ARTICLE   OPEN

Check for updates

Multi-state quantum simulations via model-space quantum imaginary time evolution

Takashi Tsuchimochi [1,2 ✉], Yoohee Ryo[3], Siu Chung Tsang[1] and Seiichiro L. Ten-no[1]

We introduce the framework of model space into quantum imaginary time evolution (QITE) to enable stable estimation of ground and excited states using a quantum computer. Model-space QITE (MSQITE) propagates a model space to the exact one by retaining its orthogonality, and hence is able to describe multiple states simultaneously. The quantum Lanczos (QLanczos) algorithm is extended to MSQITE to accelerate the convergence. The present scheme is found to outperform both the standard QLanczos and the recently proposed folded-spectrum QITE in simulating excited states. Moreover, we demonstrate that spin contamination can be effectively removed by shifting the imaginary time propagator, and thus excited states with a particular spin quantum number are efficiently captured without falling into the different spin states that have lower energies. We also investigate how different levels of the unitary approximation employed in MSQITE can affect the results. The effectiveness of the algorithm over QITE is demonstrated by noise simulations for the H_4 model system.

INTRODUCTION

Variational quantum algorithms[1–3] are expected to play a key role on noisy intermediate-scale quantum (NISQ) devices[4]. Especially, variational quantum eigensolver (VQE)[5–10] has attracted much attention for its application to quantum chemistry where quantum entanglement is essential[11,12]. The scope of VQE has extended from ground state simulations of molecular systems[7,9,13] to condensed matters[14–16] and excited states[17–22].

Building on the concept of imaginary time evolution (ITE), which aims to drive an arbitrary initial state to the exact ground state, several variants of quantum algorithms have recently been developed. McArdle and co-workers proposed variational ITE (VITE)[23], which employs a fixed ansatz and then determines the optimal parameters using McLachlan's variational principle. Therefore, VITE can be regarded as an optimizer not only for variational algorithms such as VQE[24], but also for algorithms that employ pre-processed non-Hermitian Hamiltonians such as transcorrelated methods[25–27]. Probabilistic ITE (PITE) exploits measurements to perform the non-unitary operation of ITE on quantum devices probabilistically[28–30]. This can be achieved by introducing one ancilla qubit and embedding the ITE operator acting on an N-qubit system in an $(N+1)$-qubit unitary gate[28]. In principle, ITE can be exactly performed on the N-qubit system by accepting (or discarding) the resultant state if the ancilla qubit is measured to be $|0\rangle$ (or $|1\rangle$). However, the exact PITE generally requires the singular-value-decomposition of the N-qubit Hamiltonian, which hinders its practical applications in chemistry. Furthermore, the success probability decays exponentially with the number of imaginary time steps and the system size N. As such, there have been several proposals to circumvent these problems[29,30].

Yet another scheme, Quantum ITE (QITE)[31], approximates the non-unitary short evolution of ITE by a unitary evolution that is determined by solving a set of linear equations. Therefore, it circumvents the high-dimensional noisy optimizations in variational algorithms, while driving a quantum state towards the ground state at each evolution step. The promise of QITE has been demonstrated experimentally[31–33], and numerous authors have extended the algorithm[34–40]. In our own recent study, a modified equation for the unitary approximation was presented, which enables faster convergence of QITE, thereby reducing the overall quantum resources.

Although QITE is a powerful tool for determining the ground state, there have been fewer developments that aim for obtaining excited states, especially when compared to variational algorithms that have seen a wide variety of developments[17–21,39,41–46]. The reason for this is perhaps that quantum Lanczos diagonalization (QLanczos) is expected to find reasonable excited states by increasing the size of the Krylov subspace[31,32]. However, our recent study showed that the component of excited states encoded in the initial state vanishes with imaginary time β at an exponential rate in general, and is lost in the numerical noise that is caused by the strong linear dependence of the chosen Krylov subspace[39]. This is particularly true if the excited states are separated from the ground state by large energy gaps, i.e., higher energy eigenstates.

Historically, there have been broad interests in obtaining excited states from classical ITE[47–51], and we can gain many insights from them. For instance, to retain the excited state signature throughout the QITE simulation, we followed the work of Booth and Chan[48] and adopted the folded-spectrum propagator $e^{-\beta^2(\hat{H}-\omega)^2}$ in ref. [39], an approach coined FSQITE. It was shown that FSQITE can in principle yield the desired excited states, and its convergence rate can be drastically accelerated with QLanczos. Nevertheless, FSQITE requires to estimate the target energy ω in advance and to treat the Hamiltonian squared $\hat{H}^2$, which can be quite challenging in general.

In this work, we develop the model-space QITE (MSQITE) algorithm to deliver stable and accurate solutions for excited states. MSQITE evolves an orthogonal model space to the complete subspace by ITE, simulating multiple states

¹Graduate School of System Informatics, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe, Hyogo 657-8501, Japan. ²Japan Science and Technology Agency (JST), Precursory Research for Embryonic Science and Technology (PRESTO), 4-1-8 Honcho Kawaguchi, Saitama 332-0012, Japan. ³Graduate School of Science, Technology, and Innovation, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe, Hyogo 657-8501, Japan. ✉email: tsuchimochi@gmail.com

npj

simultaneously. It also improves the behavior and accuracy for the ground states of strongly correlated systems, by directly incorporating important configurations. Because the method has many similarities to QITE and FSQITE, it can be also easily combined with QLanczos.

Furthermore, we present a scheme to deal with spin contamination in MSQITE. The spin quantum number is an essential quantity that characterizes a non-relativistic electronic state. Preserving spin symmetry is important but is more challenging in quantum simulations[52–57] than conserving other symmetries such as point-group symmetry that can be usually constrained by removing the qubits from the simulation[58,59]. It should be easily imagined the problem of spin contamination is exacerbated in excited state calculations because excited states often exhibit more complicated electronic structures than the ground state and thus are prone to spin contamination. In the following, we provide a way to circumvent this difficulty.

As will be seen, there are two different flavors of MSQITE; one uses the same unitary for all states in the model space, and the other employs different unitaries for different states. We investigate how such unitary approximations in the MSQITE algorithm can affect its representability and accuracy, and report difficulties with the former approach.

Since the quantum circuits of both QITE and MSQITE necessarily elongate with imaginary time, their applicability on NISQ computers might be severely limited due to quantum noise. However, it is expected that MSQITE could potentially outperform QITE in many respects, even on NISQ computers, providing not only excited states but also faster convergence to the ground state. We demonstrate the efficacy of MSQITE through noisy simulations using a simple error mitigation protocol.

RESULTS

Model-space QITE

In MSQITE, one prepares an orthogonal subspace that consists of zeroth-order ground and excited states, $\{|\Phi_I\rangle; I = 0, \cdots, n_{\text{states}} - 1\}$, and evolves the entire subspace by the propagator $e^{-\beta \hat{H}}$ (which is Trotterized by a short time step $\Delta\beta$). It is important to note that the imaginary time evolution makes the basis states nonorthogonal, $\langle\Phi_I|e^{-2\Delta\beta\hat{H}}|\Phi_J\rangle \neq 0$, and therefore the orthonormalization of the subspace is necessary. Hence, in MSQITE, we consider the following unitary approximation on the ℓth step:

$$\left|\Phi_I^{(\ell+1)}\right\rangle = \sum_J d_{IJ} e^{-\Delta\beta(\hat{H} - E_J)}\left|\Phi_J^{(\ell)}\right\rangle \approx e^{-i\Delta\beta\hat{A}}\left|\Phi_I^{(\ell)}\right\rangle \quad (1)$$

where $|\Phi_I^{(\ell)}\rangle$ are the Ith state at the ℓth time step, and $\hat{A}$ is a Hermitian operator parameterized by the coefficients $\mathbf{a}$,

$$\hat{A} = \sum_\mu a_\mu \hat{\sigma}_\mu \quad (2)$$

with Pauli strings $\hat{\sigma}_\mu$, which are appropriately chosen[31,34,39]. In Eq. (1), we have intentionally introduced the energy shift $E_J = \langle\Phi_J^{(\ell)}|\hat{H}|\Phi_J^{(\ell)}\rangle$ for convenience. The transformation matrix $\mathbf{d}$ is also introduced to ensure the orthonormality of the time-evolved model space $\{|\Phi_I^{(\ell+1)}\rangle\}$.

This transformation matrix can be defined in infinitely different ways; however, we require $\mathbf{d} \to \mathbf{I}$ (identity matrix) as $\Delta\beta \to 0$, because this would allow us to correctly obtain $|\Phi_I^{(\ell+1)}\rangle \equiv |\Phi_I^{(\ell)}\rangle$. We also wish the change in each state to be minimum at each time step, to be able to "follow" the Ith state between the time steps in order for Eq. (1) to be a meaningful approximation. To this end, we employ the Löwdin symmetric orthonormalization[60,61]. Remarkably, the so-obtained $\mathbf{d}$ is the one that minimizes the distance in the Hilbert space, $\mathbf{d} = \arg\min_{\mathbf{d}} \sum_I \||\Phi_I^{(\ell+1)}\rangle - e^{-\Delta\beta(\hat{H} - E_I)}|\Phi_I^{(\ell)}\rangle\|^{2}$[62,63]. In other words, the property of $|\Phi_I^{(\ell)}\rangle$ is

maximally preserved in $|\Phi_I^{(\ell+1)}\rangle$ on average, and therefore it is expected that different states do not mix strongly. In particular, when the energy shift E_J is introduced, $\mathbf{d}$ is diagonal dominant with all the diagonal elements being equal to one. In the Methods section, we have detailed the Löwdin symmetric orthonormalization procedure in MSQITE and discussed other possibilities for the definition of $\mathbf{d}$.

