Comparative study of adaptive variational quantum eigensolvers for multi-orbital impurity models

Anirban Mukherjee¹, Noah F. Berthusen¹,²,³, João C. Getelina¹, Peter P. Orth¹,⁴, & Yong-Xin Yao¹,⁴

Hybrid quantum-classical embedding methods for correlated materials simulations provide a path towards potential quantum advantage. However, the required quantum resources arising from the multi-band nature of \(d\) and \(f\) electron materials remain largely unexplored. Here we compare the performance of different variational quantum eigensolvers in ground state preparation for interacting multi-orbital embedding impurity models, which is the computationally most demanding step in quantum embedding theories. Focusing on adaptive algorithms and models with 8 spin-orbitals, we show that state preparation with fidelities better than 99.9% can be achieved using about \(2^{14}\) shots per measurement circuit. When including gate noise, we observe that parameter optimizations can still be performed if the two-qubit gate error lies below \(10^{-3}\), which is slightly smaller than current hardware levels. Finally, we measure the ground state energy on IBM and Quantinuum hardware using a converged adaptive ansatz and obtain a relative error of 0.7%.

https://doi.org/10.1038/s42005-022-01089-6
Eigensstate preparation for Hamiltonian systems is one promising application of noisy intermediate-scale quantum (NISQ) computers to achieve practical quantum advantage\textsuperscript{1–7}. One of the representative hybrid quantum-classical algorithms to achieve this task is the variational quantum eigensolver (VQE). It attempts to find the ground state of a given Hamiltonian $H$ within a variational manifold of states that are generated by parametrized quantum circuits $U(\theta)$ acting on a reference state $|\psi_0\rangle$. The parameters $\theta$ are obtained by classically minimizing the energy cost function $E(\theta) = \langle \psi_0 | U(\theta) H U(\theta) | \psi_0 \rangle$ that is measured on quantum hardware\textsuperscript{9–8,10}. The quality of a VQE calculation is tied to the ability of the variational ansatz to represent the ground state with high fidelity. In quantum computational chemistry, the unitary coupled cluster ansatz truncated at single and double excitations ($\text{UCCSD}$) has been extensively studied, owing to the success of the classical coupled cluster algorithm\textsuperscript{9} and double excitations ($\text{UCCSD}$) has been extensively studied, chemistry, the unitary coupled cluster ansatz truncated at single and double excitations ($\text{UCCSD}$) has been extensively studied, owing to the success of the classical coupled cluster algorithm\textsuperscript{9}.

It was found that the application of UCCSD ansatz is limited by the rapid circuit growth with system size and the deteriorating accuracy in the presence of static electron correlations\textsuperscript{11,12,13}. Therefore, alternative variants have been developed, including hardware-efficient ansätze, that improve the trainability and expressivity of the wave function ansatz\textsuperscript{3,13–20}.

Indeed, it was found that compact and numerically exact variational ground state ansätze can be adaptively constructed for specific problems using approaches like the adaptive derivative-assembled pseudo-trotter (ADAPT) ansatz\textsuperscript{13,16}. The adaptive ansatz is typically obtained by successively appending parametrized unitaries to a variational circuit with generators chosen from a predefined operator pool. In practice, the ADAPT-VQE algorithm works well with an operator pool composed of fermionic excitation operators in the UCCSD ansatz. The extended qubit-ADAPT VQE approach\textsuperscript{16} utilizes an operator pool composed of Pauli strings in the qubit representation of fermionic excitation operators in the UCCSD ansatz, which is shown to be capable of generating significantly more compact ansätze than the original ADAPT-VQE method at the price of introducing more variational parameters. As the circuit complexity (i.e., the number of two-qubit operations in the circuit) is a determining factor for practical calculations on NISQ devices, qubit-ADAPT is preferable and chosen for the comparative study in this work. Regarding the scalability of the qubit-ADAPT method towards larger system sizes, we note that reference\textsuperscript{18} reports a favorable linear system-size scaling for the adaptive ansatz complexity of nonintegrable mixed-field Ising model using the adaptive variational quantum imaginary time evolution method (AVQITE). AVQITE is known to generate variational circuits of comparable complexity as qubit-ADAPT VQE. As a first step to investigate the scalability in fermionic models, we here study qubit-ADAPT VQE for fermionic models with two and three spinful orbitals.

An alternative approach to constructing efficient wavefunction ansätze for problems in condensed matter physics is to exploit the sparsity of the Hamiltonian. Interacting electron systems are often simulated with reduced degrees of freedom, represented, for example, by a single-band Hubbard model. This simplified model features a sparse Hamiltonian including nearest-neighbor hopping and onsite Coulomb interactions only. Motivated by the simplicity of the Trotterized circuits for dynamics simulations due to Hamiltonian sparsity, the Hamiltonian variational ansatz (HVA) has been proposed by promoting the time in Trotter circuits to independent variational parameters\textsuperscript{21}. The HVA ansatz has attracted much attention and turns out to be very successful in reaching a compact state representation for sparse Hamiltonian systems including local spin models\textsuperscript{21–23}. Here, we propose to combine the flexibility of an adaptive approach with the efficiency of the HVA by designing a “Hamiltonian commutator” (HC) operator pool that contains pairwise commutators of operators that appear in the Hamiltonian.

To obtain a realistic description of correlated quantum materials, which typically contain partially filled d-orbitals such as transition metal compounds, or f-orbitals such as rare-earth and actinide systems, it is important to go beyond the single-orbital description of a simple Hubbard model\textsuperscript{24}. Intriguing physics arises from the local Hund’s coupling of electrons in different atomic orbitals. Examples are bad metallic behavior with suppressed quasiparticle coherence and orbital-selective Mott transitions or superconducting pairing, which naturally require a multi-orbital description\textsuperscript{25–28}. A multi-orbital model including additional inter-orbital hoppings and Hund’s couplings will necessarily make the Hamiltonian less sparse and consequently the HVA ansatz more complicated. Nevertheless, the complexity of material simulations can be greatly reduced by quantum embedding methods which map the infinite system to coupled subsystems, typically a noninteracting effective medium and some many-body interacting impurity models\textsuperscript{24,29–37}. These quantum embedding approaches have proven to be very effective to simulate correlated electron systems, including energies, electronic structure, magnetism, superconductivity, and spectral properties of multiple competing phases. The computational load in these approaches is shifted from the solution of a full lattice system to that of an interacting multi-orbital impurity model. Classical algorithms for solving the impurity problem, however, are not scalable, which can be more tractable with quantum computers\textsuperscript{35,38}.

In this paper, we compare the VQE circuit complexity for ground state preparation of multi-orbital many-body impurity models with a fixed HVA versus a qubit-ADAPT ansatz with different operator pools. An HC operator pool compatible with HVA is proposed to allow a fair comparison between qubit-ADAPT and fixed ansatz HVA calculations. For comparison, we also include results from UCCSD and qubit-ADAPT calculations with a simplified UCCSD pool. To connect with quantum embedding methods for realistic materials simulations, we use the Gutzwiller embedding approach\textsuperscript{33,39–44} to generate the impurity models that we employ for our benchmark\textsuperscript{35,45}. The quantum calculation we perform is general and could also be applied to other embedding methods. Numerical results from a noiseless state vector simulator and quantum assembly language (QASM)-based simulator with quantum sampling noise are presented. Important techniques for efficient circuit simulations of qubit-ADAPT VQE are discussed, including ways to simplify generators and reduce the operator pool size. We further investigate the impact of realistic gate noise by performing qubit-ADAPT VQE simulations with a realistic noise model including amplitude and dephasing channels. Finally, we measure the energy cost function of the converged VQE ansatz for the $e_g$ model composed of eight spin-orbitals on the IBM quantum processing unit (QPU) ibmq_casablanca and on Quantumum hardware.

**Results and discussion**

Quantum embedding model. Here we focus on a specific quantum embedding method: the well-established Gutzwiller variational embedding approach for correlated material simulations\textsuperscript{33,39–44}, which is known to be equivalent to rotationally invariant slave-boson theory at the saddle point approximation\textsuperscript{16,47}. Recently, our group has developed a hybrid Gutzwiller quantum-classical embedding approach (GQCE)\textsuperscript{35}. GQCE maps the ground state solution of a correlated electron lattice system to a coupled eigenvalue problem of a noninteracting quasiparticle Hamiltonian and one or multiple finite-size interacting embedding Hamiltonians\textsuperscript{44}. Within GQCE one
Hamiltonian system as \((N_DV)\) level is coupled with a single bath orbital. The interactions among the physical orbitals are specified by the Coulomb matrix \(V\). Due to symmetry, each physical orbital (positioned at zero energy level) is coupled with a single bath orbital at energy level \(\lambda\) with a coupling parameter \(D\). The models with \(M = 2\) and 3 correspond to that of \(e_g\) and \(t_{2g}\) orbitals in cubic crystal symmetry, respectively.

