corresponds to the energy of the state at that point in the optimization, whereas the red, green, and blue curves correspond to the energies achieved using SSE (with $K = 5000$), using accurate shadow variances for a shadow with the number of snapshots indicated in the legend. Faded regions indicate standard deviation of SSE result as determined by results from 20 independent runs.

entry in the H , is always a known quantity, we will simply output the direct estimated energy if the SSE energy is higher. In this way, we can ensure our method will always provide a better or at least an equally good estimate compared to direct estimation.

While we realize that finding a lower energy than direct estimation using SSE requires either a high-energy starting state or a large number of snapshots, we note that there are multiple techniques that could alleviate the shot burden. For example, derandomization [36] can be used after the identification of good observables for the subspace expansion. We also note the existence of alternative classical shadow techniques that can achieve greater performance on some restricted classes of observables [37] that could reduce the overhead based on problem-specific knowledge. Recent works also allow for the reduction of mean-squared errors of the estimators produced from classical shadows via a bias-variance trade-off which would be effective at reducing the error on the high-weight observables. This can be performed either through rescaling conventional shadow estimators [31] or using cumulant expansions for estimating k -RDMs of fermionic systems [38]. We also note that recent works provide a way to reduce the sampling overhead for high-weight observables by optimizing the shadow inverse map [39–41].

C. Effects of gate noise

To investigate the effect of gate noise on our algorithm, we perform density matrix simulations on a seven-qubit spin chain, using a parametrized quantum circuit with 161 single-qubit gates and 60 two-qubit gates, adding single or two-qubit depolarizing noise after each gate, with the noise strength for two-qubit gates being five times of that for single-qubit gates. We do this for a range of gate noise strengths leading to differing circuit fault rates λ , which are the expected number