In MSQITE, $\lim_{\ell\to\infty}|\Phi_I^{(\ell)}\rangle$ may *not* be the exact ground and excited states. Instead, we retain them as a model space basis and express the physical states $|\psi_I\rangle$ as a linear combination of these states, $|\psi_I\rangle = \lim_{\ell\to\infty}\sum_K c_{KI}|\Phi_K^{(\ell)}\rangle$. This corresponds to solving the eigenvalue problem

$$\mathbf{H}^{(\ell)}\mathbf{c} = \mathbf{S}^{(\ell)}\mathbf{c}\mathcal{E} \quad (3)$$

where

$$H_{IJ}^{(\ell)} = \langle\Phi_I^{(\ell)}|\hat{H}|\Phi_J^{(\ell)}\rangle \quad (4)$$

$$S_{IJ}^{(\ell)} = \langle\Phi_I^{(\ell)}|\Phi_J^{(\ell)}\rangle \quad (5)$$

and $\mathcal{E}$ contains the ground and excited state energies in the diagonal. The eigenvalues become the exact energies if the entire model space is propagated appropriately.

Now, we have two approaches to determine the unitary $e^{-i\Delta\beta\hat{A}}$. In the so-called state-specific approach, $\mathbf{a}$ is different for different $|\Phi_I\rangle$ (therefore, we write $\mathbf{a}^I$ and $\hat{A}^I$ to indicate the state dependence). Similarly to QITE[31,39], we minimize the following function

$$F^I(\mathbf{a}^I) = \left\|\sum_J d_{IJ} e^{-\Delta\beta(\hat{H} - E_J)}\left|\Phi_J^{(\ell)}\right\rangle - e^{-i\Delta\beta\hat{A}^I}\left|\Phi_I^{(\ell)}\right\rangle\right\|^2 \quad (6)$$

to the second-order of $\Delta\beta$ for each I. This results in the linear equation

$$\mathbf{M}^I\mathbf{a}^I + \mathbf{b}^I = \mathbf{0} \quad (7)$$

with

$$M_{\mu\nu}^I = 2\text{Re}\left\langle\Phi_I^{(\ell)}\left|\hat{\sigma}_\mu\hat{\sigma}_\nu\right|\Phi_I^{(\ell)}\right\rangle \quad (8)$$

$$b_\mu^I = \text{Im}\left\langle\Phi_I^{(\ell)}\left|[\hat{H}, \hat{\sigma}_\mu]\right|\Phi_I^{(\ell)}\right\rangle + \frac{2}{\Delta\beta}\sum_J d_{JI}\text{Im}\left\langle\Phi_I^{(\ell)}\left|\hat{\sigma}_\mu\right|\Phi_J^{(\ell)}\right\rangle \quad (9)$$

We have provided a detailed derivation in Supplementary Notes.

In contrast, the state-averaged approach uses the same $\mathbf{a}$ and $\hat{A}$ for all the states considered. This can be accomplished by solving

$$\sum_I \mathbf{M}^I\mathbf{a} + \sum_I \mathbf{b}^I = \mathbf{0} \quad (10)$$

Several important considerations have to be made with respect to the above derivations. Eqs. (8) and (9) are essentially the same as those corresponding to QITE[39], except that b_μ^I has the additional second term, which ensures the orthogonality of the model space. For the state-specific method, the model space is not exactly orthogonal, but it is almost so because of this term. Indeed, without the second term, Eq. (3) quickly becomes unsolvable because all elements of $S_{IJ}^{(\ell)}$ tend to become one (i.e., all states in the model space become the ground state). We note that $\frac{d_{IJ}}{\Delta\beta} \to 0$ for $I \neq J$ as $\Delta\beta \to 0$ and the diagonal term will not contribute because σ_μ is Hermitian and thus the expectation value is real; so the second term is stable. The importance of the term is less pronounced in the state-averaged method.

The state-averaged method would be preferred to the state-specific method because the model space $\{|\Phi_I^{(\ell)}\rangle\}$ is guaranteed to be orthogonal (i.e., $S_{IJ}^{(\ell)} = \delta_{IJ}$), and also because its quantum circuit is significantly simpler. However, despite the existence of a

(a)

(b)   **(c)**
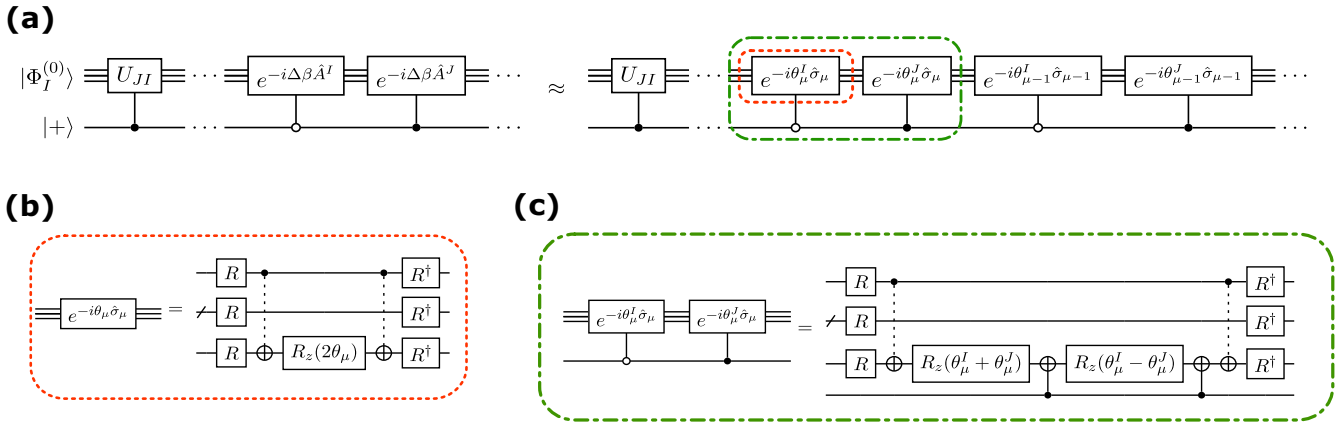
Fig. 1 Quantum circuit for obtaining matrix elements of state-specific MSQITE using the Hadamard test with an ancilla $|+\rangle$. a U_{JI} transforms $|\Phi_I^{(0)}\rangle$ to $|\Phi_J^{(0)}\rangle$. The unitary gates $e^{-i\Delta\beta\hat{A}^I}$ and $e^{-i\Delta\beta\hat{A}^J}$ are each decomposed into Pauli rotations. **b** The uncontrolled version of Pauli rotation is typically implemented by using one-qubit unitary gates (R), and a sequence of CNOT gates, which is abbreviated by the CNOT gate with a dotted line. Together with R_z, they perform $e^{-i\theta_\mu\hat{\sigma}_\mu}$. **c** Two different controlled Pauli rotations by $\theta_\mu^I\hat{\sigma}_\mu$ and $\theta_\mu^J\hat{\sigma}_\mu$ can be summarized to one controlled-R_z.

single unitary $e^{-i\Delta\beta\hat{A}}$ that correctly transforms all the states simultaneously to the desired states, it should be noted that the corresponding Hermitian $\hat{A}$ has to be quite complicated. In practice, because it is desirable to employ a simple $\hat{A}$ in Eq. (2), the representability of the unitary is considerably limited, and therefore the performance of the state-specific MSQITE may not be promising. This is quite similar to an issue recently reported by ref. [64], that the multistate contracted VQE, which minimizes the averaged energy of orthogonal states generated by the same unitary[19], experiences large errors for excited state calculations. Indeed, below, we will show that with the state-specific MSQITE a model space converges to almost the exact one using only single and double excitations in $\hat{A}$, whereas the accuracy of the state-averaged MSQITE is generally quite unsatisfactory and its errors in energy can be substantial especially when the number of states increases.

Quantum Circuit for MSQITE

The algorithmic difference between QITE and MSQITE is that the latter requires the estimation of quantities like $H_{IJ}^{(\ell)}$ for each pair I, J. Whereas the state-averaged MSQITE has a simple quantum circuit because all the states are evolved by the same unitary $e^{-i\Delta\beta\hat{A}}$, one needs the controlled gate for $e^{-i\Delta\beta\hat{A}(\mathbf{a}^I)}$ for the state-specific approach. Figure 1 illustrates how we implement the latter using the Hadamard test. We prepare the state register and an ancilla qubit as $|\Phi_I^{(0)}\rangle$ and $|+\rangle$, which controls U_{JI}, $e^{-i\Delta\beta\hat{A}^I}$ and $e^{-i\Delta\beta\hat{A}^J}$. Here, U_{JI} comprises simple gates to generate $|\Phi_J^{(0)}\rangle = \hat{U}_{JI}|\Phi_I^{(0)}\rangle$ initially. In practice, the unitary $e^{-i\Delta\beta\hat{A}(\mathbf{a}^I)}$ is Trotter-decomposed as

$$e^{-i\Delta\beta\hat{A}(\mathbf{a}^I)} \approx \prod_\mu e^{-i\theta_\mu^I\hat{\sigma}_\mu} \qquad (11)$$

with

$$\theta_\mu^I = \Delta\beta a_\mu^I. \qquad (12)$$

Since $\hat{A}^I$ and $\hat{A}^J$ only differ by the parameters $\mathbf{a}^I$ and $\mathbf{a}^J$ and share the same gate structure, it is convenient to order the controlled gates in an alternating manner as shown in Fig. 1a, noting that the controlled-$e^{-i\theta_\mu^I\hat{\sigma}_\mu}$ and controlled-$e^{-i\theta_\nu^I\hat{\sigma}_\nu}$ always commute. Without the control qubit, each Pauli rotation is performed by using the standard procedure[65–67] as shown in Fig. 1b, where (i) the qubits to be rotated are transformed to either of the X, Y, Z basis by the corresponding single-qubit unitary gates (denoted by R), (ii) their

parities are passed to the last qubit (denoted by the CNOT gate with a dotted line), and (iii) the R_z gate is applied followed by the Hermitian conjugate of (ii) and (i). Since the two adjacent controlled Pauli rotations carry out these unitary operations, the operations (i) and (ii) between them cancel out, and we can simplify the entire gate as depicted in Fig. 1c.