The Hamiltonian describes an impurity model, where the reference Hartree-Fock (HF) operator pool. The embedding Hamiltonian is much more compact, and it can be evaluated using the quantum-classical embedding framework. In quantum simulations reported below, the symmetry in a total number of electrons and spin \(z\)-component is used to transform the fermionic Hamiltonian to qubit representation.

**Variational quantum eigensolvers.** GQCE leverages quantum computing technologies to solve for the ground state of the embedding Hamiltonian, specifically the energy and one-particle density matrix. Note that the ground state is always prepared at half-filling for the embedding system, which is determined by the Gutzwiler embedding algorithm and is independent of the actual electron filling of the physical lattice model. For this purpose, we benchmark multiple versions of VQE with fixed or adaptively generated ansatz to prepare the ground state of the above embedding Hamiltonian. We consider VQE calculations with fixed UCCSD ansatz and the associated qubit-ADAPT VQE using a simplified UCCSD operator pool. The calculations are naturally performed in the molecular orbital (MO) basis representation, where the reference Hartree–Fock (HF) state becomes a simple tensor product state and fermionic excitation operators can be naturally defined. However, using a MO representation comes at the cost of reducing the sparsity of the embedding Hamiltonian compared to the atomic orbital (AO) basis representation. To take advantage of the Hamiltonian sparsity in AO representation, we consider a generalized form of the HVA and the associated qubit-ADAPT VQE with a modified HC operator pool.

**VQE algorithm.** For an \(N_q\)-qubit system with Hamiltonian \(\hat{H}\), VQE amounts to minimizing the cost function \(E(\theta) = \langle \Psi(\theta) | \hat{H} | \Psi(\theta) \rangle\) with respect to the variational parameters \(\theta\), as schematically illustrated in Fig. 2. Here, \(\langle \Psi(\theta)|U(\theta)|\Psi(\theta)\rangle\) is obtained by application of a parametrized quantum circuit \(U(\theta)\) onto a reference state \(|\Psi_0\rangle\). The cost function is evaluated on a quantum computer and the optimization is performed classically.
uses deep quantum circuits for implementations. Various approaches have been proposed to construct a more compact variational ansatz with systematically improvable accuracy. In this work, we will focus on the qubit-ADAPT VQE method\cite{Du2020}, where the ansatz takes a similar pseudo-Trotter form:

\[
|\Psi(\theta)\rangle = \prod_{j=1}^{N_g} e^{-\beta_j \hat{h}_j} |\Psi_0\rangle.
\]  

With qubit-ADAPT, the ansatz is recursively expanded by adding one unitary at a time, followed by reoptimization of parameters. The additional unitary is constructed with a generator selected from a predefined Pauli string pool which gives maximal energy gradient amplitude $|g|_{\text{max}}$ at the preceding ansatz state. The ansatz expansion process iterates until convergence, which is set by $|g|_{\text{max}} < 10^{-4}$ here. Note that we have set the half bandwidth of the original noninteracting lattice model to $D = 1$, such that $|g|_{\text{max}} \sim 0.1$ meV in physical systems with $D \sim 1$ eV.

The computational complexity of qubit-ADAPT VQE calculations is tied to the size of the operator pool, which consists of a set of Pauli strings. Naturally, one can construct an operator pool using all the Pauli strings in the qubit representation of fermionic single and double excitation operators. However, the dimension of this UCCSD-compatible pool is usually quite big and scales as $O(N_\theta^4)$. Here we propose a much-simplified operator pool, which consists of Pauli strings from single-excitation and paired double-excitation operators only. The pair excitation involves a pair of electrons with opposite spins, which are initially occupying the same spatial MO, hopping together to another initially unoccupied spatial MO. To further reduce the circuit depth, only one Pauli string is chosen from each qubit representation of the fermionic excitation operator. The qubit representation is a weighted sum of equal-length Pauli strings, and a specific choice of which one of them does not seem to be important in practical calculations reported here. This simplified pool containing operators arising from the UCC ansatz restricted to single and paired double excitation operators (suCCSpD)\cite{Du2020,Du2021} greatly reduces the number of Pauli strings compared to the UCCSD pool. The dimension of this suCCSpD pool scales as $O(N_\theta^2)$. For the $(2, 2)$ $e_g$ impurity model, the pool size reduces from 152 for UCCSD to 56 for suCCSpD, and for the $(3, 3)$ $t_2g$ impurity model it reduces from 828 to 192. The code to perform the above qubit-ADAPT VQE calculations at the state vector level with examples are available in figshare\cite{Du2020}.

Hamiltonian variational ansatz. The Hamiltonian sparsity in the AO basis naturally motivates the application of the Hamiltonian variational ansatz\cite{Peruzzo2014}, which generally takes the form of multi-layer Trotterized annealing-like circuits. While different ways of designing specific HVA forms have been developed, we propose the following ansatz with $L$ layers for the impurity model:

\[
|\Psi(\theta)\rangle = \prod_{j=1}^{L} e^{-\beta_j \hat{h}_j} |\Psi_0\rangle.
\]  

Here $\hat{H} = \sum_{\theta \in \mathcal{M}} \hat{h}_{\theta}$, with $\hat{h}_{\theta}$ being a subgroup of Hamiltonian terms which share the same coefficient and mutually commute. Such ansatz construction aims to differentiate the physical and bath orbitals while retaining the degeneracy information among the orbitals in a systematic way. For each layer of unitaries, we first apply the multi-qubit rotations that are generated by the interacting part of the Hamiltonian, since these act as entangling gates. For the $(M, M)$ impurity model, two reference states have been tried: $|\Psi_0\rangle$ is a simple tensor product state with $M$
physical orbitals fully occupied and the bath orbitals empty; \( |\Psi_0^{(i)}\rangle \) is the ground state of the noninteracting part of \( \hat{H} \), which is equivalent to the one-electron core Hamiltonian in quantum chemistry. We did not find any significant difference between the two choices of reference state in practical simulations of the impurity models. Therefore, only HVA calculations with the reference state \( |\Psi_0^{(i)}\rangle \) are reported here. We adopt the gradient-based Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm as the classical optimizer. Proper parameter initialization for HVA optimization is crucial, as barren plateaus and local energy minima are generally present in the variational energy landscape. In practice, we find that a uniform initialization of the parameters, such as setting all to \( \pi/7 \), overall works well for simulations reported here.

Inspired by the idea of adaptive ansatz generation\(^{13}\), we also tried constructing and optimizing an \( L \)-layer HVA ansatz by adaptively adding layers from 1 to \( L \). Specifically, the calculation starts with optimizing a single-layer ansatz, followed by appending another layer to the ansatz while keeping the first layer at previously obtained optimal angles. The two-layer ansatz is then optimized with the parameters for the new layer initialized randomly or uniformly. The procedure continues with the optimization of \( l \)-layer ansatz leveraging the \((l-1)\)-layer solution until the ansatz reaches \( L \) layers.

Let the number of cost function evaluations for optimizing an \( l \)-layer ansatz be \( N_l^{(1)} \). The total number of function evaluations amounts to \( N^{(2)} = \sum_{l=1}^N N_l^{(2)} \). In practice, we find that the direct optimization of the \( L \)-layer ansatz using a uniform initialization takes \( N^{(1)} \) function evaluations with \( N^{(1)} \sim N_1^{(2)} < N_l^{(2)} \), and reaches the same accuracy. Starting with \( L \) layers is therefore more efficient than growing the ansatz layer by layer.

Intuitively, this can be related to the fact that successive HVA optimization introduces discontinuities in the variational path toward the ground state whenever a new layer of unitaries is added. Since the energy gradient associated with new variational parameters that are initialized to zero (for continuity) vanishes (see the “Methods” section), they have to be initialized away from zero. In other words, the \((l-1)\)-layer HVA solution is not a good starting point for the optimization of the \( l \)-layer ansatz. The open source code to perform the above HVA calculations at the state vector level with examples are available in figshare\(^{54}\).