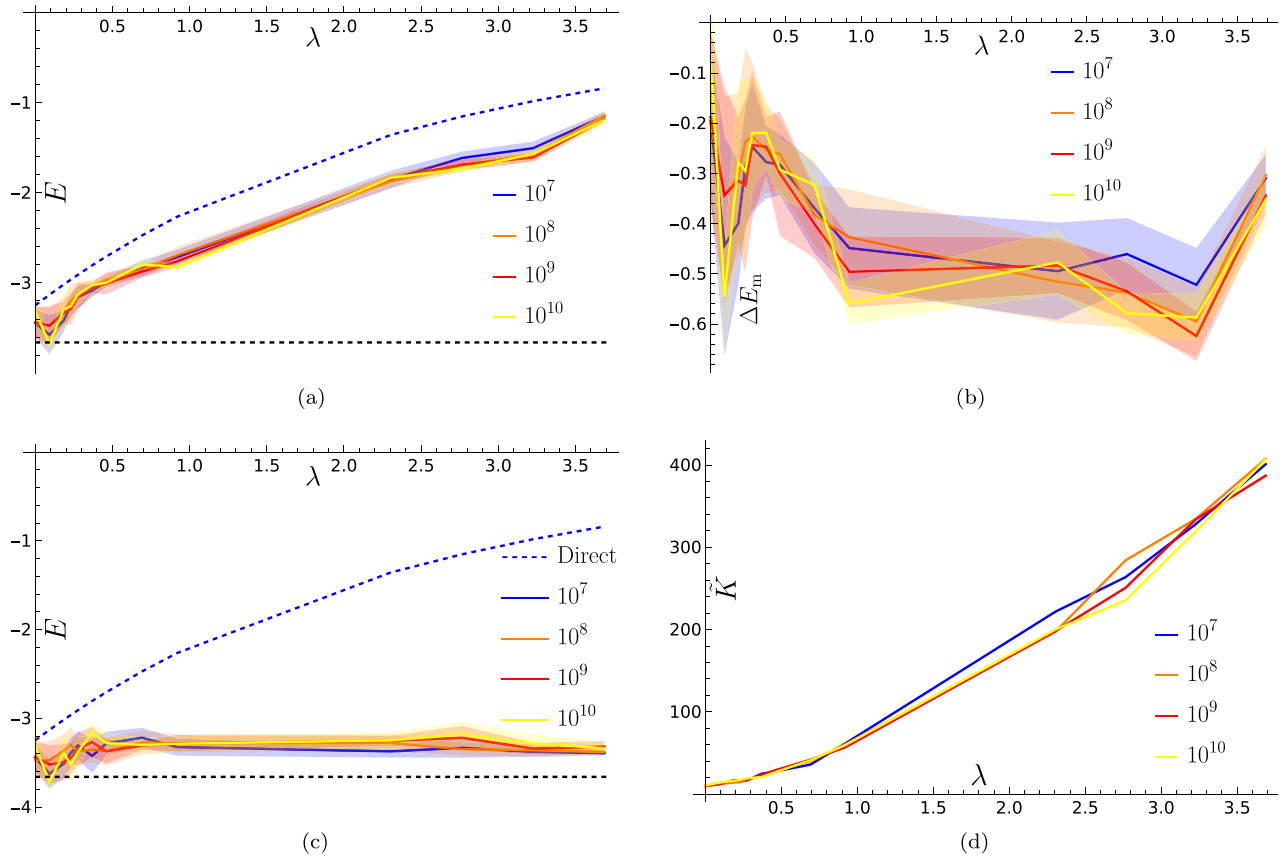


FIG. 3. (a) Energies obtained via simulations with gate noise, comparing the direct measurement of energy with the use of SSE with varying levels of gate noise, denoted by the circuit fault rate λ and shot noise from snapshot counts as given in the legend. SSE lines correspond to an average over 20 separately sampled instances of shot noise with outliers removed, the shaded regions are the standard deviations of these distributions. The effective dimension of the subspace \tilde{K} is allowed to be 15 at maximum. (b) Energy improvements of the SSE mitigated observables over the direct estimation on the noisy state in Fig. 3(a) showing an energy improvement of ~ 0.5 over the whole noise range, approximately independent of shot count. (c) The same calculation, as in Fig. 3(a) but with \tilde{K} allowed to vary freely. The average values of \tilde{K} found by the regularizer for the numerics in this plot are shown in Fig. 3(d). (d) Average value of \tilde{K} found for the circuit used to generate Fig. 3(c) with fault rate λ . The value of \tilde{K} increases past the dimension of the Hilbert space in this small example, but remains well below the 4^n possible Pauli terms that can be used in the expansion for gate noise mitigation.

of faults per circuit execution, given by the sum of error probabilities over all gates $\lambda = \sum_i p_i$ and simulate shot noise as in Sec. IV B for a range of shot noise strengths denoted by the number of snapshots taken. We take $K = 500$ and in Fig. 3(a) show the performance of SSE on noisy circuits, producing an energy improvement over the direct noisy estimate of approximately 0.5 over the noise range, with little dependence on the shot count as shown in Fig. 3(b), demonstrating the robustness of the method as a form of error mitigation, distinct from existing error mitigation methods that combine shadows and subspace expansion through reconstruction and expansion in terms of the density matrix ρ [33,42].

Noisy simulations of quantum circuits are resource intensive and thus we have only simulated up to seven qubits, which corresponds to a Hilbert space dimension of $d = 2^7$ with 256 real parameters. The Hermitian and trace-preserving operator space that density operators live in has the dimension of $d^2 - 1 \sim 15\,000$. The small system size runs the risk of having an expansion subspace with a dimension larger than the degree of freedom needed to simulate the full quantum system. Hence, we have avoided this by clamping the

maximum possible value of \tilde{K} to be 15 in Fig. 3(a). In Fig. 3(c), we show the same simulations without any restrictions on \tilde{K} , and show the corresponding values of \tilde{K} found in Fig. 3(d), where they increase with the noise level. We see that indeed increasing the dimension of the subspace is an excellent way to mitigate gate noise in the circuit. In fact, as shown in Fig. 3(c), we are able to achieve an estimation accuracy that is largely independent of the circuit noise level if we put no restriction on the dimension of the expansion.

D. Quantum chemistry and excited states

Besides spin chain Hamiltonians, we also examine the performance of SSE on a quantum chemistry Hamiltonian. We use the 12-qubit LiH Hamiltonian in the STO-3G basis [43] and run VQE using a UCCSD ansatz [44]. We then perform SSE on the output VQE states along the optimization trajectory. In Fig. 4, we show the VQE energies achieved by the orange line and the SSE energies achieved from the state at that step number using blue markers, with a shot noise of $\epsilon = 10^{-6}$. SSE is unable to improve on the Hartree-Fock state

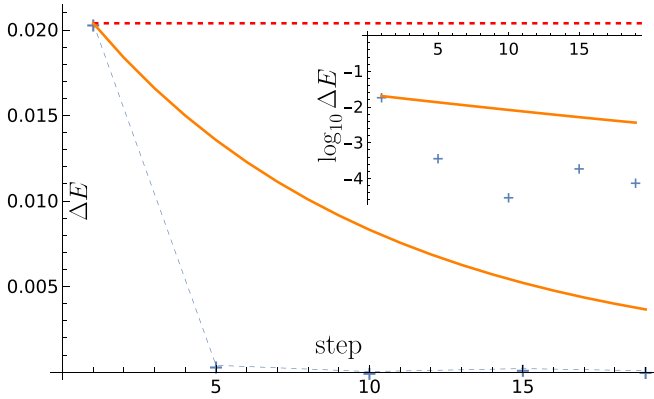


FIG. 4. Energy estimation errors of direct estimation (orange curve) and SSE (blue markers) as a function of the number of VQE steps (orange, using UCCSD ansatz with a single layer) used to generate the initial expansion state, starting from the Hartree-Fock state for LiH. The Hartree-Fock energy is shown by the red dotted line, and SSE is unable to improve on this due to the long Pauli strings require but even after a small number of VQE steps, SSE is able to rapidly make improvements on the ground state energy. The inset shows the log of the energy error for SSE. Shot noise level $\epsilon = 10^{-6}$.

due to its restriction to local Pauli strings, but even after a small number of steps, two orders of magnitude improvement in the energy estimation error ΔE can be achieved over the directly estimated VQE energy. In this case, the Hamiltonian of LiH used had a ground state with the correct electron number, and so no accommodation needed to be made for the electron number symmetry of the Hamiltonian. However, if this were not the case and the Hamiltonian had a nonphysical ground state with a different number of electrons, the techniques in Appendix A would be required to ensure the solution was not corrupted with lower energy states of the wrong electron number.