Therefore, the additional complexity in the quantum circuit of the state-specific MSQITE arises from the two CNOT operations and one additional R_z rotation. We consider this additional effort may not be a significant overhead cost compared with the circuit shown in Fig. 1b.

For the diagonal terms $H_{II}^{(\ell)}$, they represent energy expectation values of $|\Phi_I^{(\ell)}\rangle$, and therefore their quantum circuits are identical to that of QITE, without any controlled-$e^{-i\Delta\beta\hat{A}}$ operations. Hence, for MSQITE with a model space comprising n_{states} states, one needs to prepare $n_{\text{states}}(n_{\text{states}} - 1)/2$ circuits at each ℓ to measure $H_{IJ}^{(\ell)}$ (noting that $\mathbf{H}^{(\ell)}$ is Hermitian), in addition to n_{states} circuits to perform the same measurements as QITE.

To investigate the relative complexity of the quantum circuits required for the state-specific MSQITE, in Fig. 2, we have depicted the number of CNOT gates for each noiseless simulation conducted in this study, as a function of imaginary time β. It is evident that the number of CNOT gates for the off-diagonal circuit is generally approximately 1.2 to 1.3 times greater than that for the diagonal circuit. This observation underscores the efficacy and effectiveness of our quantum circuit designed for the Hadamard test. Nevertheless, it should be mentioned that these quantum circuits are too deep if noise is of concern, and in practice, we need to introduce appropriate simplifications, which will be discussed later.

MS-QLanczos

Similar to many other Krylov methods[46,68–70], QLanczos forms and diagonalizes the effective Hamiltonian to generate a wave function as a linear combination of time-evolved states. Here, we can generalize QLanczos to the model space formalism, which we call MS-QLanczos. Let us consider to expand the Krylov model subspace as

$$\left\{e^{-\ell\Delta\beta\hat{H}}|\Phi_I^{(0)}\rangle; (\ell = 0, \cdots, n); (I = 0, \cdots, n_{\text{states}} - 1)\right\} \qquad (13)$$

which comprises the basis for the effective Hamiltonian to be diagonalized. Here, we choose to use the normalized states $|\Phi_I^{(\ell)}\rangle$
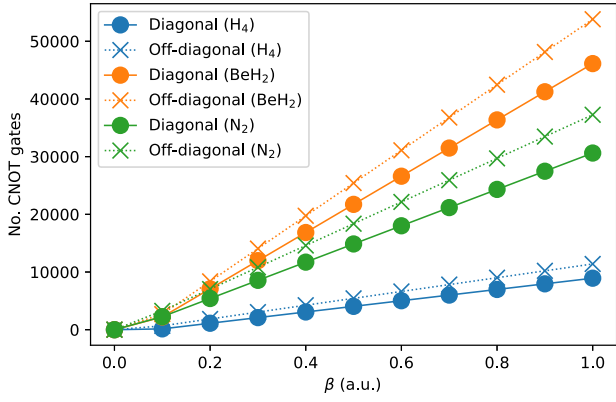
Fig. 2 Number of CNOT gates required for each of noiseless simulations. The number of CNOT gates required to estimate the energy expectation value (diagonal) $\langle\Phi_I|\hat{H}|\Phi_I\rangle$ and coupling (off-diagonal) $\langle\Phi_I|\hat{H}|\Phi_J\rangle$ are plotted as circles and crosses, respectively, as a function of imaginary time. The blue, orange, and green plots represent the results for simulations of H_4, BeH_2, and N_2, respectively.

to ease the derivation:

$$\left\{\left|\Phi_I^{(\ell)}\right\rangle; (\ell = 0, \cdots, n); (I = 0, \cdots, n_{\text{states}} - 1)\right\} \quad (14)$$

Note that it spans the same space as Eq. (13). At an arbitrary time step $\ell\Delta\beta$, the Ith quantum state is given by,

$$\begin{aligned}\left|\Phi_I^{(\ell)}\right\rangle &= \sum_J d_{JI}^{(\ell-1)} e^{-\Delta\beta(\hat{H} - E_J^{(\ell-1)})}\left|\Phi_J^{(\ell-1)}\right\rangle \\ &= \sum_J \tilde{d}_{JI}^{(\ell-1)} e^{-\Delta\beta(\hat{H} - E_0)}\left|\Phi_J^{(\ell-1)}\right\rangle\end{aligned} \quad (15)$$

where E_0 is some reference energy that is fixed throughout the imaginary time evolution (e.g., the average energy of the initial model space), and

$$\Delta E_I^{(\ell)} = E_I^{(\ell)} - E_0 \quad (16)$$

$$\tilde{d}_{JI}^{(\ell)} = d_{JI}^{(\ell)} e^{\Delta\beta\Delta E_J^{(\ell)}} \quad (17)$$

The global energy shift E_0 is introduced to ensure that the propagator is independent of both state and imaginary time, while avoiding the vanishing norm due to $e^{\Delta\beta E_0}$. Using the relation (15) recursively, we find

$$\left|\Phi_I^{(\ell)}\right\rangle = \sum_J \left(\tilde{\mathbf{d}}^{(\ell')} \cdots \tilde{\mathbf{d}}^{(\ell-1)}\right)_{JI} e^{-(\ell-\ell')\Delta\beta(\hat{H} - E_0)}\left|\Phi_J^{(\ell')}\right\rangle \quad (18)$$

for arbitrary $\ell' < \ell$.

Then, one can write the overlap matrix among the model space (14) as

$$\begin{aligned}\mathcal{S}_{I(\ell),J(\ell')} &\equiv \langle\Phi_I^{(\ell)}|\Phi_J^{(\ell')}\rangle \\ &= \left[\left(\mathbf{D}^{(\frac{\ell+\ell'}{2}\to\ell-1)}\right)^\top \mathbf{S}^{(\frac{\ell+\ell'}{2})}\left(\mathbf{D}^{(\ell'\to\frac{\ell+\ell'}{2}-1)}\right)^{-1}\right]_{IJ}\end{aligned} \quad (19)$$

where

$$\mathbf{D}^{(\ell'\to\ell)} \equiv \tilde{\mathbf{d}}^{(\ell')}\tilde{\mathbf{d}}^{(\ell'+1)} \cdots \tilde{\mathbf{d}}^{(\ell)} \quad (\ell > \ell') \quad (20)$$

Since we expect n_{states} to be small, the computational cost of $\mathbf{D}$ is negligible. The MS-QLanczos Hamiltonian matrix elements are similarly derived as

$$\begin{aligned}\mathcal{H}_{I(\ell),J(\ell')} &\equiv \langle\Phi_I^{(\ell)}|\hat{H}|\Phi_J^{(\ell')}\rangle \\ &= \left[\left(\mathbf{D}^{(\frac{\ell+\ell'}{2}\to\ell-1)}\right)^\top \mathbf{H}^{(\frac{\ell+\ell'}{2})}\left(\mathbf{D}^{(\ell'\to\frac{\ell+\ell'}{2}-1)}\right)^{-1}\right]_{IJ}.\end{aligned} \quad (21)$$

and one simply solves the generalized eigenvalue problem using $\mathcal{H}$ and $\mathcal{S}$. Note that the derivation reduces to that of the modified single-state QLanczos[39] when $n_{\text{states}} = 1$.

Illustrative noiseless simulations

Here, we assess the performance of MSQITE and MS-QLanczos, using molecular systems without the impact of noise. For this reason, we use the unitary coupled-cluster generalized singles and doubles (UCCGSD) ansatz[66,71], which was found to be suitable for $\hat{A}$ in simulating molecules[39]. Since the UCCGSD ansatz has the same number of parameters as the given Hamiltonian, the gate complexity for each time step in UCCGSD-based (MS)QITE scales similarly to the real time evolution. Therefore, the cost of UCCGSD-based MSQITE is expected to be comparable to algorithms that exploit real time evolution[68,70].

We first consider the BeH_2 molecule at equilibrium ($R_e = 1.334$ Å). As the initial model space for MSQITE, we choose the following three configurations: the HF configuration, and the configurations where two electrons are promoted from the highest occupied orbital to π_u orbitals, as listed in Fig. 3. In the same figure, the performances of various methods for the ground and excited states are depicted. Because the ground state of the system is only weakly correlated, QITE and especially QLanczos quickly converge. Also shown in the figure is the results of FSQITE (using the exact target energy $E_{2\Sigma_g^+} = 15.2263$ Hartree) and its extension to QLanczos (FS-QLanczos). Although FSQITE and FS-QLanczos eventually converge to the exact states, their evolutions are rather slow. Moreover, it cannot determine the ground state because it is far from the target state.

In contrast, clearly, MSQITE delivers remarkably fast convergence to the desired excited states when compared with FSQITE. Nevertheless, we note the convergence profile is state-dependent. For example, while the ground state $X\Sigma_g^+$ converged rapidly, the convergence of the second excited state $1\Delta_g$ required approximately $\beta = 5$ a.u., followed by the convergence of the first excited state $2\Sigma_g^+$ approximately $5\sim6$ a.u. later. This difference is attributed to the fact that these excited states are strongly correlated. To see this, we have tabulated the coefficients of the exact eigenstates in Supplementary Table 1. It is verified that the initial configurations (ii) $|000000110011\rangle$ and (iii) $|000011000011\rangle$ are the dominant ones for $1\Delta_g$, each with a coefficient of about 0.5, but it also contains other dominant configurations such as $|000000111100\rangle$ (see the Methods section for our qubit mapping: here, two electrons are promoted from $2\sigma_g$ to $1\pi_u$ with respect to HF). Such additional configurations need to be generated by the (MS)QITE procedure, and the imaginary time evolution typically takes more steps if their coefficients are non-negligible. From Supplementary Table 1, it is seen that the $2\Sigma_g^+$ state is even more strongly correlated than $1\Delta_g$, resulting in slower convergence in MSQITE.