Hamiltonian commutator pool. It has been demonstrated that the qubit-ADAPT VQE in the MO basis outperforms VQE-UCCSD calculations regarding circuit complexity and numerical accuracy\(^{13,16}\). Motivation by this observation, we compare the corresponding qubit-ADAPT VQE with a Hamiltonian-compatible pool in AO basis and HVA calculations. Following HVA, we choose the simple tensor product state \( |\Psi_0\rangle \) as the reference state. In the qubit-ADAPT step, the energy gradient criterion \( g_0 = 21\text{Im}[\langle \Psi[\theta] | \hat{P} \hat{H} | \Psi[\theta]\rangle] \) to append a new unitary generated by \( \hat{P} \) vanishes due to symmetry with \( |\Psi[\theta]\rangle \), if the number of Pauli-Y operators in the Pauli string \( P \) is even\(^{13,35}\). This can be simply shown by the following argument. Because the impurity model in this study respects time-reversal symmetry and spin-flip \((Z_2)\) symmetry, both Hamiltonian \( \hat{H} \) and wavefunction are real \((\hat{H} = \hat{H}^*, \Psi[\theta] = \Psi[\theta]^*)\). The Pauli string \( \hat{P} \) is also real \((\hat{P} = \hat{P}^\dagger )\) if it has an even number of Pauli-Y operators. Consequently, the expectation value of \( \langle \Psi[\theta] | \hat{P} \hat{H} | \Psi[\theta]\rangle \) is real and \( g_0 \) vanishes if the associated generator \( \hat{P} \) has an even number of Pauli-Y operators.

By construction, the sUCCSpD pool consists of Pauli strings of an odd number of \( Y \)’s. However, the Hamiltonian of the impurity models studied here is all real. Consequently, all the Pauli strings in the qubit representation of the Hamiltonian contain an even number of \( Y \)’s, which excludes the option of directly constructing the operator pool from the Hamiltonian operators. Nevertheless, the practical usefulness of HVA implies that the Hamiltonian-like pool can be constructed by commuting the Hamiltonian terms, which we call the Hamiltonian commutator (HC) pool \( \mathcal{P}_{\text{HC}} \):

Mathematically, \( \mathcal{P}_{\text{HC}} \) is constructed in the following manner:

\[
\mathcal{P}_{\text{HC}} = \{ \hat{H} [\hat{P}, \hat{P}^\dagger] | \hat{P}, \hat{P}^\dagger \in \mathcal{P}_H \}, \quad \text{and } N_Y(\hat{P}, \hat{P}^\dagger)(\text{mod } 2) = 1,
\]

Here \( \mathcal{P}_H \) is the set of Pauli strings \( \{\hat{P}\} \) present in the qubit representation of Hamiltonian \( \hat{H} = \sum_i \omega_i \hat{P}_i \), \( N_Y(\hat{P}) \) counts the number of \( Y \) operators in the Pauli string \( \hat{P} \). Therefore, the size of \( \mathcal{P}_{\text{HC}} \) can scale as \( N_Y^{10} \), where \( N_Y \) is the total number of Hamiltonian terms. Clearly, the pool \( \mathcal{P}_{\text{HC}} \) should only be applied to sparse Hamiltonian systems. The dimension of the HC pool is 56 for the \( e_x \) impurity model, and 192 for the \( e_z \) model.

Quantum circuit implementation. Performing a calculation on a quantum computer always needs to deal with the presence of noise. Even for ideal fault-tolerant quantum computers, quantum sampling (or shot) noise is present due to a finite number of measurements that are used to estimate expectation values. The current noisy quantum devices exhibit additional noise originating from qubit relaxation and dephasing as well as hardware imperfections when implementing unitary gate operations. In this subsection, we describe several techniques adopted in our simulations to most efficiently use the available quantum resources and stabilize the calculations against sampling noise. We discuss how to mitigate gate noise in the final subsection.

Measurement circuit reduction. The quantum circuit implementation for VQE and its adaptive version amounts to the direct measurement of the Hamiltonian as a weighted sum of Pauli string expectation values, \( \langle \hat{H} \rangle = \sum_i \omega_i \langle \hat{P}_i \rangle \), with respect to parametrized circuits \( U[\theta] \). Here, \( \hat{H} = \sum_i \omega_i \hat{P}_i \) is the Hamiltonian in qubit representation. Because the number of shots (or repeated measurements) scales with the desired precision \( \epsilon \) as \( N_{\text{sh}} \propto 1/\epsilon^2 \) due to the central limit theorem, \( N_{\text{sh}} \) is often huge in practical calculations. Therefore, it is desirable to group the Pauli strings into mutually commuting sets such that the number of distinct measurement circuits is reduced to a minimum. Indeed, many techniques to achieve such measurement reduction have been developed\(^{56–61}\). In this work, we adopt the measurement reduction strategy based on the Hamiltonian integral factorization\(^{61}\), which shows a favorable linear system-size scaling of the number of distinct measurement circuits and embraces a diagonal representation for the operators to be measured.

Specifically, we transform the physical subsystem Hamiltonian as follows:

\[
\hat{H}_S = \sum_{\alpha \beta} \tilde{\epsilon}_{\alpha \beta} \sigma^{\alpha \beta \dagger} \sigma^{\alpha \beta} + \frac{1}{4} \sum_{\alpha \beta \gamma} \sum_{\sigma \delta} V_{\alpha \beta \gamma \delta} \sigma^{\alpha \beta \dagger} \sigma^{\beta \gamma \dagger} \sigma^{\gamma \delta} \sigma^{\delta \alpha}, \tag{11}
\]

with \( \tilde{\epsilon}_{\alpha \beta} = \epsilon_{\alpha \beta} - \frac{1}{4} \sum_{\gamma} V_{\alpha \beta \gamma \gamma} \). A typical way to simplify the measurement of the two-body terms \( \hat{H}_S^{(2)} \) in Eq. (11) is to perform nested matrix factorization for the Coulomb \( V \) tensor.

Namely, we first rewrite \( \hat{H}_S^{(2)} \) in the following factorized form by diagonalizing the real symmetric positive semidefinite
supermatrix $V_{apf_1f_2}$:

$$ \hat{\mathcal{H}}^{(2)}_S = \frac{1}{2} \sum_{l=1}^{L} \sum_{ap} \sum_{\sigma} \left( \epsilon_{ap}^f \hat{c}_a^\dagger \hat{c}_{p\sigma} \right)^2. $$

(12)

Here $l$ runs through the $L$ positive eigenvalues of the supermatrix $V$, and the $lh$ component of the auxiliary tensor $L$ is obtained by multiplying the $lh$ eigenvector with the square root of $lh$ positive eigenvalue. Each tensor component, $\epsilon_{ap}^f$, which is a real symmetric matrix, is subsequently diagonalized to reach the following decomposition:

$$ \sum_{apf_1f_2} \epsilon_{ap}^{f_1f_2} \hat{c}_a^\dagger \hat{c}_{p\sigma} = \sum_{m=1}^{M_l} \lambda_{m}^{l} \sum_{ap} \hat{c}_{a\sigma}^{\dagger} \hat{c}_{m}^{l} + \sum_{m=1}^{M_N} \sum_{a} \lambda_{m}^{N} \hat{c}_{a\sigma}^{\dagger} \hat{c}_{m}^{N}. $$

(13)

Here, we have defined $\epsilon_{ap}^{f_1f_2} = \sum_{m=1}^{M_l} \lambda_{m}^{l} \sum_{ap} \epsilon_{ap}^{f_1f_2} U_{ap}^{f_1f_2} \hat{c}_m$. The index $m$ goes through the $M_l$ nonzero eigenvalues $\lambda_{m}^{l}$ and associated eigenvectors $\hat{U}_{ap}^{f_1f_2}$, which determines the single-particle basis transformation for the $lh$ component. The whole embedding Hamiltonian of Eq. (1) can then be cast into the following doubly-factorized form with a unitary transformation similar to Eq. (13) for the one-body part:

$$ \hat{\mathcal{H}} = \sum_{m=1}^{M_l} \sum_{a} \sum_{\sigma} \lambda_{m}^{l} \hat{c}_{a\sigma}^{\dagger} \hat{c}_{m}^{l} + \sum_{m=1}^{M_N} \sum_{a} \lambda_{m}^{N} \hat{c}_{a\sigma}^{\dagger} \hat{c}_{m}^{N}, $$

(14)

which is composed of $L + 1$ groups characterized by unique single-particle basis transformations $\{\epsilon_{ap}^{f_1f_2}\}$, including one from the single-electron component. This form allows efficient measurement of the Hamiltonian expectation value using $L + 1 \propto O(N)$ distinct circuits for a generic quantum chemistry problem with a single-particle basis dimension given by $N$.

The expectation value of $\hat{\mathcal{H}}$ is obtained by measuring each group $l$ independently in the variational state $|\Psi(\theta)\rangle$. The variational state is transformed to the same representation used in the $lh$ group by applying a series of Givens rotations, $\{\theta_{apc}, \theta_{apc}, \theta_{ap}^\dagger\}$, with the set of $\{\theta_{apc}\}$ determined by the single-particle transformation matrix $\{\epsilon_{ap}^{f_1f_2}\}$. Here $\mu$ and $\nu$ are generic indices for physical and bath orbital sites. Therefore, the number of distinct measurement circuits is $N_c = L + 1$. As an example, we have $N_c = 4$ for $e_g$ model. We refer to the "Methods" section for further details.