As the generalized eigenvalue problem [Eq. (3)] outputs an entire spectrum of energies within the subspace, it is interesting to compare the spectrum produced by SSE and the true spectrum of the Hamiltonian. In Fig. 5, we show the energy difference between the two spectra for LiH for a run of SSE with $K = 2000$ and without shot noise from the final state output by the variational algorithm in Fig. 4. The ground state energy difference found is approximately $\Delta E = 10^{-4}$, with the inset showing the two spectra individually (SSE in blue, true spectrum in orange). We can see that for the first ~ 80 eigenstates, the energies are in strong agreement with an energy difference of $\Delta E \lesssim 10^{-2}$. This falls off rapidly as we move to higher energy eigenstates. The agreement at the low-energy level suggests the possibility of further extending our techniques to applications involving excited states.

V. CONCLUSION

In this paper, we present SSE which leverages classical shadow data to perform high-dimensional subspace expansions for estimating ground state energies of quantum systems. We numerically investigate our approach by finding the ground state of a 14-qubit spin chain as well as obtaining

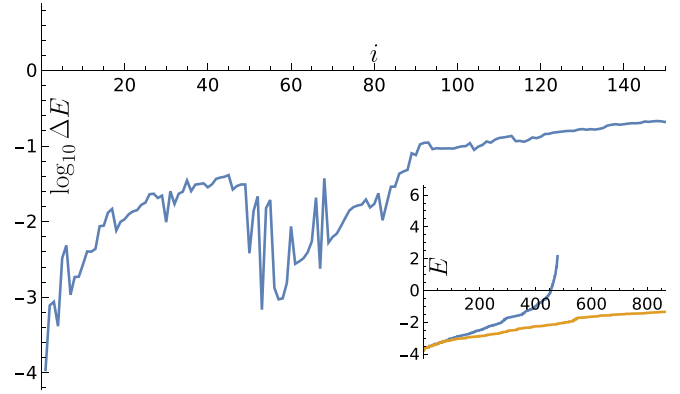


FIG. 5. The energy difference between the spectrum produced by SSE and the true spectrum of LiH for the first 150 energy levels, indicating some agreement for the lowest-lying eigenstates but rapidly diverges. The inset plot shows the SSE spectrum (blue) and the true spectrum (orange), extending into higher energy levels.

excited state energies of a LiH Hamiltonian as relevant in quantum chemistry.

We demonstrate that our method remains effective for subspace expansions of dimension up to the thousands—which is beyond the size limitations of previous approaches by orders of magnitude. We observe more than one order of magnitude error reduction when using observable expectation values under a typical level of shot noise $\propto 10^{-3}$, while indeed further error reduction can be achieved by further suppressing shot noise (increasing the number of snapshots or circuit repetitions).

We specifically considered simple local Clifford shadows for estimating expansion operators as local Pauli strings and found that our method is robust against relatively small overlaps with the true ground state and, indeed, outperforms direct energy estimation (using the available expected values) in all examples. However, when the initial state is already very close to the true ground state, then we observe that an increased number of shadow snapshots are needed to obtain improvements beyond direct energy estimation. Furthermore, we demonstrate that the approach is also robust against gate noise in the state-preparation circuits.

Based on the potential advantages outlined above, when one performs ground-state energy estimation using classical shadows, our approach can be additionally performed fully in postprocessing at no quantum cost and can potentially further reduce errors in the energy estimates; of course, in the worst case our approach yields no improvement in which case one still has access to direct energy estimate.

Given SSE is particularly effective when the initial state has relatively small overlap with the true ground state a potentially powerful application could be the following: If we are given classical shadows of the ground state of a given Hamiltonian, we can use this as an initial state for SSE to probe the ground state energy of slightly modified Hamiltonians, enabling us to probe the entire nearby energy landscape from the shadow data of one single state. Furthermore, in our numerical simulations we have primarily focused on estimating ground state energies, however, as we note in Sec. III A, our method

can directly be used for estimating other observables from the ground states.

Besides performing expansion of a given state within a subspace, one can also use postprocessing techniques to project the state into a given subspace using techniques like symmetry verification [45,46], in which the symmetry can arise from the target problems or qubit encoding [19]. Reference [7] have applied classical shadows to symmetry verification in the context of quantum error correction codes and similar ideas are also discussed in Ref. [8]. It will be interesting to see if classical shadows can be applied to more general techniques that incorporate both symmetry verification and subspace expansion [47] as discussed in Appendix A.

Many further improvements are possible for SSE. The current version of SSE is still very susceptible to shot noise due to the finite number of shadow snapshots. This can be potentially alleviated by studying alternative shadow schemes and expansion bases, and also better ways to select the relevant expansion bases to perform the regularization.