We can expect a better performance of MSQITE if these additional configurations are included in the initial model space; however, of course, such detailed information may not be accessible a priori. Instead, MS-QLanczos can automatically detect and extract these states much earlier than MSQITE, as shown in Fig. 3. In contrast to FS-QLanczos, MS-QLanczos was not able to obtain the $3\Sigma_g^+$ state. This is simply because we have truncated the Krylov vectors to avoid numerical instabilities. If such higher states are desired, one needs to add more states in the model space, and MSQITE (MS-QLanczos) can find the eigenstates in the energy order.

It should be noted that the MSQITE method should bring a certain advantage not only for excited states but also for strongly correlated ground states, because the model space by definition can naturally provide multi-configuration states. To observe this advantage, we take the square H_4 molecule with a bond length of 1 Å as an example. As shown in Fig. 4, QITE and QLanczos take

(a)

(b)

(c)

(d)
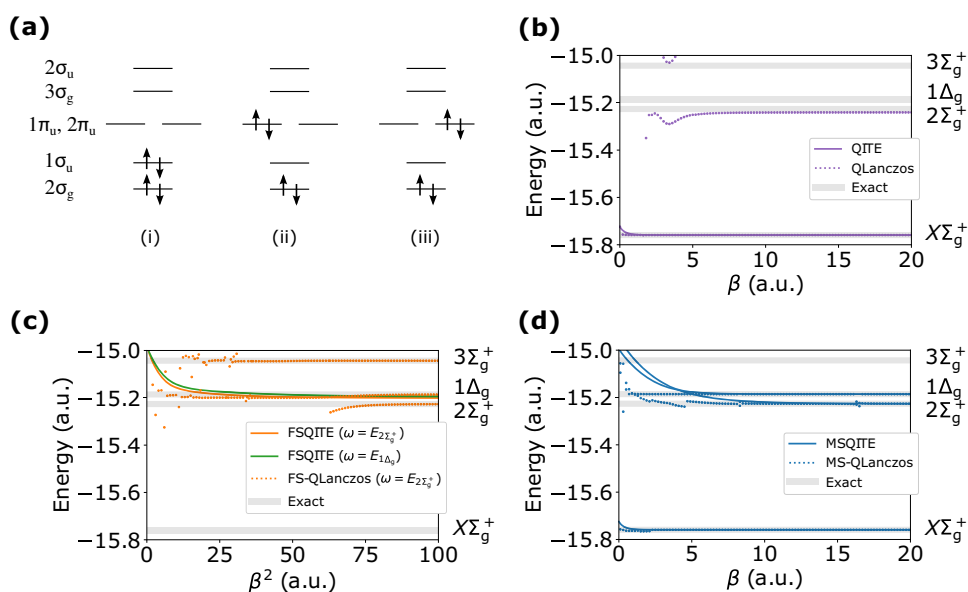
Fig. 3   Summary for BeH$_2$ simulation results. a Orbital diagrams. **b** Convergence behavior of energy using the QITE algorithm. **c** Convergence behavior of energy using FSQITE. **d** Convergence behavior of energy using MSQITE.

(a)

(b)
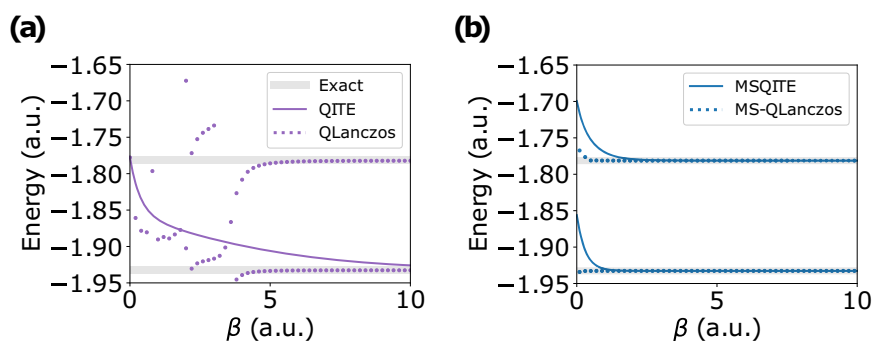
Fig. 4   Comparison between different algorithms for H$_4$. a QITE. **b** MSQITE.

more than 10 and 4 a.u. in imaginary time, respectively, to reach the ground state within the 1 mHartree accuracy. The slow convergence of the former is ascribed to the strong correlation in H$_4$, which is a two-determinant system with $|00001111\rangle$ and $|00110011\rangle$.

The energies of MSQITE with the model space comprised of $|00001111\rangle$ and $|00110011\rangle$ approach the (near) exact energies very rapidly, within less than a few a.u. in imaginary time; MS-QLanczos convergence is even faster. We emphasize, however, that each of the resulting MSQITE basis states $|\Phi_I\rangle$ are *not* the exact eigenstates. They are rather states that have either $|00001111\rangle$ or $|00110011\rangle$ as the dominant configuration, but possess almost no component of the other configuration. Nevertheless, the model space is developed to the complete space during the MSQITE procedure, such that linear combinations of $\{|\Phi_I\rangle\}$ are the exact states, as described in the preceding section.

Avoiding spin contamination with shifted propagator
For a non-relativistic molecular Hamiltonian, the exact wave function is an eigenstate of the number operator $\hat{N}$ and the spin operators $\hat{S}^2$ and $\hat{S}_z$. However, since each of the Pauli rotations applied in QITE does not necessarily commute with these symmetry operators, both the number of electrons and spin quantum numbers fluctuate during the evolution. Nevertheless,

for the one-particle symmetry operators ($\hat{N}$ and $\hat{S}_z$), such fluctuations are moderate and do not affect the result in our numerical experiments. It is also relatively easy to constrict the quantum state to the fixed quantum numbers by using fermionic operators instead of Pauli operators, i.e., one can employ the parametrization of Eq. (2) and treat linear combinations of Pauli operators[57].

However, we found that the $\hat{S}^2$ symmetry is difficult to preserve, especially for excited states. Usually, the initial model space is prepared such that only the target spin states (e.g., singlets in the above cases) are included. However, due to the approximate nature of (MS)QITE, the model space often starts to leak into different spin symmetry spaces and finds higher spin states (e.g., triplets $s = 1$ and quintets $s = 2$) with lower energies than states with the desired spin, by virtue of imaginary time propagation. For variational simulations, one can use the projection operator[55–57], but it is not straightforward to apply it in the framework of ITE.

Such spin-contamination, and the resulting "spin-collapse", are practical yet significant issues in MSQITE, as demonstrated below. In essence, when spin-collapse occurs, it becomes necessary to increase the number of states in a model space, n_{states}, and rerun the calculation to obtain the true target states with low spins. However, it is important to emphasize that states with higher spin (s) than the target spin (i.e., singlet) can be obtained much more efficiently through a separate calculation, in which the initial states

are prepared with $s = m_s = (N_\alpha - N_\beta)/2$, where N_α and N_β are the numbers of α and β electrons, respectively. Such simulations preclude the convergence of low-spin states (the fluctuation of $\hat{S}_z$ and hence m_s is negligible in MSQITE). In fact, this protocol is a de facto standard frequently employed in quantum chemical calculations to track different spin states.

As an example, here we consider N_2 at a stretched bond distance of 1.6 Å. We employed three configurations (i), (ii), and (iii) given in Fig. 5 to make an initial model space, and aimed to obtain the lowest singlet states with the Σ_g^+ symmetry including the ground state.

Figure 6a plots the changes in energy of MSQITE state along with the exact energies of different spin symmetries (shown in
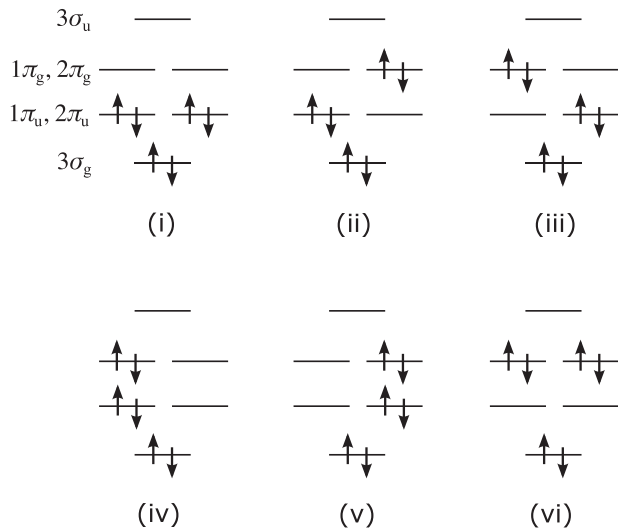
Fig. 5 Configurations used in MSQITE for N_2. In the MSQITE simulation of N_2, we have used up to six electronic configurations as the model space. Configuration (i) is the HF determinant. Two electrons are excited from π_u to π_g in configurations (ii), (iii), (iv), and (v), and four electrons are excited in configuration (vi). All configurations belong to Σ_g^+.

different colors). Whereas the lowest state of MSQITE converges to the singlet ground state quickly, the two excited states suffer from slow convergence. More importantly, both converge to the quintet $1^5\Sigma_g^+$ state. It is worth reiterating that this lowest quintet state can be obtained simply by preparing an initial state with $m_s = 2$ in QITE, rather than adopting the more intricate MSQITE with $m_s = 0$. In Fig. 6b, we monitor the change in $\langle \hat{S}^2 \rangle$, and the third state is trapped in some unphysical spin state with $\langle \hat{S}^2 \rangle \approx 2$. The problem here is that, in general, the exact eigenstates are unknown and therefore, one may get confused as if the MSQITE state achieved a stationary triplet state, as $\|\mathbf{b}_I\| \approx 0$ for $90 < \beta < 120$. However, this is an artifact of spin contamination. In fact, this contaminated state was found to be a half-and-half mixture between $3^1\Sigma_g^+$ and $1^5\Sigma_g^+$: the exact energies are $E_{3^1\Sigma_g^+} = -108.293193$ and $E_{1^5\Sigma_g^+} = -108.463729$, and the energy expectation value of the state is trapped at about $-108.378 \approx (E_{3^1\Sigma_g^+} + E_{1^5\Sigma_g^+})/2$.