In practice, it is advantageous to isolate the one-body and two-body terms that contain only density operators before the double factorization procedure, because they are already in a diagonal representation. For the $e_g$ model we have carried out the double-factorization with explicit calculations in the "Methods" section and we ultimately find $N_c = 3$ for the $e_g$ model. This can be compared with the Hamiltonian measurement procedure using the mutual qubit-wise commuting groups: operators that commute with respect to every qubit site are placed in the same group. This commuting Pauli approach generally needs $N_c \propto O(N^3)$ distinct circuits for Hamiltonian measurement. And for the $e_g$ model, it requires $N_c = 5$.

**Noise-resilient optimization.** Although classical optimization approaches such as BFGS, which rely on a computation of the energy gradient, are effective, they rely on very accurate cost function evaluations. Because of the inherent noise in quantum computing, optimization algorithms that are robust to cost function noise are highly desirable. In the noisy quantum simulations reported here, we adopt two optimization techniques that are more tolerant to noise than BFGS: sequential minimal optimization (SMO)\(^62\) and Adadelta\(^63\). Because of their similar performance in the noisy simulations, we only discuss SMO in the main text and leave the discussions of Adadelta in the "Methods" section.

SMO is the first technique we use for our noisy quantum simulations. Tailored to the qubit-ADAPT ansatz of Eq. (8) where each variational parameter is associated with a single Pauli string generator, the optimization consists of $N_{\text{var}}$ sweeps of sequential single parameter minimization of the cost function. At a specific optimization step with varying parameter $\theta$, by keeping others fixed, the cost function has a simple form of $a \cos(2\theta - b) + c$, with the optimal $\theta = b/2$ if $a < 0$ and $(b + n)/2$ otherwise. To determine the parameters $a$, $b$, and $c$, one requires knowledge of function values for at least three mesh points in the range of $[-n/2, n/2)$. In practice, we use eight uniformly spaced mesh points to better mitigate the effect of noise in the cost function. Consequently, least square fitting is used to determine the values of $a$, $b$, and $c$. In SMO calculations, we use the number of sweeps as the parameter to control the convergence, which we set to $N_{\text{var}} = 40$. Alternative control parameters, such as energy and gradient, usually are required to be evaluated at higher precision, which can be challenging and introduce additional quantum computation overhead.

In this work, we perform noisy simulations with classical optimizations that include sampling noise due to a finite number of measurements or shots ($N_{\text{shot}}$) as well as both sampling and gate noise. The purpose is to investigate the performance of the qubit-ADAPT algorithm in the presence of sampling and gate noise and to separate the effects of sampling noise, which is controlled by a single parameter $N_{\text{sh}}$ from the effect of gate noise. The code with the circuit implementation of qubit-ADAPT VQE with examples on QASM simulator and quantum hardware are available at github\(^64\).

**Statevector simulations.** In this section, we present numerical simulation results using a statevector simulator, which is equivalent to a fault-tolerant quantum computer with an infinite number of measurements ($N_{\text{sh}} = \infty$). Figure 3 shows the ground state energy calculations of the $(2, 2)$ $e_g$ and $(3, 3)$ $t_{eg}$ impurity models using VQE-HVA as well as qubit-ADAPT VQE with sUCCSpD and HC pools. The reference UCCSD energy is 0.029 higher than the exact ground state energy $E_{\text{UGS}}$ for the $e_g$ model and 0.128 higher for the $t_{eg}$ model. This implies that both models are in the strong electron correlation region. For calculations of the $e_g$ model, the energy converges below $10^{-5}$ with $N_{\text{var}} = 20$ variational parameters for VQE-HVA, $N_{\text{var}} = 59$ for ADAPT-sUCCSpD, and $N_{\text{var}} = 31$ for ADAPT-HC. Although the qubit-ADAPT VQE calculation on a statevector simulator is in principle deterministic, the operator selection from a predefined operator pool can introduce some randomness due to the numerical accuracy and near degeneracy of scores (i.e., the associated gradient components) for some operators. As a result, the converged $N_{\text{var}}$ can slightly change by about one between runs.

As a simple estimation of the circuit complexity for NISQ devices, we provide the number of CNOT gates $N_{\text{cnot}}$ assuming full qubit connectivity, which can be realized in trapped ion systems. The converged circuit has $N_{\text{cnot}} = 288$ for VQE-HVA, $N_{\text{cnot}} = 292$ for ADAPT-sUCCSpD, and $N_{\text{cnot}} = 150$ for ADAPT-HC. As a reference, the UCCSD ansatz has $N_{\text{cnot}} = 26$ and $N_{\text{cnot}} = 1096$. The HVA calculation converges with the smallest number of variational parameters, but the number of CNOT gates ($N_{\text{cnot}}$) is in between that of ADAPT-HC and ADAPT-sUCCSpD because each variational parameter in HVA is associated with a generator composed of a weighted sum of Pauli strings. The ADAPT-HC calculation starts from a reference state $|\psi_0\rangle$, a simple tensor product state on an AO basis, with energy higher than the HF reference state used by ADAPT-sUCCSpD, yet ADAPT-HC converges faster to the ground state. In fact, the initial state
fidelity, defined as $\mathcal{F} \equiv \langle \Psi_0 | \Psi_{\text{GS}} \rangle^2$, is 0.19 for ADAPT-HC, compared with 0.76 for ADAPT-sUCCSpD. Therefore, the final ansatz complexity does not show a simple positive correlation with the state fidelity, which implies that both the Hamiltonian structure and operator pool are determining factors.

Compared with ADAPT-sUCCSpD, the advantage of ADAPT-HC becomes more prominent when applied to the $t_2g$ model. To reach energy convergence below $10^{-5}$, ADAPT-HC needs $N_d = 270$ parameters and $N_{cx} = 2052$ CNOTs, while ADAPT-sUCCSpD requires as many as $N_d = 1020$ parameters and $N_{cx} = 3066$ CNOTs. For reference, the UCCSD ansatz has $N_d = 117$ parameters and $N_{cx} = 2400$ CNOTs. The HVA calculation is carried out with up to $L = 10$ layers, which amounts to $N_d = 70$ and $N_{cx} = 2420$, and the energy converges close to $10^{-6}$.

We emphasize that strong electron correlation effects are present in our chosen model that lies deep in the bad metallic state\cite{18,49}. This state cannot be accurately captured within a mean-field description and hence requires the application of an appreciable number of unitary gates to the reference state. Generally, the circuit depth of a variational ansatz is tied to both the complexity of the problem (i.e. the complexity of the ground state wavefunction) and the desired state fidelity. As shown in Fig. 4, when we require a state fidelity close to 99.9% or an energy error close to 0.001, which is typically necessary for practical calculations, one observes a sharp rise in $N_d$ when the system is tuned from the weak correlation ($U < 1$) to the strong correlation ($U > 2$) regime by increasing Hubbard $U$.

Simulations with shot noise. The ADAPT VQE calculations are often reported at the statevector level, and a systematic study including the effect of noise is not yet available\cite{13,16,65,67}. Here we present qubit-ADAPT VQE calculations of the $(2, 2)$ $t_2g$ model including shot noise.

Figure 5 shows the representative convergence behavior of the qubit-ADAPT energy with an increasing number of variational parameters $N_d$ calculated using different numbers of shots per observable measurement: Fig. 5a is for $N_{sh} = 2^{12}$, and Fig. 5b is for $N_{sh} = 2^{16}$. We use SMO for the classical optimization. The adaptive ansatz energy $E$ overall decreases as the circuit grow and more variational parameters are used. The energy uncertainty is tied to the number of shots $N_{sh}$. The energy spread roughly reduces by a factor of 4 when $N_{sh}$ increases from $2^{12}$ to $2^{16}$, consistent with the 16-fold increase in $N_{sh}$ due to the central limit theorem.

The energy points shown include not only the final SMO optimized energies of the qubit-ADAPT ansatz with $N_d$ parameters but also the intermediate energies after each of the $N_{sw} = 40$ sweeps during SMO optimizations to provide more detailed convergence information. The above-reported $N_{sh}$ is referred to as measurements for SMO optimizations. At the operator screening step of the qubit-ADAPT calculation to expand the ansatz by appending an additional optimal unitary, we fix $N_{sh} = 2^{16}$ shots for energy evaluations in all cases and determine the energy gradient by the parameter-shift rule\cite{68}.