ACKNOWLEDGMENTS

The authors thank M. Goh and H. Jnane for helpful comments. Z.C. is supported by the Junior Research Fellowship from St. John's College, Oxford. B.K. thanks the University of Oxford for a Glasstone Research Fellowship and Lady Margaret Hall, Oxford for a Research Fellowship. The numerical modeling involved in this study made use of the QUANTUM EXACT SIMULATION TOOLKIT (QUEST) [48] and the recent development QUESTLINK [49], which permits the user to use MATHEMATICA as the integrated front end, and PYQUEST [50], which allows access to QUEST from PYTHON. We are grateful to those who have contributed to all of these valuable tools. B.K. thanks UKRI for the Future Leaders Fellowship Theory to Enable Practical Quantum Advantage (MR/Y015843/1). The authors also acknowledge support from the EPSRC projects Robust and Reliable Quantum Computing (RoarQ, EP/W032635/1), Software Enabling Early Quantum Advantage (SEEA, EP/Y004655/1), and the EPSRC QCS Hub grant (Agreement No. EP/T001062/1). This research was funded in whole, or in part, by UKRI (MR/Y015843/1).

APPENDIX A: SYMMETRY CONSIDERATIONS

It may be desirable for the subspace expansion to find the lowest energy state with a certain symmetry, that is, within a subspace of the Hilbert space which has the correct eigenvalue of a symmetry operator A . For example, we may want to restrict our solution to the space with the correct number of electrons in a quantum chemistry application to ensure the ground state energy estimate is not corrupted by nonphysical states with the wrong number of electrons.

If the state ρ has the correct value of this symmetry, it is straightforward to modify the subspace expansion to only choose operators that commute with the symmetry operator $[G_i, A] = 0$. However, when our state ρ does not itself lie within the symmetry subspace, we must instead use postprocessing to mitigate for the component with the wrong symmetry. One way of achieving this is to perform an analog

of symmetry verification [46], and use existing techniques and perform an internal subspace expansion to individually mitigate all the matrix elements required [8,47]. However, this requires increasing the weight of all observables measured by the weight of the symmetry operator(s) and may produce a large overhead due to the exponentially scaling shot requirements for higher weight observables.

There is an alternative method which we describe for the case of a single symmetry operator A (although it is easily generalized). We can additionally determine the elements of the matrix

$$A_{ij} = \text{Tr}(G_i A G_j \rho) \quad (\text{A1})$$

and, before solving for the ground state energy, solve the generalized eigenvalue problem for the symmetry operator:

$$A\vec{w} = A_{\vec{w}}\mathbf{S}\vec{w}. \quad (\text{A2})$$

From solving this equation, we can identify a subspace within our subspace expansion with a particular eigenvalue of the symmetry operator A . That is, we can diagonalize the matrix $\mathbf{S}^{-1}\mathbf{A}$ to find a transformation from the full subspace to the subspace with the desired symmetry eigenvalue. This will produce a transformation similar to Eq. (7) to be performed to produce symmetry adapted matrices $\mathbf{H}_A, \mathbf{S}_A$ which will then be used for a better estimate of the lowest energy within the correct symmetry subspace.

APPENDIX B: IDENTIFYING GOOD OBSERVABLES THROUGH TIME DEPENDENCE

We present here an alternative method for finding the K -significant observables based on algorithmic shadow spectroscopy [10]. Given an initial state $|\psi\rangle = \sum_{k=1}^d c_k |\psi_k\rangle$, we time evolve it to obtain the quantum state $|\psi(t)\rangle := e^{-it\mathcal{H}}|\psi\rangle$. Measuring the expected value of any observable O , the signal $S(t)$ is the following function of time:

$$S(t) := \langle \psi(t) | O | \psi(t) \rangle = \sum_{k,l=1}^d I_{kl} e^{-it(E_l - E_k)}. \quad (\text{B1})$$

Here, the intensity carries all dependency on the observable as $I_{kl} = c_k^* c_l \langle \psi_k | O | \psi_l \rangle$.

It follows that each signal has a phase-shifted cosinusoidal time dependence as

$$S(t) = \sum_{k=1}^d |c_k|^2 \langle \psi_k | O | \psi_k \rangle + \sum_{k<l}^d |I_{kl}| \cos[t(E_l - E_k) + \phi_{kl}]. \quad (\text{B2})$$

Here we have used that $I_{kl} = |I_{kl}| e^{-i\phi_{kl}} = c_k^* c_l \langle \psi_k | O | \psi_l \rangle$.

Assuming a low-energy initial state that has a dominant overlap with the ground state via $c_1 \gg c_k$ with $k > 1$, we can approximate the signal by dropping all terms nonlinear in c_k as

$$S(t) \approx |c_1|^2 \langle \psi_1 | O | \psi_1 \rangle + \sum_{k=2}^d |I_{1,k}| \cos[t(E_1 - E_k) + \phi_{1,k}]. \quad (\text{B3})$$

To tell whether an observable O can induce a transition between the ground state and an excited state as $\langle \psi_k | O | \psi_1 \rangle$, it

suffices to verify that the signal $S(t)$ is constant. As we need not estimate the frequency of the signal $\cos[t(E_1 - E_k) + \phi_{1,k}]$, it suffices to perform a relatively short time evolution to obtain the signal $S(t)$ and apply a Ljung-Box test to verify whether the signal is statistically significantly different from shot noise. The approach is robust to algorithmic errors in the time evolution since the frequency and the phase shift $\cos[t(E_1 - E_k) + \phi_{1,k}]$ need not be estimated.