Moreover, note that there exist several spin states (triplets and quintets) that have lower energies than singlet states as is clear from the figure. Hence, the convergence of MSQITE to these wrong spin states is highly likely. Of course, one could add more configurations in the model space to obtain higher singlet states; however, one can easily imagine that this approach is inefficient and is best avoided. We should also mention that, as the representability and flexibility of $\hat{A}$ increase, MSQITE would become even more prone to spin-contamination.

Hence, we introduce the spin-shift to the propagator,

$$e^{-\beta\hat{H}} \rightarrow e^{-\beta\left(\hat{H} + \lambda(\hat{S}^2 - s(s+1))\right)} \tag{22}$$

where λ is an arbitrary positive number and s is the designated spin quantum number. Because MSQITE is expected to transform the initial model space into the complete subspace, the use of the spin-shift should also be able to fix the spin at the same time. The trick here is that, while the target spin component with s in the model space remains unaffected by the shifted propagator, the spin contaminants with $s' > s$ rapidly vanish. Note that we can always assume $s' \geq s$ by appropriately constructing the initial model space (namely, we set $m_s = s$). Thereby, the model space will be projected to spin s. We would like to emphasize that this effect shares similarities with the widely-used penalty function
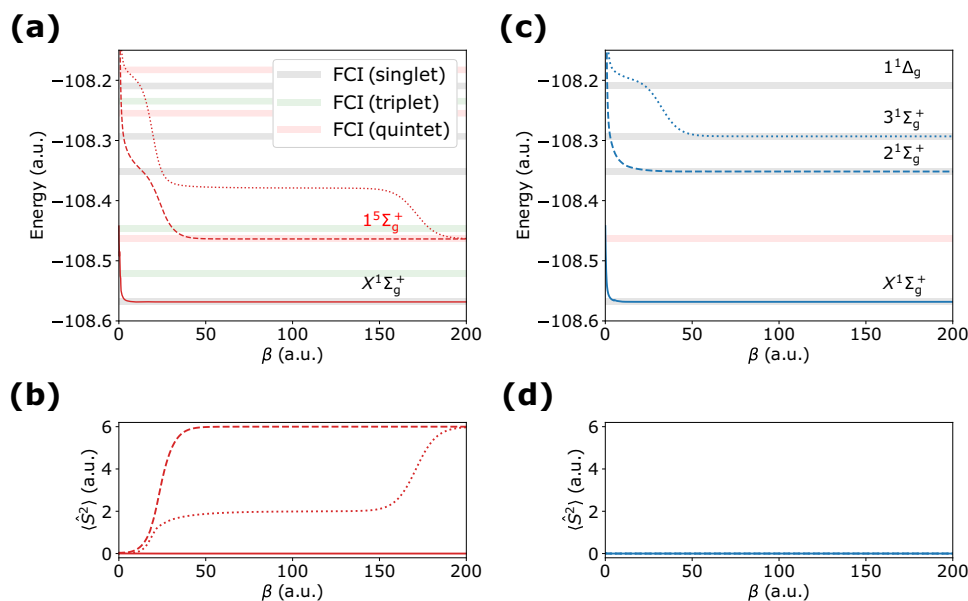
Fig. 6 Importance of spin-shift in MSQITE to remove spin-contamination (N_2 at a bond length of 1.6 Å). a Energy convergence profile of MSQITE. **b** Spin expectation value of MSQITE. **c** Energy convergence profile of MSQITE with spin-shift. **d** Spin expecation value of MSQITE with spin-shift.

method in variational algorithms, which penalizes the energy associated with incorrect spin states[7,55,72,73].

As λ becomes large, the spin-projection acts more strongly; however, it could spoil the convergence of MSQITE because of the large Trotter error. In principle, it suffices to use $\lambda > E_{s'} - E_s$ where $E_{s'}$ is the energy of excited state with spin s'.

In Fig. 6c, we show the results of MSQITE with the spin-shift using $\lambda = 0.5$. As expected, all the states nicely converge to the desired singlet states. Throughout all imaginary time, these states retain $\langle \hat{S}^2 \rangle = 0$ approximately, and get rid of spin contamination appropriately (Fig. 6d). We notice that the third state of MSQITE initially approaches the $1^1\Delta_g$ state instead of directly converging to the $3^1\Sigma_g^+$ state, and then starts to find the latter state as the lower state. However, this is not the weakness of the method; it is rather an indication of the ability of MSQITE to find the lowest states.

State-specific and state-averaged MSQITE

In the preceding section, we have discussed the advantages that MSQITE has to offer, focusing on the state-specific algorithm. As the state-averaged scheme is more attractive in terms of circuit complexity, we also carried out the state-averaged MSQITE to evaluate its accuracy. To properly evaluate the potential of the state-averaged MSQITE, we have performed noiseless simulations.

Table 1 compares the final energies obtained at convergence of the state-specific and state-averaged MSQITE methods. In addition to H_4 and BeH_2, N_2 at equilibrium (a bond distance of 1.098 Å) was tested with two configurations (i) and (ii) in Fig. 5, as an initial model space. Whereas the state-specific MSQITE yields quite accurate energies independent of systems, the state-averaged MSQITE results become significantly inaccurate for larger systems. Its accuracy is satisfactory for H_4 but deteriorates for N_2 with an

error of 47 mHartree for the $2\Sigma_g^+$ state. In general, increasing the model space tends to result in larger errors in energy, as shown in Fig. 7.

Another prominent example of the failure of the state-averaged MSQITE is the N_2 molecule with two π_u orbitals and two π_g orbitals and four electrons (comprising an eight qubit system). With a model space comprising six configurations—HF and all five pair-excited configurations derived from it (those listed in Fig. 5)—the UCCGSD-based state-averaged MSQITE methods immediately converge at $\beta = 0$, because $\sum_l \mathbf{b}^l = \mathbf{0}$ by symmetry. Note that this convergence does not indicate, of course, $\mathbf{b}^l = \mathbf{0}$ for each state; in fact, the state-specific MSQITE performs quite well, yielding very accurate energies. It is worth noting that $\sum_l \mathbf{b}^l$ is equivalent to the averaged energy derivative that appear in the VQE-based state-averaged UCCGSD method[74]; indeed, we applied the method to this system and found that it suffers from the same problem and no optimization of parameters was carried out. Overall, this strongly implies the limitation of other state-averaged methods for general systems[19,64].

It should be clear that this ill-behavior of the state-averaged MSQITE does *not* necessarily imply a possible theoretical flaw in our derivation. The failure is rather ascribed to the limitation of the form of $\hat{A}$ that we employed, i.e., single and double substitutions. In other words, it is unlikely the same UCCGSD amplitudes can evolve any arbitrary states to the desired ones all at once through $e^{-i\Delta\beta\hat{A}}$, even qualitatively. That being said, with triples (T) and quadruples (Q) included, we can rigorously obtain the exact eigenstates by definition: such UCCGSDTQ ansatz is complete for a four-electron system. For this particular case, the UCCGSDT-based state-averaged MSQITE already delivers almost the exact result (with less than 10^{-12} Hartree error). Overall, therefore, we are led to conclude that the state-averaged MSQITE does not seem practical because one needs way more Pauli operators from higher rank excitations than double excitations, to achieve a satisfactory accuracy, and this will become quickly infeasible with the increase in number of electrons.
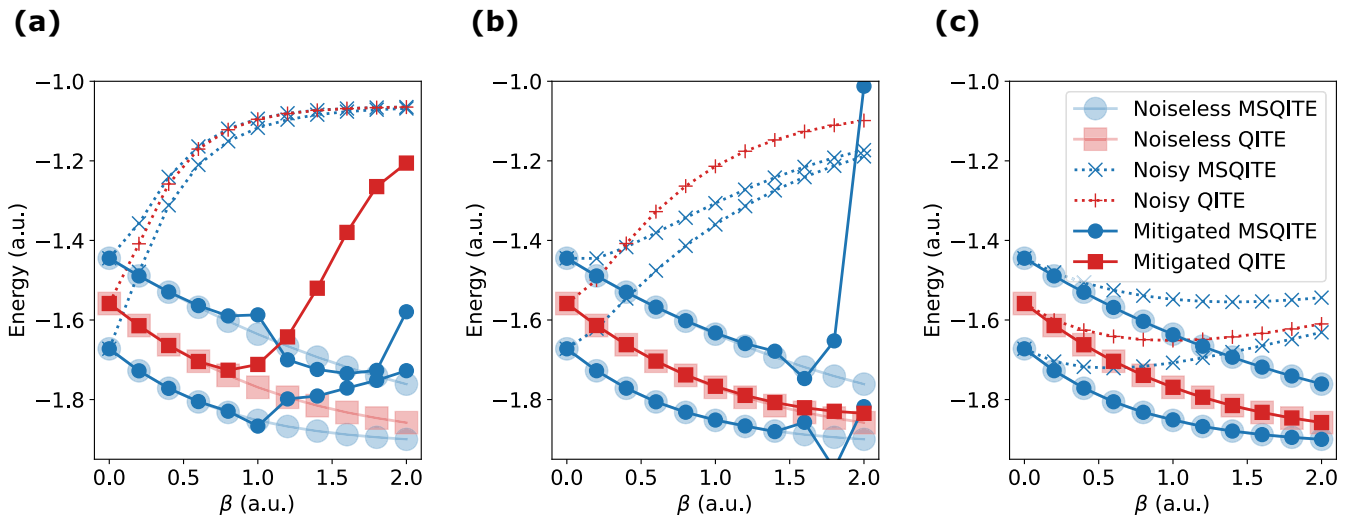
Noise simulations

To investigate the effect of quantum noise in state-specific MSQITE, here we conducted noise simulations on squared H_4 with each side being 2 Å. To simplify the quantum circuits, we have used several techniques (refer to the METHODS section for details). Figure 8 showcases both noiseless and noisy results for QITE and MSQITE, using different error rates for depolarizing noise. As is evident from the figure, both noisy QITE and noisy MSQITE experience significant error accumulation as imaginary time extends. For a relatively conservative error rate for each CNOT,

Table 1. Exact energy and error of MSQITE for each system (in Hartree).