To further assess the quality of the qubit-ADAPT ansatz obtained in these QASM simulations, we plot in Fig. 5c the ansatz

---

**Fig. 3 Energy convergence of variational quantum eigensolver (VQE) calculations with four types of ansätze.** Panels a, b show the energy difference between the variational and the exact ground state energy $E_{\text{GS}}$ as a function of a number of variational parameters $N_d$. Panels c, d show the energy difference versus the number of CNOT gates $N_{cx}$. Panels a, c are for the degenerate ($N_d = 2, N_{cx} = 2$) impurity model and panels b, d correspond to the $(3, 3) t_2g$ impurity model. VQE calculations are reported with fixed Hamiltonian variational ansatz (HVA, orange dashed line) and unitary coupled cluster ansatz with single and double excitations (UCCSD, black cross) as well as with adaptive ansätze constructed from a simplified unitary coupled cluster pool with single and paired double excitation operators (sUCCSpD, black line) and a Hamiltonian commutator pool (HC, sky blue line). Here $N_{cx}$ is estimated according to each multi-qubit rotation gate with a Pauli string generator $P$ of lengthcontributing $2(l-1)$ CNOT gates, which assumes a full qubit connection. The Hamiltonian parameters are $\epsilon = -9.8\langle 12.7 \rangle, \lambda = 0.3\langle 0.1 \rangle, D = -0.3\langle -0.3 \rangle$ with the same Hubbard $U = 7$ for the $e_g$ ($t_2g$) model, corresponding to the correlated bad metallic regime. The energy unit is the half-band width $D$ of the noninteracting DOS for the multi-band lattice model (see Fig. 1).
energies evaluated using a statevector simulator at the end of each noisy SMO optimization. The four solid curves are calculated using the variational parameters that are obtained by QASM optimizations with different numbers of shots $N_{sh}$ as indicated and noiseless optimization results are shown for comparison as dashed line. While there is no clear order of the energies during the early stages of the simulation, the final convergence is consistently improved with more shots. Specifically, the error converges close to and below $10^{-3}$ for $N_{sh} = 2^{14}$ and $2^{16}$ and the fidelity $f$ improves beyond 99.9%. The associated single-particle density matrix elements also converge to an accuracy better than $10^{-2}$.

Similar QASM simulations of qubit-ADAPT VQE have been performed using the Adadelta optimizer, as specified in the “Methods” section. Generally, we find the numerical results and the dependence on the number of shots to be comparable to SMO. Compared with SMO, Adadelta can potentially take advantage of multiple QPUs by evaluating the gradient vector in parallel.

Discussion of optimal pool size. One important factor determining the computational load of qubit-ADAPT VQE calculations is the size of the operator pool $N_{p}$. One simple strategy to reduce $N_{p}$ is to strip off Pauli $Z$'s in the pool of operators because they contribute negligibly to the ground state energy as pointed out in refs. 16,66. This reduces $N_{p}$ of the Hamiltonian commutator (HC) pool from 56 to 16 for the $e_{q}$ model, and from 192 to 60 for the $t_{36}$ model, due to a large degeneracy. Furthermore, some qualitative guidance has been laid out in the literature to construct a minimal complete pool (MCP) of size $2(N_{q} - 1)$16,69, where $N_{q}$ is the number of qubits. Indeed, we find that an MCP can be constructed using a subset of operators in the HC pool.

We discover a dichotomy that the reduction of the pool size can potentially make the optimization of the qubit-ADAPT ansatz more challenging, especially in the presence of noise. Figure 6 compares qubit-ADAPT calculations using three different pool sizes of dimensions 56, 16, and 10, which were introduced above. Figure 6a shows the qubit-ADAPT energies with increasing $N_{p}$ from statevector simulations of the $e_{q}$ model using the three pools. All the simulations converge with $31 \times 10^{3}$ shots. The details of the convergence rate of the three runs differ significantly. When the pool dimension decreases, the region of $N_{p}$ with minimal energy change expands, as seen by the almost flat segments of the curves of Fig. 6a. The minimal energy gain implies that small noise in the cost function evaluation could deteriorate the parameter optimization.

Indeed as shown in Fig. 6b, the qubit-ADAPT energy from noisy simulation converges slower as the pool size decreases. The flat segments in the energy curves become more evident owing to
the stochastic energy errors. We further analyze the quality of the qubit-ADAPT ansatz by evaluating the energy at optimal angles obtained in noisy simulations, as plotted in Fig. 6c. The energy difference is 0.001, 0.027, 0.135 at \( N_p = 31 \) where the statevector simulation converges, and 0.006, 0.001, 0.005 at the end of \( N_p = 40 \) for calculations with pools of size 56, 16, and 10, respectively.

Our analysis clearly shows the strikingly distinct convergence behaviors of qubit-ADAPT calculations using different complete operator pools in the presence of sampling noise. This indicates that the optimal pool in practical calculations can be a trade-off between choosing a small pool size and guaranteeing sufficient connectivity of the operators in the pool.

Simulations with noise models. Besides the inherent sampling noise in quantum computing, NISQ hardware is subject to various other error effects. These include coherent errors due to imperfect gate operations as well as stochastic errors due to qubit decoherence, dephasing, and relaxation. Here, we perform a preliminary investigation of the impact of hardware imperfections on qubit-ADAPT VQE calculations by adopting a realistic decoherence noise model proposed by Kandala et al. in ref. 3. The model includes an amplitude-damping channel \( \rho \rightarrow \sum_{i=1}^{2} E_i^\rho E_i^{\dagger} \) and a dephasing channel \( \rho \rightarrow \sum_{i=1}^{2} E_i^\rho E_i^{\dagger} \). These act on the qubit density matrix following each single-qubit or two-qubit gate operation. The Kraus operators are given as:

\[
E_1 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p^d} \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & \sqrt{p^d} \\ 0 & 0 \end{pmatrix}.
\]

\[
E_1^d = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p^d} \end{pmatrix}, \quad E_2^d = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{p^d} \end{pmatrix}.
\]

The error rates \( p^d = 1 - e^{-\tau/T_1} \) and \( p^d = 1 - e^{-2\tau/T_2} \) are determined by the gate time \( \tau \), the qubit relaxation time \( T_1 \) and the dephasing time \( T_2 = 2T_1T_2/(2T_1-T_2) \), where \( T_2 \) is the qubit coherence time. For the sake of simplicity of the analysis, we choose a uniform single-qubit gate error rate \( p^d_1 = p^d_2 = p_1 = 10^{-4} \), which is close to the error found in current hardware. We also assume a uniform two-qubit error rate \( p^d_1 = p^d_2 = p_2 \) that we vary between \( 10^{-4} \) and \( 10^{-2} \), in order to study the impact of two-qubit noise on the VQE optimization.

Figure 7a shows a typical qubit-ADAPT energy curve \( E-E_{\text{GS}} \) during optimization as a function of the number of variational parameters \( N_p \) obtained in noisy simulations with \( p_2 = 10^{-2}, 10^{-3}, \) and \( 10^{-4} \). Here, \( E_{\text{GS}} \) is the exact ground state energy. The results with only single-qubit noise are also shown for reference. Figure 7b contains the associated exact energies for the ansatz states, which we obtain by evaluating the VQE ansatz on a statevector simulator.

For \( p_2 = 10^{-2} \), which represents the current hardware noise level, the noisy energy increases with \( N_p \), indicating that the error rate is too large to get reliable energy estimation. Nevertheless, as shown in the corresponding statevector analysis in Fig. 7b, one still observes a sizable energy reduction in the early stage of the optimization. The evaluated ansatz state fidelity is found to improve from 0.19 in the initial state to about 0.70 with \( 4 < N_p < 9 \). When further increasing \( N_p \), however, the statevector ansatz energy shows an upward trend due to noise accumulation, signifying a failure of the noisy optimization. For a smaller error rate, \( p_2 = 10^{-3} \), which was demonstrated recently with the IBM Falcon device70, the noisy energy initially decreases and reaches a minimum near \( N_p = 7 \). This is again followed by an upturn as the number of variational parameters \( N_p \) grows. On the other hand, the statevector simulator shows a clear continuous energy improvement up to \( N_p = 25 \), followed by saturation with small fluctuations. We find the ansatz state fidelity saturates near 0.97. Similar observations apply to the noisy simulations with other two-qubit error rates. The statevector analysis shows that the energy converges at an error \( \approx 3 \times 10^{-3} \) with a fidelity \( \approx 0.9997 \) for \( p_2 = 10^{-4} \). When including only single-qubit errors, we find an error \( \approx 1 \times 10^{-3} \) with a fidelity \( \approx 0.9992 \).

The observed improvement of the ansatz (revealed using statevector analysis), even though the noisy energy expectation value increases, is intriguing. This effect is most clearly seen in results for \( p = 10^{-3} \) between \( 7 \leq N_p < 25 \). It demonstrates the robustness of VQE to certain types of noise effects and can be rationalized as follows. Assuming for simplicity a global depolarizing error channel, we can relate the expectation value of an observable \( \langle O \rangle \) with respect to a noisy density matrix to the noiseless result \( \langle O \rangle = (1-p)\langle O \rangle + p \text{Tr}[O] \). Since any observable can be shifted to be traceless \( \text{Tr}[O] = 0 \), \( \langle O \rangle \) is equivalent to \( \langle O \rangle \) up to a constant scaling factor. The noise thus only rescales the energy landscape of the variational ansatz and maintains the optimal parameters. The fact that we find the
The converged qubit-ADAPT ansatz used for the ground state energy is obtained with the HC pool using statevector simulations. The converged state at the end of VQE calculations to ensure accurate observable measurements.