APPENDIX C: PROOF OF EQ. (6)

As detailed in the main text, we consider the solution to the generalized eigenvalue equation as the eigenvalue equation for the matrix $M = S^{-1}H$ as

$$S^{-1}H\vec{w} = E_{\vec{w}}\vec{w}. \quad (C1)$$

We denote entries of the q th column vector of M as $u^{(q)}$ which is equivalent to $u^{(q)} = S^{-1}v^{(q)}$, where $v^{(q)}$ is the q th column vector of the Hamiltonian matrix H . Lemma 2 in Ref. [30] quantifies the shot noise propagation in this inverse-matrix-vector equation $u^{(q)} = A^{-1}v^{(q)}$ in terms of the total variance, i.e., the expected Euclidean distance squared, of the vector $u^{(q)}$ due to shot noise as $\epsilon_q = \sum_k \text{Var}[u_k^{(q)}]$.

The following bound was obtained for the shot-noise propagation:

$$\epsilon_q^2 = \sum_{k,l=1}^v a_{kl} \text{Var}\{[S]_{kl}\} + \sum_{k=1}^v b_k \text{Var}[v_k^{(q)}],$$

$$\text{where } a_{kl} := \sum_{i,j=1}^v [S^{-1}]_{ik}^2 [S^{-1}]_{lj}^2 [v_l^{(q)}]^2$$

$$\text{and } b_k := \sum_{l=1}^v \{[S^{-1}]_{kl}\}^2.$$

A bound on the following coefficients was proved as

$$a_{kl} \leq \|v^{(q)}\|_\infty^2 \sum_{i,j=1}^v [S^{-1}]_{ik}^2 [S^{-1}]_{lj}^2. \quad (C2)$$

We assume that the shot noise in the matrix entries $[S]_{kl}$ and $[H]_{kl}$ is independent (zero covariance). We expect the bounds on the shot-noise propagation to be loose, and we are primarily interested in establishing the scaling of the error rather than a guaranteed bound.

We can apply the variance bound for classical shadows given that in the present work we assume the matrix entries $\text{Var}\{[S]_{kl}\}$ are estimated from classical shadows. Given all matrix entries are of the form of Eq. (4), the variance is upper bounded as

$$\text{Var}\{[S]_{kl}\} \leq N_s^{-1} 3^{2w'}. \quad (C3)$$

Above we used the largest weight w' among the expansion operators and N_s is the number of shots used.

Similarly, we can bound the variance of the Hamiltonian matrix entries as

$$\text{Var}[v_k^{(q)}] = \text{Var}\{[H]_{kq}\} \leq N_s^{-1} 3^{w+2w'},$$

where we used the largest weight w among the Pauli terms in the Hamiltonian.

Substituting this back into the error propagation formula, we obtain

$$\epsilon_q^2 \leq N_s^{-1} 3^{2w'} \sum_{k,l=1}^v a_{kl} + N_s^{-1} 3^{w+2w'} \sum_{k=1}^v b_k. \quad (C4)$$

Substituting back the explicit form of b_k and our upper bound for a_{kl} , we finally obtain

$$\epsilon_q^2 \leq N_s^{-1} 3^{2w'} \|v^{(q)}\|_\infty^2 \|S^{-1}\|_2^4 + N_s^{-1} 3^{w+2w'} \|S^{-1}\|_2^2, \quad (C5)$$

where $\|\cdot\|_2$ is the Hilbert-Schmidt or Frobenius norm of a matrix.

We have so far quantified the total variance in the q th column vector of M . We define the total variance in M as $\epsilon_M^2 = \sum_{q=1}^K \epsilon_q^2 = \sum_{qk} \text{Var}\{[M]_{kq}\}$, which is the expected Hilbert-Schmidt distance in the matrix M due to shot noise, using that

$$\sum_{q=1}^K \|v^{(q)}\|_\infty^2 \leq K \max_{kl} [H]_{kl} \leq K \|H_{\text{tot}}\|_\infty, \quad (C6)$$

where $\|H_{\text{tot}}\|_\infty$ is the absolute largest eigenvalue of the problem Hamiltonian. We finally obtain the bound

$$\epsilon_M^2 \leq N_s^{-1} 3^{2w'} K [\|H_{\text{tot}}\|_\infty \|S^{-1}\|_2^4 + 3^w \|S^{-1}\|_2^2]. \quad (C7)$$

APPENDIX D: VARIANCE OF LOCAL CLIFFORD SHADOW

Suppose we are given the ability to prepare a given state ρ , and we want to construct its classical shadow using local Clifford twirling. In a given circuit run, we will prepare the state ρ and then perform measurements corresponding to a set of random local Pauli operators \hat{G} , which give the set of random outcomes $\hat{g} \in \{\pm 1\}^{\otimes N}$. The output state $|\hat{G}, \hat{g}\rangle$ is stabilized by the set of Pauli generators $\hat{g} \odot \hat{G}$, where \odot is elementwise multiplication, and we will call this a *projection* of the input state ρ . If we use \mathcal{D} to denote the channel where each qubit is locally depolarized with 2/3 probability, then by applying the inverse of this channel on the output state, $|\hat{\rho}\rangle = \mathcal{D}^{-1}|\hat{G}, \hat{g}\rangle$, will give a classical snapshot of the input state ρ and the collection of them form the classical shadows of ρ . We can obtain the expectation value of a Pauli observable O with respect to the input state ρ by taking the expectation value of O with respect to these snapshots:

$$\langle\langle O|\rho\rangle\rangle = \mathbb{E}[\langle\langle O|\hat{\rho}\rangle\rangle] = \mathbb{E}[\langle\langle O|\mathcal{D}^{-1}|\hat{G}, \hat{g}\rangle\rangle].$$