System	Exact	State-specific	State-Averaged
H_4 XA_g	−1.932 645	$< 1 \times 10^{-8}$	2×10^{-7}
H_4 $2A_g$	−1.781 254	$< 1 \times 10^{-8}$	2×10^{-7}
BeH_2 $X\Sigma_g^+$	−15.759 026	$< 1 \times 10^{-8}$	2×10^{-4}
BeH_2 $2\Sigma_g^+$	−15.226 336	8×10^{-8}	4×10^{-4}
BeH_2 $1\Delta_g$	−15.185 771	$< 1 \times 10^{-8}$	2×10^{-4}
N_2 $X\Sigma_g^+$	−108.669 173	6×10^{-5}	5×10^{-3}
N_2 $2\Sigma_g^+$	−107.968 085	8×10^{-5}	5×10^{-2}
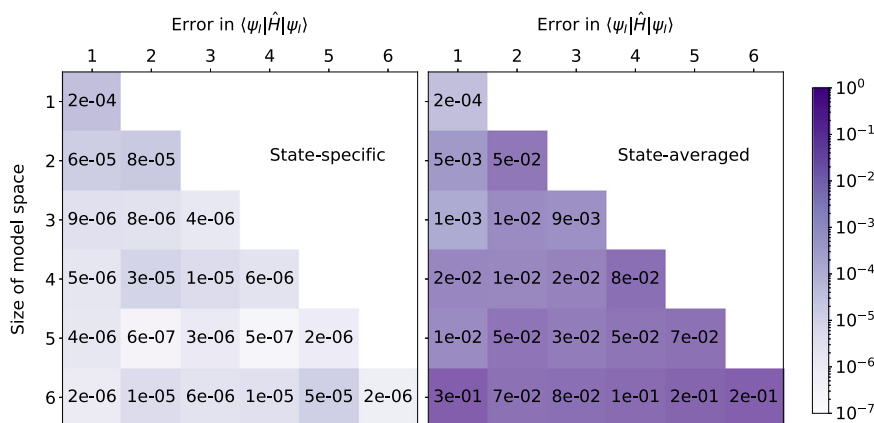
Fig. 7 Energy errors of the converged UCCGSD-based MSQITE states with the state-specific and state-averaged schemes for N_2 at equilibrium. States used for each MSQITE simulation are chosen from Fig. 5 in serial order.

(a) **(b)** **(c)**

Fig. 8 Noisy QITE and MSQITE for H₄ using depolarizing error. The error rate for one-qubit gates p_1 is set to 0.1 times the error rate for two-qubit gates p_2. (**a**) $p_2 = 10^{-2}$ (**b**) $p_2 = 5 \times 10^{-3}$, and (**c**) $p_2 = 10^{-3}$.
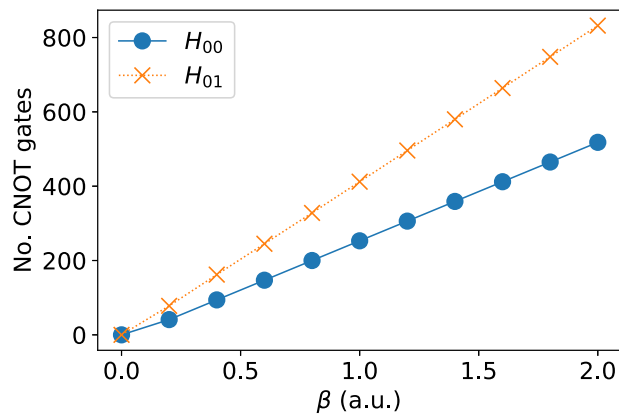
Fig. 9 Number of CNOT gates required for the noisy simulations of H₄. The numbers of CNOT gates required to estimate the energy expectation value (diagonal) $\langle \Phi_I | \hat{H} | \Phi_I \rangle$ and coupling (off-diagonal) $\langle \Phi_I | \hat{H} | \Phi_J \rangle$ in the noisy simulations of H₄ are plotted as circles and crosses, respectively, as a function of imaginary time.

$p_2 = 10^{-2}$, the energies for both noisy QITE and noisy MSQITE become exceed that of HF already at $\beta = 0.2$ a.u. (see Fig. 8a).

Such large errors arise from the fact that the circuit depth increases linearly with the imaginary-time step in these methods. Figure 9 plots the number of CNOT gates required to estimate the diagonal and off-diagonal matrix elements of the effective Hamiltonian, H_{00} and H_{01}, respectively, as a function of β. We will omit the results for H_{11} because they are almost the same as the results of H_{00}. As discussed earlier, the quantum circuit for H_{01} is not significantly more complicated compared to that for H_{00}. Nevertheless, the number of CNOT gates exceeds 400 at $\beta = 1$ a.u. (corresponding to five imaginary-time steps, as we used $\Delta\beta = 0.2$ a.u. for these simulations), which introduces tremendous errors even with a small error rate of $p_2 = 10^{-3}$ for each CNOT gate (Fig. 8c). Especially, as the noise accumulates, the expectation values of off-diagonal matrix elements tend to zero, leaving almost no advantage in performing MSQITE over QITE.

Therefore, it is essential to mitigate errors with NISQ computers[75,76], and we have employed the simple protocol of zero-noise-extrapolation (ZNE)[77–79]. The error-mitigated QITE and MSQITE results are also presented in Fig. 8 as filled squares and circles, respectively. Clearly, ZNE significantly improves the noisy results,

successfully approximating the noiseless energies at short β values. Interestingly, we observe that error-mitigated MSQITE is almost always as stable as mitigated QITE, despite the slightly longer circuit required for H_{01}.

This result can be rationalized by recognizing that the off-diagonal elements generally have a much smaller absolute value than the diagonal ones, serving as a correction to the QITE energies. As a result, errors in the off-diagonal elements have a less impact on the accuracy of the final energies of MSQITE, compared to errors in the diagonal elements.

In order to examine this result further, we plotted the noisy values ($p_2 = 5 \times 10^{-3}$) of H_{00} and H_{01} along side the extrapolated values in Fig. 10a. The "ideal" noiseless estimates, generated using the same quantum circuits and parameters **a**, are also shown. We observe that the noisy H_{01} converges to 0, as expected. In most instances, ZNE successfully approximates the noiseless values of H_{00} and H_{01} with notable accuracy. Figure 10b illustrates the differences between the ZNE and noiseless values, namely, $\Delta H_{IJ} = |H_{IJ}(\text{ZNE}) - H_{IJ}(\text{Noiseless})|$ for $IJ \in (00, 01)$, which serve as a metric for the reliability of ZNE. Remarkably, both ΔH_{00} and ΔH_{01} exhibit similar precision, even though H_{01} is presumed to be more susceptible to noise due to its higher CNOT gate count. Therefore, the errors introduced in the effective Hamiltonian are relatively balanced between diagonal and off-diagonal elements, yielding an overall noise impact on the MSQITE energies similar to that on the QITE energies.

In conclusion, MSQITE continues to offer its advantages over QITE even on NISQ computers, as long as suitable error-mitigation techniques are utilized to minimize the effects of noise.

DISCUSSION

In this work, we introduced a model space into QITE to enable excited state simulations. The orthogonality condition was retained using the Löwdin symmetric orthonormalization, which minimizes the state change during time steps and thus is suitable for the short-time unitary approximation of the non-unitary imaginary time propagation. MSQITE was shown to be a promising route to obtaining both ground and excited states, and its extension to QLanczos allowed for further acceleration in obtaining approximate eigenstates.

This study also proposed the spin-shift in the propagator. Because excited states frequently suffer from spin-contamination, it is necessary to remove the irrelevant spin configurations from

(a)

(b)
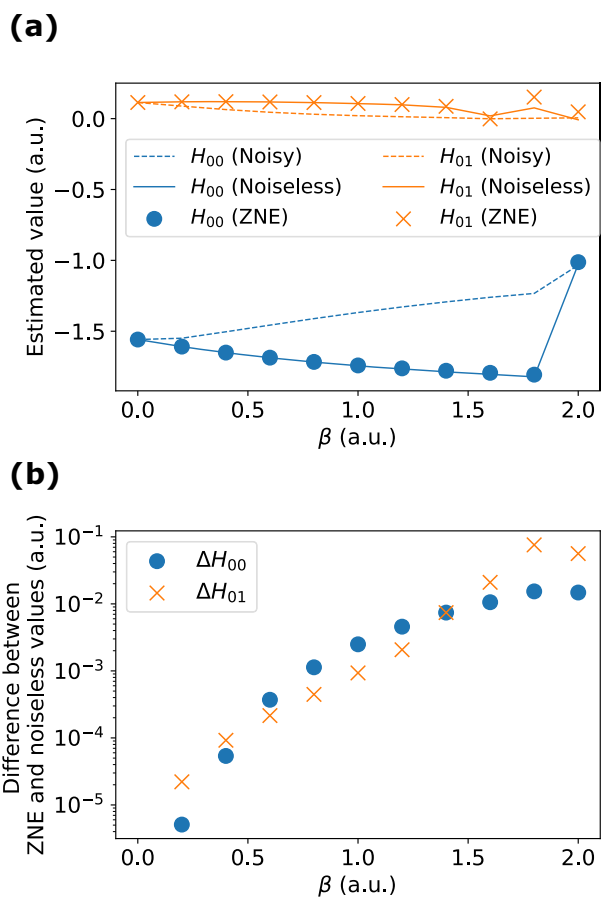
Fig. 10 Accuracy of effective Hamiltonian matrix elements in the presence of noise. a Noisy, noiseless, and extrapolated values of the diagonal and off-diagonal matrix elements H_{00} and H_{01} for $p_2 = 5 \times 10^{-3}$. **b** Difference between extrapolated and noiseless values.

the MSQITE simulation. We have shown that the proposed spin-shift approach achieves this feat by projecting out the desired spin symmetry through ITE.