To reduce the noise in the cost function measurement, it is essential to utilize a range of error mitigation techniques. We employ the standard readout error mitigation using the full confusion matrix approach, as implemented in Qiskit. The adopted measurement circuits based on Hamiltonian integral factorization also allow convenient symmetry detection and filtering with respect to how well the ansatz preserves the total electron number $N_e = 4$ and total spin $z$-projection $S_z = 0$. The gate error is mitigated using zero noise extrapolation (ZNE) with Richardson second-order polynomial inference. The noise scale factor increases from 1 to 2 and 3 for each measurement circuit by local random unitary folding following the implementation in Mitiq. Because of the random gate folding and the stochastic SWAP mapping during transpilation to native gates, we perform ten runs for each measurement circuit at each noise level to smooth out the nondeterministic effects with averaging. For each run, we apply $N_{sh} = 2^{14}$ shots for the measurements.

Figure 8a shows the Richardson extrapolation for the ground state energy with measured points at noise scale factors $\lambda = 1, 2, 3$, taking all 10 runs for each $\lambda$ into account. The estimated energy has an absolute error $\Delta(E) = 0.6 \pm 1.4$ compared with the exact result indicated by the horizontal dashed line. This corresponds to a relative error of 3%. The standard deviation is obtained by fitting the sample points with a second-order polynomial using the SciPy function curve_fit which takes both the mean values and standard deviations into account. In the postprocessing for the mean value of the energy cost function from statistical samplings, we first apply readout calibration, followed by symmetry filtering which discards the configurations with total electron number $N_e \neq 4$ or total spin $S_z \neq 0$. We observe that the ten runs can be divided into two groups based on the average $N_z$ and $S_z$ evaluated before symmetry filtering, as shown in Fig. 8c and d. A subgroup of five runs denoted by square symbols has much less bias away from the correct conserved quantum numbers $N_e = 4$ and $S_z = 0$ than the other five runs shown as circles. A more accurate ground state energy can be obtained when restricting to this optimal subgroup, as shown in Fig. 8b. The estimated energy error reduces significantly to $\Delta(E) = 0.1 \pm 0.2$, with a relative error of 0.7%.

In the above calculations on QPU, the circuits are transpiled into the basis gates of ibmq_casablanca device using the qubit layout and coupling map illustrated in the inset of Fig. 8a. Due to the limited qubit connectivity between nearest neighbors, each of the three transpiled measurement circuits for the $e_g$ model contains about 350 CNOT gates, which amounts to over two-fold increase compared to about 150 CNOTs without qubit swapping. Therefore, we also benchmark the calculations on other types of QPUs with full qubit connectivity such as trapped-ion devices. As an initial reference, we perform an energy estimation with the same ansatz on Quantumium’s trapped-ion Honeywell System Model H1-2. The transpiled circuits have about 150 two-qubit ZZMax gates as expected. Due to limited access to the device, we apply only $N_{sh} = 450$ shots per circuit for the measurements without utilizing any error mitigation. The energy thus obtained is $-17.6 \pm 2$, which should be compared with data points in Fig. 8a at a scale factor 1, and is found to be located near the lower end of that range. Here the error bar is estimated using multiple runs of simulations with the associated system Model H1-2 emulator (H1-2E) including a realistic noise model.

Estimating ground state energy on NISQ devices. As a further step to benchmark the realistic noise effect on qubit-ADAPT VQE calculations of the multi-orbital quantum impurity models, we measure the Hamiltonian expectation value of the $e_g$ model with a converged qubit-ADAPT ansatz on the IBM quantum device ibmq_casablanca. The ansatz with optimal parameters is obtained with the HC pool using statevector simulations. The converged qubit-ADAPT ansatz used for the ground state energy estimate has 32 parameters, and the associated 32 generators for multi-qubit unitary gates are listed in the “Methods” section.

To reduce the noise in the cost function measurement, it is essential to utilize a range of error mitigation techniques. We employ the standard readout error mitigation using the full confusion matrix approach, as implemented in Qiskit. The adopted measurement circuits based on Hamiltonian integral factorization also allow convenient symmetry detection and filtering with respect to how well the ansatz preserves the total electron number $N_e = 4$ and total spin $z$-projection $S_z = 0$. The gate error is mitigated using zero noise extrapolation (ZNE) with Richardson second-order polynomial inference. The noise scale factor increases from 1 to 2 and 3 for each measurement circuit by local random unitary folding following the implementation in Mitiq. Because of the random gate folding and the stochastic SWAP mapping during transpilation to native gates, we perform ten runs for each measurement circuit at each noise level to smooth out the nondeterministic effects with averaging. For each run, we apply $N_{sh} = 2^{14}$ shots for the measurements.

Figure 7 Noisy qubit adaptive derivative-assembled pseudo-trotter (ADAPT) simulations and analysis of $e_g$ model. a Difference between the exact ground state energy $E_{GS}$ and qubit-ADAPT noisy simulation results with a uniform two-qubit gate error rate $p_2 = 10^{-1}$ (black line), $10^{-3}$ (orange line), $10^{-4}$ (sky blue line) and 0 (bluish green line). We use a uniform single-qubit error rate $p_1 = 10^{-4}$ and $N_{sh} = 2^{16}$ shots per measurement circuit. b Energy differences evaluated using statevector for the adaptive ansatz obtained in the noisy simulations. The noisy simulations are performed with the Hamiltonian commutator pool of size 56.

Fig. 7 Noisy qubit adaptive derivative-assembled pseudo-trotter (ADAPT) simulations and analysis of $e_g$ model. a Difference between the exact ground state energy $E_{GS}$ and qubit-ADAPT noisy simulation results with a uniform two-qubit gate error rate $p_2 = 10^{-1}$ (black line), $10^{-3}$ (orange line), $10^{-4}$ (sky blue line) and 0 (bluish green line). We use a uniform single-qubit error rate $p_1 = 10^{-4}$ and $N_{sh} = 2^{16}$ shots per measurement circuit. b Energy differences evaluated using statevector for the adaptive ansatz obtained in the noisy simulations. The noisy simulations are performed with the Hamiltonian commutator pool of size 56.

Conclusions
In an effort towards performing hybrid quantum-classical simulations of realistic correlated materials using a quantum embedding approach, we assess the gate depth and accuracy of variational ground state preparation with fixed and adaptive ansätze for two representative interacting multi-orbital, $e_g$ and $t_{2g}$, impurity models. To take advantage of the sparsity of the

$$E - E_{GS}$$

Noisy simulation

Statevector analysis

$$\begin{align*}
E - E_{GS} & = 10^{-2} \\
E - E_{GS} & = 10^{-1} \\
E - E_{GS} & = 10^{-2} \\
E - E_{GS} & = 10^{-3}
\end{align*}$$

$$\begin{align*}
N_{th} & = 2^{14} \\
N_{sh} & = 2^{16}
\end{align*}$$

$$\begin{align*}
\rho_2 & = 10^{-2} \\
\rho_2 & = 10^{-4} \\
\rho_2 & = 10^{-3} \\
\rho_2 & = 0
\end{align*}$$
Hamiltonian in the atomic orbital representation in real space, we consider the HVA ansatz and an adaptive variant in the qubit-encoded atomic orbital basis. An HC pool composed of pairwise commutators of the Hamiltonian terms is developed to allow fair comparison between the qubit-ADAPT and HVA ansatz. For reference, the standard UCCSD and related qubit-ADAPT calculations using UCCSD-compatible pools are also presented. The qubit-ADAPT calculation with an HC pool generally produces the most compact circuit representation with a minimal number of CNOTs in the final converged circuit. The fixed HVA ansatz follows very closely and has the additional advantage of requiring the least variational parameters $N_P$.

To address the effect of quantum shot noise, we report QASM simulations of qubit-ADAPT VQE in the presence of shot noise for different numbers of shots ($N_{sh}$) that allow controlling the stochastic error. For our benchmark, we adopt state-of-the-art techniques such as low-rank tensor factorization to reduce the number of distinct measurement circuits and a noise resilient optimization including sequential minimal optimization and Adadelta. We find a modest number of shots $N_{sh} = 2^{14}$ per measurement circuit can lead to a variational representation of the ground state with fidelity $f > 99.9\%$.

We further discuss ways to simplify the pool operators and reduce the pool size using $e_f$ model as an example. It is pointed out that a minimal complete pool, as defined in refs. 16,69, can be constructed using a subset of the HC pool. While a simplified pool can reduce the quantum computation resource in the adaptive operator screening procedure, it can make classical optimization more complicated, especially in the presence of noise. This suggests both the dimension and connectivity of operators are joint determining factors to design a practically optimal pool.