Using the fact that $\langle\langle O|\mathcal{D}^{-1}|\hat{G}, \hat{g}\rangle\rangle = 3^w \langle\langle O|\hat{G}, \hat{g}\rangle\rangle$, where w is the weight of O , we can estimate $\langle\langle O|\rho\rangle\rangle$ via two steps. Step one is to obtain the random variable $\hat{Q} = \langle\langle O|\hat{G}, \hat{g}\rangle\rangle$ through postprocessing for each snapshot. Two possible scenarios can arise:

(1) $O \in \langle\hat{G}\rangle$, i.e., O can be obtained through multiplying the set of local Pauli measurements in the given snapshot, or in other words, $+O$ or $-O$ is a stabilizer of the snapshot. This happens with the probability 3^{-w} since each local Pauli measurement is chosen randomly from $\{X, Y, Z\}$. In this case, the output random variable \hat{O} follows the same distribution

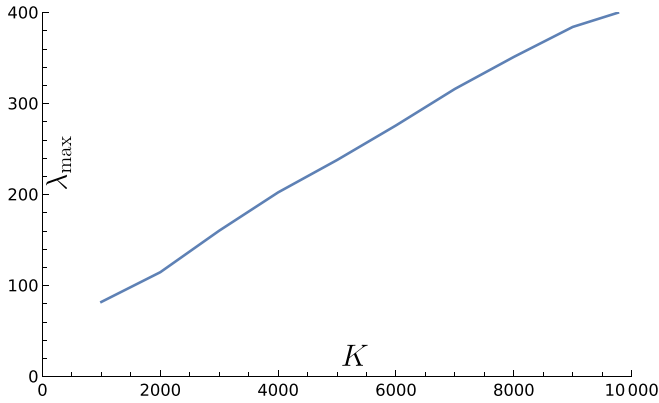


FIG. 6. Demonstration of the largest eigenvalue of \mathbf{S} increasing linearly with K , as expected for a matrix of overlaps $S_{ij} = \langle \psi_i | \psi_j \rangle$, where the states are mostly similar to each other (consisting of only small local changes). For example, in the extreme case where we take all $G_i = I$, then the largest eigenvalue is K .

as directly measuring O on state ρ . Thus, we have $\mathbb{E}[\hat{O}] = \langle \langle O | \rho \rangle \rangle$.

(2) $O \notin \langle \hat{G} \rangle$, which happens with the probability $1 - 3^{-w}$. In this case, we simply have $\langle \langle O | \hat{G}, \hat{g} \rangle \rangle = 0$.

From the description above, we have

$$\mathbb{E}[\hat{Q}] = 3^{-w} \mathbb{E}[\hat{O}] + (1 - 3^{-w})0 = 3^{-w} \mathbb{E}[\hat{O}],$$

$$\mathbb{E}[\hat{Q}^2] = 3^{-w} \mathbb{E}[\hat{O}^2] + (1 - 3^{-w})0^2 = 3^{-w} \mathbb{E}[\hat{O}^2],$$

$$\text{Var}[\hat{Q}] = \mathbb{E}[\hat{Q}^2] - \mathbb{E}[\hat{Q}]^2 = 3^{-w} - 3^{-2w} \mathbb{E}[\hat{O}]^2.$$

Step two is rescaling the result obtained from \hat{Q} by 3^w , which gives rise to the estimator $\hat{R} = 3^w \hat{Q}$. This gives

$$\mathbb{E}[\hat{R}] = 3^w \mathbb{E}[\hat{Q}] = \mathbb{E}[\hat{O}] = \langle \langle O | \rho \rangle \rangle$$

as expected, so the shadow estimator \hat{R} is an unbiased estimator of $\langle \langle O | \rho \rangle \rangle$. Its variance is given as

$$\text{Var}[\hat{R}] = 3^{2w} \text{Var}[\hat{Q}] = 3^w - \mathbb{E}[\hat{O}]^2.$$

APPENDIX E: PRIOR WORK ON SHOT NOISE PROPAGATION

As derived in the Appendix of Ref. [33], when solving

$$\mathbf{H} \vec{w} = E_{\vec{w}} \mathbf{S} \vec{w},$$

the shot noise in \mathbf{H} and \mathbf{S} , denoted as $\delta \mathbf{H}$ and $\delta \mathbf{S}$, will also lead to noise in the calculated weight vectors and energies. One can show that

$$\delta E_{\vec{w}} = \vec{w}^\dagger (\delta \mathbf{H} - E_{\vec{w}} \delta \mathbf{S}) \vec{w}.$$

As shown in Ref. [33], this error is bounded by

$$\delta E_{\vec{w}} \leq \|\mathbf{S}^{-1}\|_\infty (\|\delta \mathbf{H}\|_\infty + |E_{\vec{w}}| \|\delta \mathbf{S}\|_\infty).$$

Furthermore, we can bound the energy as $|E_{\vec{w}}| \leq \|\mathbf{H}_{\text{tot}}\|_\infty$, which yields the bound

$$\delta E_{\vec{w}} \leq \|\mathbf{S}^{-1}\|_\infty (\|\delta \mathbf{H}\|_\infty + \|\mathbf{H}_{\text{tot}}\|_\infty \|\delta \mathbf{S}\|_\infty).$$

The dependence on $\|\mathbf{S}^{-1}\|_\infty$ means that truncation of the expansion basis is essential such that \mathbf{S} is not close to being singular.