Based on the results obtained in this work, we conclude that the state-averaged MSQITE is likely to necessitate substantially complicated $\hat{A}$ to appropriately evolve all the states considered in a model space, compared to the state-specific scheme. Namely, the former method requires fermionic excitations beyond double excitations in $\hat{A}$ for $e^{-i\Delta\beta\hat{A}}$ to achieve reasonable accuracy in ideal, noiseless computations; this is deemed unappealing because of the increasing number of Pauli operators that need to be included. It should be pointed out that the recently proposed state-averaged orbital-optimized VQE[74] shares the same difficulty because it uses the same unitary for multiple orthonormal states as in the state-averaged MSQITE; thus, the scalability of the method with the increase in number of electrons and states remains to be an open question. In contrast, the state-specific MSQITE is potentially more promising than the state-averaged one, requiring only single and double excitations (i.e., UCCGSD) to achieve high accuracy.

We have also demonstrated that the effectiveness of state-specific MSQITE compared to QITE remains unaffected by quantum noise when appropriate error mitigation techniques are applied. This resilience can be attributed to two main factors. First, the quantum circuit for state-specific MSQITE was designed to minimize the CNOT gate counts in the estimation of off-

diagonal elements. Second, both the diagonal and off-diagonal elements display a comparative magnitude of errors. Overall, we found that the impact of quantum noise on MSQITE is similar to that on QITE.

Finally, many of the ideas developed in this work are versatile, and we expect that they can be applied to fields beyond quantum chemistry, such as nuclear physics. Moreover, MSQITE can be extended to combine with adaptive algorithms[32,34] and variational algorithms[23,35]. This integration holds potential for new synergies that can reduce the circuit depth. We are currently working along these directions.

METHODS
Orthonormalization of model space

In the Results section, the transformation matrix **d** was introduced in MSQITE to preserve the orthonormality of the model space after a short-time propagation. We chose the Löwdin symmetric orthonormalization for this purpose. First, we form the overlap matrix of the target imaginary time evolved model space,

$$
\begin{aligned}
\tilde{S}_{IJ}^{(\ell)} &= \langle \Phi_I^{(\ell)} | e^{-\Delta\beta(\hat{H}-E_I)} e^{-\Delta\beta(\hat{H}-E_J)} | \Phi_J^{(\ell)} \rangle \\
&= S_{IJ}^{(\ell)} - 2\Delta\beta \left(H_{IJ}^{(\ell)} - \tfrac{1}{2}(E_I + E_J) S_{IJ}^{(\ell)} \right) + O(\Delta\beta^2)
\end{aligned}
\tag{23}
$$

which is truncated after the first-order of $\Delta\beta$ to obtain the approximate overlap. Note that here, we assume the model space is orthonormal $S_{IJ}^{(\ell)} = \delta_{IJ}$; however, even if this assumption is not satisfied, one can still find such a basis and the argument does not lose generality, see Supplementary Notes. Diagonalizing $\tilde{\mathbf{S}}$ gives

$$
\tilde{\mathbf{S}}\mathbf{U} = \mathbf{U}\tilde{\mathbf{s}}
\tag{24}
$$

where **s** is the diagonal matrix with the eigenvalues and **U** the eigenvectors. The **d** matrix from the Löwdin symmetric ortho-normalization is then uniquely obtained as

$$
\mathbf{d} = \mathbf{U}\tilde{\mathbf{s}}^{-1/2}\mathbf{U}^{\dagger}.
\tag{25}
$$

We note that, instead of the above **d**, it would be also tempting to employ the transformation that diagonalizes $\langle \Phi_I^{(\ell)} | e^{-\Delta\beta(\hat{H}-E_I)} \hat{H} e^{-\Delta\beta(\hat{H}-E_J)} | \Phi_J^{(\ell)} \rangle$, such that $\lim_{\ell\to\infty} |\Phi_I^{(\ell)}\rangle$ is the exact ground or excited state, $|\psi_I\rangle$. However, it is easily seen that the unitary matrix obtained from the diagonalization of such an effective Hamiltonian matrix is inadequate because it can flip the signs and even the ordering of the states, and thus Eq. (1) cannot be a valid approximation.

Following Blunt et al.[50], one may perform the Gram-Schmidt orthogonalization to define **d**. However, the Gram-Schmidt orthogonalization is not unique about the order of orthogonaliza-tion steps and also leads to a biased update of $\{|\Phi_I\rangle\}$. Importantly, the propagation of the first state $|\Phi_0\rangle$ will remain unaffected by the presence of other states $|\Phi_I\rangle (I > 0)$. Therefore, it will naturally become the exact ground state at $\beta \to \infty$. It is highly desirable that $|\Phi_0\rangle$ is initially chosen to be the closest to the ground state among all the states in the model space at $\beta = 0$. Otherwise, the model space would experience large reorganization, which the short-time unitary evolution of Eq. (1) would find difficult to express. This requirement may be easily satisfied for the ground state (i.e., HF may be the most reasonable starting point). However, for excited states, the appropriate ordering is generally unknown.

Simulation details

MSQITE and MS-QLanczos were implemented in our Python-based emulator package, QUKET[80], which compiles other useful libraries such as OPENFERMION[81], PYSCF[82], and QULACS[83], to perform quantum simulations. In all simulations, we used the STO-6G basis set and HF orbitals. The Jordan-Wigner transformation was employed to

map the fermion operators to the qubit representation, such that α and β spin orbitals were aligned alternately with the rightmost qubit represents the lowest energy α spin orbital. $\Delta\beta$ was set to 0.1 a.u. for QITE and MSQITE, and 0.05 a.u. for FSQITE, and the UCCGSD ansatz[39,66,71] was employed for the form of $\hat{A}$, except for the noise simulations, for which the computational details are provided below. To ensure stable solutions to Eqs. (7) and (10), we applied a truncation scheme to the singular values of the $\mathbf{M}$ matrix. Specifically, we retained singular values that were greater than 10^{-7} times the largest singular value, as implemented in scipy.linalg.lstsq[84]. The Be 1s orbital and the N 1s and 2s orbitals were not considered in the simulations. For H_4, the initial HF calculation was performed with the C_{2h} symmetry instead of D_{4h}, to relax the orbitals.

To perform noisy QITE simulations, we used Qiskit[85], and the quantum noise was modeled by depolarizing error. The error rate for two-qubit gates was set to $p_2 = 10^{-2}, 5 \times 10^{-2}$, and 10^{-3}, and that for one-qubit gates was set to $p_1 = p_2 \times 0.1$. To mitigate the error, we used the zero-noise-extrapolation technique with exponential fitting[79], only for two-qubit (CNOT) gates. For this simulation, the Hamiltonian of H_4 is transformed to the reduced representation by using the tapering-off technique[58,59]. This resulted in the following four-qubit-Hamiltonian:

$$
\begin{aligned}
\hat{H} = {} & h_0 + h_1 Z_2 Z_3 + h_2 Z_0 Z_2 + h_2 Z_1 Z_2 + h_3 Z_0 Z_1 Z_2 Z_3 \\
& + h_4 Z_0 Z_1 + h_5 (Z_0 Z_3 + Z_1 Z_3) + h_6 (Z_0 Z_1 Z_2 + Z_0 Z_1 Z_3) \\
& + h_7 (-X_1 Y_2 Y_3 + Y_1 Y_2 X_3 - X_0 Y_2 Y_3 + Y_0 Y_2 X_3) \\
& + h_8 (-Y_0 Z_1 Y_2 Z_3 + Y_0 Z_1 Z_2 Y_3 + Y_0 Y_2 + X_0 Z_1 Z_2 X_3 \\
& \quad - Z_0 X_1 Z_2 X_3 + Z_0 Y_1 Y_2 Z_3 - Z_0 Y_1 Z_2 Y_3 - Y_1 Y_2) \\
& + h_9 (-X_0 Z_1 Z_2 + X_0 Z_2 Z_3 - Z_0 X_1 Z_2 + X_1 Z_2 Z_3) \\
& + h_{10} (X_0 X_1 Z_2 Z_3 + Y_0 Y_1 Z_2 Z_3) \\
& + h_{11} (X_0 Y_1 Y_2 - Y_0 X_1 Y_2 + Z_0 Z_2 X_3 - Z_1 Z_2 X_3) \\
& + h_{12} (Z_0 Z_1 Y_2 Y_3 - Y_2 Y_3) \\
& + h_{13} (Z_0 + Z_1 + Z_0 Z_2 Z_3 + Z_1 Z_2 Z_3) \\
& + h_{14} (-Z_2 - Z_3) \\
& + h_{15} (-Y_0 X_1 Y_3 + X_0 Y_1 Y_3 + Z_0 X_2 Z_3 - Z_1 X_2 Z_3)
\end{aligned}
\tag{26}
$$

with

$$
\begin{aligned}
h_0 &= -1.0613356242517709 \\
h_1 &= 0.3752318182963852 \\
h_2 &= 0.3736748877137335 \\
h_3 &= 0.3722054115496151 \\
h_4 &= 0.26033869205932514 \\
h_5 &= 0.22299864958557725 \\
h_6 &= 0.09825586237928423 \\
h_7 &= 0.07901175885991207 \\
h_8 &= 0.0746839486241239 \\
h_9 &= 0.07166447926823467 \\
h_{10} &= 0.05690174793752179 \\
h_{11} &= 0.05592385851947026 \\
h_{12} &= 0.05496497155276822 \\
h_{13} &= 0.03491578410706995 \\
h_{14} &= 0.021711508654056723 \\
h_{15} &= 0.018760090104653643
\end{aligned}
$$

whose first two lowest eigenvalues are -1.91552763 and -1.87493645. We also simplified the quantum circuit by using a reduced form of UCCGD (neglecting single substitutes) where only one of eight Pauli tensors arising from a double excitation is used.