To assess the effects of realistic noise on VQE calculations of multi-orbital impurity models, we perform qubit-ADAPT VQE calculations with a realistic decoherence noise model that includes amplitude and dephasing error channels. We find the impact of two-qubit errors to dominate over those of single-qubit errors, also since they are larger in NISQ hardware. We report that practically useful results can be obtained for $p_2 = 10^{-3}$, which is closer to current hardware levels. Importantly, we observe that the classical optimization continues to improve the ansatz even in a regime, where the noisy energy expectation value starts to rise. We reveal this behavior by executing the ansatz state on statevector simulators. Such persisting ansatz state improvement masked by noise shows that VQE is robust to certain noise effects and implies that costly error mitigation methods can potentially be reserved for the evaluation of expectation values in the final converged state.

Finally, we measure the energy for a converged qubit-ADAPT ansatz of the $e_f$ model on the ibmq_casablanca QPU and Quantumum’s H1-2 device. Using the results from IBM hardware, we obtain an error of 0.1 (0.7\%) for the total energy by adopting error mitigation techniques such as zero-noise extrapolation, combined with a careful post-selection based on symmetry and the conservation of quantum numbers.

Moving forward, the full qubit-ADAPT VQE calculations of quantum impurity models will be extended from noisy QASM simulations to simulations that include device-specific noise effects beyond our decoherence model and finally to experiments on real hardware. Our study shows that an array of error mitigation techniques, including readout calibration, zero-noise extrapolation, and potentially probabilistic error cancellation, Clifford data regression, and probabilistic
machine-learning-based techniques\textsuperscript{82}, need to be adopted to reach sufficiently accurate results. This is especially important when using VQE as an impurity solver in a quantum embedding approach as sufficiently accurate impurity model results are needed in order to enable the convergence of the classical self-consistency loop. Our results constitute an important step forward in demonstrating high-fidelity ground state preparation of impurity models on quantum devices. This is especially relevant for realizing correlated material simulations through hybrid quantum-classical embedding approaches, where the ground state preparation of a generic \( f \)-electron impurity model consisting of 28 spin-orbital is on the verge of achieving practical quantum advantage\textsuperscript{35}.

**Methods**

**Energy gradient of HVA.** Here we show that the outermost \( \mathbf{t} \) layer gradient component vanishes \( \frac{\delta E}{\delta \mathbf{t}}_{|E=0} = 0 \) for an \( \mathbf{l} \)-layer HVA ansatz

\[
|\Psi_{f}(\mathbf{\theta})\rangle = \Pi_{i=1}^{\mathbf{t}} e^{-\lambda_{ii} f_{i}} |\Psi_{l-1}(\mathbf{\theta})\rangle
\]

Because the system Hamiltonian \( \mathcal{H} \) under study is real due to time-reversal symmetry, HVA is also real by construction. Therefore, \( |\Psi_{l-1}(\mathbf{\theta})\rangle \mathcal{H} |\Psi_{l-1}(\mathbf{\theta})\rangle \) is real, and \( \frac{\delta E}{\delta \mathbf{t}}_{|E=0} \) vanishes. Note that the exactly same reason motivates the development of the HC pool for qubit-ADAPT calculations.

**Hamiltonian factorization of the impurity model.** Here we explain explicitly how the Hamiltonian factorization is obtained using the \( e_{\mathbf{t}} \) model as an example, whose Hamiltonian takes the following specific form:

\[
\mathcal{H} = D \sum_{i=1}^{2} \sum_{\sigma} \left( c_{i \sigma}^\dagger c_{i \sigma} + h.c. \right)
= + \mathcal{J} \left( \mathbf{e}_{1 \mathbf{t} \mathbf{1}} \right)
+ + \mathcal{J} \left( \mathbf{e}_{1 \mathbf{t} \mathbf{1}} + \mathbf{e}_{2 \mathbf{t} \mathbf{1}} \right)
+ + \mathcal{J} \left( \mathbf{e}_{1 \mathbf{t} \mathbf{1}} + \mathbf{e}_{2 \mathbf{t} \mathbf{1}} + \mathbf{e}_{3 \mathbf{t} \mathbf{1}} \right)
+ + \mathcal{J} \left( \mathbf{e}_{1 \mathbf{t} \mathbf{1}} + \mathbf{e}_{2 \mathbf{t} \mathbf{1}} + \mathbf{e}_{3 \mathbf{t} \mathbf{1}} + \mathbf{e}_{4 \mathbf{t} \mathbf{1}} \right)
\]

This can be derived conveniently in the matrix formulation:

\[
\begin{pmatrix}
\mathcal{J} \left( \mathbf{e}_{1 \mathbf{t} \mathbf{1}} \right)
\end{pmatrix}
= \begin{pmatrix}
\mathcal{J} \left( \mathbf{e}_{1 \mathbf{t} \mathbf{1}} \right)
\end{pmatrix}
= \begin{pmatrix}
\mathcal{J} \left( \mathbf{e}_{1 \mathbf{t} \mathbf{1}} \right)
\end{pmatrix}
= \begin{pmatrix}
\mathcal{J} \left( \mathbf{e}_{1 \mathbf{t} \mathbf{1}} \right)
\end{pmatrix}
= \begin{pmatrix}
\mathcal{J} \left( \mathbf{e}_{1 \mathbf{t} \mathbf{1}} \right)
\end{pmatrix}
= \begin{pmatrix}
\mathcal{J} \left( \mathbf{e}_{1 \mathbf{t} \mathbf{1}} \right)
\end{pmatrix}
= \begin{pmatrix}
\mathcal{J} \left( \mathbf{e}_{1 \mathbf{t} \mathbf{1}} \right)
\end{pmatrix}
= \begin{pmatrix}
\mathcal{J} \left( \mathbf{e}_{1 \mathbf{t} \mathbf{1}} \right)
\end{pmatrix}
\]

The pair hopping and spin-flip terms of the second line of Eq. (18) can be rewritten as:

\[
\mathcal{L}^{(1)} = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\]

The above expression is obtained by diagonalizing the Coulomb supermatrix \( V_{\mathbf{ag}}(\nu_{0}) \) with density-density elements set to zero, \( V_{\mathbf{ag}}(\nu_{0}) \equiv 0 \), which gives a single eigenvector associated with nonzero eigenvalue. Following the similar derivation in Eq. (23), the pair hopping and spin-flip terms have the following diagonal representation:

\[
\mathcal{L}^{(1)} = \begin{pmatrix}
\mathbf{e}_{1 \mathbf{t} \mathbf{1}} & \mathbf{e}_{1 \mathbf{t} \mathbf{1}} & \mathbf{e}_{1 \mathbf{t} \mathbf{1}} & \mathbf{e}_{1 \mathbf{t} \mathbf{1}} \\
\mathbf{e}_{1 \mathbf{t} \mathbf{1}} & \mathbf{e}_{1 \mathbf{t} \mathbf{1}} & \mathbf{e}_{1 \mathbf{t} \mathbf{1}} & \mathbf{e}_{1 \mathbf{t} \mathbf{1}} \\
\mathbf{e}_{1 \mathbf{t} \mathbf{1}} & \mathbf{e}_{1 \mathbf{t} \mathbf{1}} & \mathbf{e}_{1 \mathbf{t} \mathbf{1}} & \mathbf{e}_{1 \mathbf{t} \mathbf{1}} \\
\mathbf{e}_{1 \mathbf{t} \mathbf{1}} & \mathbf{e}_{1 \mathbf{t} \mathbf{1}} & \mathbf{e}_{1 \mathbf{t} \mathbf{1}} & \mathbf{e}_{1 \mathbf{t} \mathbf{1}}
\end{pmatrix}
\]

with \( \mathbf{e}_{1 \mathbf{t} \mathbf{1}} \equiv \frac{\mathbf{e}_{1 \mathbf{t} \mathbf{1}}}{\mathbf{e}_{1 \mathbf{t} \mathbf{1}}} \) and

\[
\mathbf{e}_{1 \mathbf{t} \mathbf{1}} = \frac{\mathbf{e}_{1 \mathbf{t} \mathbf{1}}}{\mathbf{e}_{1 \mathbf{t} \mathbf{1}}} = \frac{1}{\sqrt{2}} (\mathbf{e}_{1 \mathbf{t} \mathbf{1}} - \mathbf{e}_{1 \mathbf{t} \mathbf{1}}).
\]