APPENDIX F: DEPENDENCE OF EIGENVALUES OF \mathbf{S} ON K

Here we show the dependence of the largest eigenvalue of the matrix \mathbf{S} on the subspace dimension K (Fig. 6), demonstrating that the inclusion of more subspace dimensions does increase the amount of information present about the most significant directions within the subspace. However, as this increase is linear, we are not able to significantly increase the number of eigenvalues above the noise floor shown in Fig. 1(d) by increasing K .

-
- [1] P. Shor, Algorithms for quantum computation: Discrete logarithms and factoring, in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science* (IEEE Computer Society 1730 Massachusetts Ave., NW Washington, DC USA, 1994), pp. 124–134.
- [2] 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**, 4213 (2014).
- [3] D. Grinko, J. Gacon, C. Zoufal, and S. Woerner, Iterative quantum amplitude estimation, *npj Quantum Inf.* **7**, 52 (2021).
- [4] 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).
- [5] S. Chen, W. Yu, P. Zeng, and S. T. Flammia, Robust shadow estimation, *PRX Quantum* **2**, 030348 (2021).
- [6] D. E. Koh and S. Grewal, Classical shadows with noise, *Quantum* **6**, 776 (2022).
- [7] H.-Y. Hu, R. LaRose, Y.-Z. You, E. Rieffel, and Z. Wang, Logi-cal shadow tomography: Efficient estimation of error-mitigated observables, [arXiv:2203.07263](https://arxiv.org/abs/2203.07263).
- [8] H. Jnane, J. Steinberg, Z. Cai, H. C. Nguyen, and B. Koczor, Quantum error mitigated classical shadows, *PRX Quantum* **5**, 010324 (2024).
- [9] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, *Nat. Phys.* **16**, 1050 (2020).
- [10] H. H. S. Chan, R. Meister, M. L. Goh, and B. Koczor, Algorithmic shadow spectroscopy, [arXiv:2212.11036](https://arxiv.org/abs/2212.11036).
- [11] G. Boyd and B. Koczor, Training variational quantum circuits with CoVaR: Covariance root finding with classical shadows, *Phys. Rev. X* **12**, 041022 (2022).
- [12] K. Nakaji, S. Endo, Y. Matsuzaki, and H. Hakoshima, Measurement optimization of variational quantum simulation by classical shadow and derandomization, *Quantum* **7**, 995 (2023).
- [13] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, Probing Rényi entanglement entropy via randomized measurements, *Science* **364**, 260 (2019).