Finally, we can represent the embedding Hamiltonian for \( e_{\mathbf{t}} \) model in the following doubly-factorized form:

\[
\mathcal{H} = D \sum_{\sigma} (-\mathbf{e}_{1 \mathbf{t} \mathbf{1}} + \mathbf{e}_{1 \mathbf{t} \mathbf{1}} + \mathbf{e}_{1 \mathbf{t} \mathbf{1}} + \mathbf{e}_{1 \mathbf{t} \mathbf{1}})
+ j/2 \left( \mathbf{e}_{1 \mathbf{t} \mathbf{1}} + \mathbf{e}_{1 \mathbf{t} \mathbf{1}} + \mathbf{e}_{1 \mathbf{t} \mathbf{1}} + \mathbf{e}_{1 \mathbf{t} \mathbf{1}} \right)
+ U \sum_{\sigma} \mathbf{n}_{1 \mathbf{t} \mathbf{1}} \mathbf{n}_{1 \mathbf{t} \mathbf{1}}
+ j/2 \left( \mathbf{e}_{1 \mathbf{t} \mathbf{1}} + \mathbf{e}_{1 \mathbf{t} \mathbf{1}} + \mathbf{e}_{1 \mathbf{t} \mathbf{1}} + \mathbf{e}_{1 \mathbf{t} \mathbf{1}} \right)
\]

The Hamiltonian integral factorization we find that three distinct measurement circuits are needed for the Hamiltonian expectation value: (i) the diagonal terms in the original atomic orbital basis, (ii) the hybridization terms in the basis of \( e_{\mathbf{t} \mathbf{1}}(22) \), (iii) the pair hopping and spin-flip terms in the basis of \( e_{\mathbf{t} \mathbf{1}}(27) \).

**Quantum simulation with Adadelta optimizer.** In the main text, we reported the qubit-ADAPT VQPE calculation with shots using the SMO optimizer. Here we additionally perform the calculations using the Adadelta optimization method, which is potentially tolerant to cost function errors\textsuperscript{65}. Below we describe the implementation of the algorithm followed by the results.

The algorithm minimizes the cost function along the steepest descent direction in parameter space, with a parameter update at step \( t \) as \( \theta_{t} = \theta_{t-1} - \alpha_{t} \), where \( \alpha_{t} \) is determined from the derivative of the energy function along
Fig. 9 Energy convergence of qubit adaptive derivative-assembled pseudo-trotter (ADAPT) noisy simulations of $e_9$ model with Adadelta optimizer. The difference between the exact ground state energy $E_{GS}$ and qubit-ADAPT noisy simulation results, obtained with number of shots $N_{sh} = 2^{16}$ in panel (a). The energy differences evaluated using statevector for the adaptive ansätze obtained in the noisy simulations with $N_{sh} = 2^{16}$ (orange line) are shown in panel (b). The statevector simulation results ($N_{sh} = \infty$) of the qubit-ADAPT method are also shown in a blue dashed line for reference.

Received: 18 March 2022; Accepted: 17 November 2022; Published online: 04 January 2023

References

1. Aspuru-Guzik, A., Dutoi, A. D., Love, P. J. & Head-Gordon, M. Simulated quantum computation of molecular energies. Science 309, 1704–1707 (2005).
2. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. Nat. Commun. 5, 1–7 (2014).
3. Kandala, A. et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. Nature 549, 242–246 (2017).
4. O’Malley, P. J. et al. Scalable quantum simulation of molecular energies. Phys. Rev. X 6, 031007 (2016).
5. Preskill, J. Quantum computing in the NISQ era and beyond. Quantum 2, 79 (2018).
6. McArdis, S., Endo, S., Aspuru-Guzik, A., Benjamín, S. C. & Yuan, X. Quantum computational chemistry. Rev. Mod. Phys. 92, 015003 (2020).
7. Cerezo, M. et al. Variational quantum algorithms. Nat. Rev. Phys. 3, 625–644 (2021).
8. McClain, J. R., Romero, J., Babbush, R. & Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. New J. Phys. 18, 032032 (2016).
9. Hoffmann, M. R. & Simons, J. A unitary multiconfigurational coupled-cluster method: theory and applications. J. Chem. Phys. 88, 993–1002 (1988).
10. Barlett, R. J., Kucharski, S. A. & Noga, J. Alternative coupled-cluster ansätze ii. the unitary coupled-cluster method. Chem. Phys.Lett. 155, 133–140 (1989).
11. Barlett, R. J. & Musial, M. Coupled-cluster theory in quantum chemistry. Rev. Mod. Phys. 79, 291 (2007).
12. Romero, J. et al. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. Quantum Sci. Technol. 4, 014008 (2018).
13. Grimsley, H. R., Economou, S. E., Barnes, E. & Mayhall, N. J. An adaptive variational algorithm for exact molecular simulations on a quantum computer. Nat. Commun. 10, 3007 (2019).
14. Ryabinkin, I. G., Yen, T.-C., Genin, S. N. & Isaylov, A. F. Qubit coupled cluster method: a systematic approach to quantum chemistry on a quantum computer. J. Chem. Theory Comput. 14, 6317–6326 (2018).
15. Lee, J., Huggins, W. J., Head-Gordon, M. & Whaley, K. B. Generalized unitary variational hybrid quantum-classical algorithms. Nat. Phys. 15, 311–324 (2018).
16. Tang, H. L. et al. Qubit-adapt-vcq: an adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor. PRX Quantum 2, 020310 (2021).
17. Zhang, F. et al. Shallow-circuit variational quantum eigensolver based on symmetry-inspired hilbert space partitioning for quantum chemical calculations. Phys. Rev. Res. 3, 013039 (2021).
18. Gomes, N. et al. Adaptive variational quantum imaginary time evolution approach for ground state preparation. Adv. Quantum Technol. 4, 2100114 (2021).
19. Fedorov, D. A., Peng, B., Govind, N. & Alexeev, Y. VQE method: a short survey and recent developments. Mater. Theory 6, 2 (2022).
20. Tilly, J. et al. The variational quantum eigensolver: a review of methods and best practices. Phys. Rep. 896, 1–128 (2022).
21. Wecker, D., Hastings, M. B. & Troyer, M. Progress towards practical quantum variational algorithms. Phys. Rev. A 92, 042303 (2015).
22. Ho, W. W. & Hsieh, T. H. Efficient variational simulation of non-trivial quantum states. SciPost Phys. 6, 29 (2019).
23. Wiersma, R. et al. Exploring entanglement and optimization within the hamiltonian variational ansatz. PRX Quantum 1, 020319 (2020).
24. Kent, P. R. & Kotliar, G. Toward a predictive theory of correlated materials. Science 361, 348–354 (2018).
25. Yin, Z., Haule, K. & Kotliar, G. Kinetic frustration and the nature of the magnetic and paramagnetic states in iron pnictides and iron chalcogenides. Nat. Mater. 10 12, 932–5 (2011).
26. Georges, A., de Medici, L. & Marzari, N. Strong correlations from Hund’s coupling. Annu. Rev. Condensed Matter Phys. 4, 137–178 (2013).
27. de Medici, L., Giovannietti, G. & Capone, M. Selective Mott physics as a key to iron superconductors. Phys. Rev. Lett. 112 17, 177001 (2014).
28. Sprau, P. O. et al. Discovery of orbital-selective cooperator pairing in FESF. Science 357, 75 – 80 (2017).
29. Georges, A., Kotliar, G., Kuroth, W. & Rozenberg, M. J. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. Rev. Mod. Phys. 68, 13 (1996).
30. Kotliar, G. et al. Electronic structure calculations with dynamical mean-field theory. Rev. Mod. Phys. 78, 885 (2006).
31. Sun, Q. & Chan, G. K.-L. Quantum embedding theories. Acc. Chem. Res. 49, 2705–2712 (2016).
32. Sakurai, R., Mizukami, W. & Shinaoka, H. Hybrid quantum-classical.
33. Bauer, B., Wecker, D., Millis, A. J., Hastings, M. B. & Troyer, M. Hybrid.
34. Knizia, G. & Chan, G. K.-L. Density matrix embedding: a simple alternative to.
35. Kotliar, G. & Ruckenstein, A. E. New functional integral approach to strongly.
36. Yao, Y.-X. Python driver of Gutzwiller quantum-classical embedding.
37. Bünemann, J. & Gebhard, F. Equivalence of Gutzwiller and slave-boson.
38. de’ Medici, L., Mravlje, J. & Georges, A. Janus-faced in field theories for multiband Hubbard models.
39. Abraham, H. et al. Qiskit: an open-source framework for quantum computing.
40. Crawford, O. et al. Efficient quantum measurement of Pauli operators in the presence of finite sampling error.
41. Yen, T.-C. & Izmaylov, A. F. Cartesian subalgebra approach to efficient measurements of quantum observables.
42. Huggins, W. J. et al. Efficient and noise resilient measurements for quantum chemistry on near-term quantum computers.
43. Nakahara, K. M., Fujii, K. & Todo, S. Sequential minimal optimization for quantum-classical hybrid algorithms.