- [14] G. I. Struchalin, Ya. A. Zagorovskii, E. V. Kovlakov, S. S. Straupe, and S. P. Kulik, Experimental estimation of quantum state properties from classical shadows, *PRX Quantum* **2**, 010307 (2021).
- [15] A. Abbas, R. King, H.-Y. Huang, W. J. Huggins, R. Movassagh, D. Gilboa, and J. McClean, On quantum backpropagation, information reuse, and cheating measurement collapse, *Adv. Neural Inform. Process. Syst.* **36**, 44792 (2023).
- [16] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, *Rev. Mod. Phys.* **92**, 015003 (2020).
- [17] A. J. Daley, I. Bloch, C. Kokail, S. Flannigan, N. Pearson, M. Troyer, and P. Zoller, Practical quantum advantage in quantum simulation, *Nature (London)* **607**, 667 (2022).
- [18] A. Abbas, A. Ambainis, B. Augustino, A. Bärttschi, H. Buhrman, C. Coffrin, G. Cortiana, V. Dunjko, D. J. Egger, B. G. Elmegreen, N. Franco, F. Fratini, B. Fuller, J. Gacon, C. Gonciulea, S. Gribling, S. Gupta, S. Hadfield, R. Heese, G. Kircher *et al.*, Quantum optimization: Potential, challenges, and the path forward, *Nat. Rev. Phys.* **6**, 718 (2024).
- [19] J. R. McClean, Z. Jiang, N. C. Rubin, R. Babbush, and H. Neven, Decoding quantum errors with subspace expansions, *Nat. Commun.* **11**, 636 (2020).
- [20] 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).
- [21] T. Takeshita, N. C. Rubin, Z. Jiang, E. Lee, R. Babbush, and J. R. McClean, Increasing the representation accuracy of quantum simulations of chemistry without extra quantum resources, *Phys. Rev. X* **10**, 011004 (2020).
- [22] J. Cohn, M. Motta, and R. M. Parrish, Quantum filter diagonalization with double-factorized Hamiltonians, *PRX Quantum* **2**, 040352 (2021).
- [23] S. Aaronson, Shadow tomography of quantum states, in *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2018 (Association for Computing Machinery, New York, 2018), pp. 325–338.
- [24] K. Wan, W. J. Huggins, J. Lee, and R. Babbush, Matchgate shadows for fermionic quantum simulation, *Commun. Math. Phys.* **404**, 629 (2023).
- [25] H.-Y. Hu, S. Choi, and Y.-Z. You, Classical shadow tomography with locally scrambled quantum dynamics, *Phys. Rev. Res.* **5**, 023027 (2023).
- [26] M. Ippoliti, Y. Li, T. Rakovszky, and V. Khemani, Operator relaxation and the optimal depth of classical shadows, *Phys. Rev. Lett.* **130**, 230403 (2023).
- [27] 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).
- [28] B. Wu and D. E. Koh, Error-mitigated fermionic classical shadows on noisy quantum devices, *npj Quantum Inf.* **10**, 1 (2024).
- [29] 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).
- [30] B. van Straaten and B. Koczor, Measurement cost of metric-aware variational quantum algorithms, *PRX Quantum* **2**, 030324 (2021).
- [31] Z. Cai, A. Chapman, H. Jnane, and B. Koczor, Biased estimator channels for classical shadows, [arXiv:2402.09511](https://arxiv.org/abs/2402.09511).
- [32] R. King, D. Gosset, R. Kothari, and R. Babbush, Triply efficient shadow tomography, [arXiv:2404.19211](https://arxiv.org/abs/2404.19211).
- [33] N. Yoshioka, H. Hakoshima, Y. Matsuzaki, Y. Tokunaga, Y. Suzuki, and S. Endo, Generalized quantum subspace expansion, *Phys. Rev. Lett.* **129**, 020502 (2022).
- [34] N. J. Higham, Computing a nearest symmetric positive semidefinite matrix, *Linear Algebra Appl.* **103**, 103 (1988).
- [35] 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. USA* **115**, 9456 (2018).
- [36] H.-Y. Huang, R. Kueng, and J. Preskill, Efficient estimation of pauli observables by derandomization, *Phys. Rev. Lett.* **127**, 030503 (2021).
- [37] M. Ippoliti, Classical shadows based on locally-entangled measurements, *Quantum* **8**, 1293 (2024).
- [38] N. Takemori, Y. Teranishi, W. Mizukami, and N. Yoshioka, Balancing error budget for fermionic k-RDM estimation, [arXiv:2312.17452](https://arxiv.org/abs/2312.17452).
- [39] A. Caprotti, J. Morris, and B. Dakić, Optimising quantum tomography via shadow inversion, *Phys. Rev. Res.* **6**, 033301 (2024).
- [40] J. Malmi, K. Korhonen, D. Cavalcanti, and G. García-Pérez, Enhanced observable estimation through classical optimization of informationally over-complete measurement data—beyond classical shadows, *Phys. Rev. A* **109**, 062412 (2024).
- [41] L. E. Fischer, T. Dao, I. Tavernelli, and F. Tacchino, Dual frame optimization for informationally complete quantum measurements, *Phys. Rev. A* **109**, 062415 (2024).
- [42] B. Yang, N. Yoshioka, H. Harada, S. Hakkaku, Y. Tokunaga, H. Hakoshima, K. Yamamoto, and S. Endo, Dual-GSE: Resource-efficient generalized quantum subspace expansion, [arXiv:2309.14171](https://arxiv.org/abs/2309.14171).
- [43] J. R. McClean, N. C. Rubin, K. J. Sung, I. D. Kivlichan, X. Bonet-Monroig, Y. Cao, C. Dai, E. S. Fried, C. Gidney, B. Gimby, P. Gokhale, T. Häner, T. Hardikar, V. Havlíček, O. Higgott, C. Huang, J. Izaac, Z. Jiang, X. Liu, S. McArdle *et al.*, OpenFermion: The electronic structure package for quantum computers, *Quantum Sci. Technol.* **5**, 034014 (2020).
- [44] J. Lee, W. J. Huggins, M. Head-Gordon, and K. B. Whaley, Generalized unitary coupled cluster wave functions for quantum computation, *J. Chem. Theory Comput.* **15**, 311 (2019).
- [45] S. McArdle, X. Yuan, and S. Benjamin, Error-mitigated digital quantum simulation, *Phys. Rev. Lett.* **122**, 180501 (2019).
- [46] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O’Brien, Low-cost error mitigation by symmetry verification, *Phys. Rev. A* **98**, 062339 (2018).
- [47] Z. Cai, Quantum error mitigation using symmetry expansion, *Quantum* **5**, 548 (2021).
- [48] T. Jones, A. Brown, I. Bush, and S. C. Benjamin, QuEST and high performance simulation of quantum computers, *Sci. Rep.* **9**, 10736 (2019).
- [49] T. Jones and S. C. Benjamin, QuESTlink—Mathematica embiggened by a hardware-optimized quantum emulator, *Quantum Sci. Technol.* **5**, 034012 (2020).
- [50] R. Meister, pyQuEST—A Python interface for the Quantum Exact Simulation Toolkit (2024), <https://github.com/rmeister/pyQuEST>.