The resulting form of the ansatz is comprised of 14 unitaries,

$$
e^{-i\Delta\beta\hat{A}} = \prod_{i=0}^{13} e^{-i\Delta\beta a_i \hat{P}_i}
\tag{27}
$$

with

$$
\begin{aligned}
\hat{P}_0 &= X_1 X_2 Y_3 \\
\hat{P}_1 &= X_0 X_2 Y_3 \\
\hat{P}_2 &= Z_1 Z_2 Y_3 \\
\hat{P}_3 &= X_0 Y_1 X_2 \\
\hat{P}_4 &= X_2 Y_3 \\
\hat{P}_5 &= X_0 Y_3 \\
\hat{P}_6 &= X_1 Y_2 \\
\hat{P}_7 &= Y_0 Z_1 X_2 Z_3 \\
\hat{P}_8 &= Z_0 Y_1 Z_2 X_3 \\
\hat{P}_9 &= Y_0 X_1 X_3 \\
\hat{P}_{10} &= Z_0 Y_2 Z_3 \\
\hat{P}_{11} &= X_0 Y_1 \\
\hat{P}_{12} &= Y_1 \\
\hat{P}_{13} &= Y_0
\end{aligned}
$$

where we have used $\Delta\beta = 0.2$ a.u.. We used the stabilization procedure for QLanczos as presented in ref. [39] to describe excited states. However, for MS-QLanczos, the numerical instability arising from the linear dependence in the Krylov subspace becomes even more challenging compared with the standard QLanczos. Hence, we adopted the same procedure as ref. [31], i.e., we use Krylov vectors that satisfy $\mathcal{S}_{\ell\ell'} < 0.99$ to alleviate the linear dependence. However, we have made the following modifications: the selection of Krylov vectors is performed backwards (i.e., starting from the current time instead of from the initial time) in order to ensure the latest states are always included in the basis, and the number of states included in the subspace is limited to 5. The selection is based on the assumption that excessively old time states do not play an important role but only cause numerical instabilities.

DATA AVAILABILITY
The data that support the findings of this study are available from the corresponding author upon reasonable request.

CODE AVAILABILITY
The code that is used to produce the data presented in this study is available from the authors upon reasonable request.

REFERENCES
1. Moll, N. et al. Quantum optimization using variational algorithms on near-term quantum devices. *Quantum Sci. Technol.* **3**, 030503 (2018).
2. Cerezo, M. et al. Variational quantum algorithms. *Nat. Rev. Phys.* **3**, 625–644 (2021).
3. Farhi, E., Goldstone, J. & Gutmann, S. A quantum approximate optimization algorithm. Preprint at https://arxiv.org/abs/1411.4028 (2014).
4. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
5. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. *Nat. Comm.* **5**, 4213 (2014).
6. Wang, D., Higgott, O. & Brierley, S. Accelerated variational quantum eigensolver. *Phys. Rev. Lett.* **122**, 140504 (2019).

7. McClean, J. R., Romero, J., Babbush, R. & Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. *New J. Phys.* **18**, 023023 (2016).

8. Wecker, D. et al. Solving strongly correlated electron models on a quantum computer. *Phys. Rev. A* **92**, 062318 (2015).

9. O'Malley, P. J. J. et al. Scalable quantum simulation of molecular energies. *Phys. Rev. X* **6**, 031007 (2016).

10. Kandala, A. et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* **549**, 242–246 (2017).

11. Cao, Y. et al. Quantum chemistry in the age of quantum computing. *Chem. Rev.* **119**, 10856–10915 (2019).

12. McArdle, S., Endo, S., Aspuru-Guzik, A., Benjamin, S. C. & Yuan, X. Quantum computational chemistry. *Rev. Mod. Phys.* **92**, 015003 (2020).

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. Comm.* **10**, 3007 (2019).

14. Cerasoli, F. T., Sherbert, K., Sławińska, J. & Buongiorno Nardelli, M. Quantum computation of silicon electronic band structure. *Phys. Chem. Chem. Phys.* **22**, 21816–21822 (2020).

15. Fan, Y., Liu, J., Li, Z. & Yang, J. Equation-of-motion theory to calculate accurate band structures with a quantum computer. *J. Phys. Chem. Lett.* **12**, 8833–8840 (2021).

16. Yoshioka, N., Sato, T., Nakagawa, Y. O., Ohnishi, Y.-y & Mizukami, W. Variational quantum simulation for periodic materials. *Phys. Rev. Res.* **4**, 013052 (2022).

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

18. Higgott, O., Wang, D. & Brierley, S. Variational quantum computation of excited states. *Quantum* **3**, 156 (2019).

19. Parrish, R. M., Hohenstein, E. G., McMahon, P. L. & Martínez, T. J. Quantum computation of electronic transitions using a variational quantum eigensolver. *Phys. Rev. Lett.* **122**, 230401 (2019).

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

21. Zhang, F., Gomes, N., Yao, Y., Orth, P. P. & Iadecola, T. Adaptive variational quantum eigensolvers for highly excited states. *Phys. Rev. B* **104**, 075159 (2021).

22. Xie, Q.-X., Liu, S. & Zhao, Y. Orthogonal state reduction variational eigensolver for the excited-state calculations on quantum computers. *J. Chem. Theory Comput.* **18**, 3737–3746 (2022).

23. McArdle, S. et al. Variational ansatz-based quantum simulation of imaginary time evolution. *npj Quantum Inf.* **5**, 75 (2019).

24. Benedetti, M., Fiorentini, M. & Lubasch, M. Hardware-efficient variational quantum algorithms for time evolution. *Phys. Rev. Res.* **3**, 033083 (2021).

25. McArdle, S. & Tew, D. P. Improving the accuracy of quantum computational chemistry using the transcorrelated method. Preprint at https://arxiv.org/abs/2006.11181 (2020).

26. Kumar, A. et al. Quantum simulation of molecular electronic states with a transcorrelated Hamiltonian: higher accuracy with fewer qubits. *J. Chem. Theory Comput.* **18**, 5312–5324 (2022).

27. Sokolov, I. O., Dobrautz, W., Luo, H., Alavi, A. & Tavernelli, I. Orders of magnitude increased accuracy for quantum many-body problems on quantum computers via an exact transcorrelated method. *Phys. Rev. Res.* **5**, 023174 (2023).

28. Lin, S.-H., Dilip, R., Green, A. G., Smith, A. & Pollmann, F. Real- and imaginary-time evolution with compressed quantum circuits. *PRX Quantum* **2**, 010342 (2021).

29. Liu, T., Liu, J.-G. & Fan, H. Probabilistic nonunitary gate in imaginary time evolution. *Quantum Inf. Process.* **20**, 204 (2021).

30. Kosugi, T., Nishiya, Y., Nishi, H. & Matsushita, Y.-i Imaginary-time evolution using forward and backward real-time evolution with a single ancilla: First-quantized eigensolver algorithm for quantum chemistry. *Phys. Rev. Res.* **4**, 033121 (2022).

31. Motta, M. et al. Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution. *Nat. Phys.* **16**, 205–210 (2020).

32. Yeter-Aydeniz, K., Pooser, R. C. & Siopsis, G. Practical quantum computation of chemical and nuclear energy levels using quantum imaginary time evolution and Lanczos algorithms. *npj Quantum Inf.* **6**, 63 (2020).

33. Yeter-Aydeniz, K., Siopsis, G. & Pooser, R. C. Scattering in the Ising model with the quantum Lanczos algorithm. *New J. Phys.* **23**, 043033 (2021).

34. Gomes, N. et al. Efficient step-merged quantum imaginary time evolution algorithm for quantum chemistry. *J. Chem. Theory Comput.* **16**, 6256–6266 (2020).

35. Gomes, N. et al. Adaptive variational quantum imaginary time evolution approach for ground state preparation. *Adv. Quantum Technol.* **4**, 2100114 (2021).

36. Sun, S.-N. et al. Quantum computation of finite-temperature static and dynamical properties of spin systems using quantum imaginary time evolution. *Phys. Rev. X Quantum* **2**, 010317 (2021).

37. Huang, Y., Shao, Y., Ren, W., Sun, J. & Lv, D. Efficient quantum imaginary time evolution by drifting real-time evolution: an approach with low gate and measurement complexity. *J. Chem. Theory Comput.* **19**, 3868–3876 (2023).

38. Amaro, D. et al. Filtering variational quantum algorithms for combinatorial optimization. *Quantum Sci. Technol.* **7**, 015021 (2022).

39. Tsuchimochi, T., Ryo, Y., Ten-no, S. L. & Sasasako, K. Improved algorithms of quantum imaginary time evolution for ground and excited states of molecular systems. *J. Chem. Theory Comput.* **19**, 503–513 (2023).

40. Jouzdani, P., Johnson, C. W., Mucciolo, E. R. & Stetcu, I. Alternative approach to quantum imaginary time evolution. *Phys. Rev. A* **106**, 062435 (2022).

41. Santagati, R. et al. Witnessing eigenstates for quantum simulation of Hamiltonian spectra. *Sci. Adv.* **4**, eaap9646 (2018).

42. Colless, J. I. et al. Computation of molecular spectra on a quantum processor with an error-resilient algorithm. *Phys. Rev. X* **8**, 011021 (2018).

43. Ollitrault, P. J. et al. Quantum equation of motion for computing molecular excitation energies on a noisy quantum processor. *Phys. Rev. Res.* **2**, 043140 (2020).

44. Tkachenko, N. V. et al. Quantum Davidson algorithm for excited states. Preprint at https://arxiv.org/abs/2204.10741 (2022).

45. Heya, K. et al. Subspace variational quantum simulator. *Phys. Rev. Res.* **5**, 023078 (2023).

46. Takahiro, Y., Ten-no, S. L. & Tsuchimochi, T. Quantum inverse algorithm via adaptive variational quantum linear solver: applications to general eigenstates. *J. Phys. Chem. A* **127**, 6577–6592 (2023).

47. Ohtsuka, Y. & Nagase, S. Projector Monte Carlo method based on Slater determinants: test application to singlet excited states of H_2O and liF. *Chem. Phys. Lett.* **485**, 367–370 (2010).

48. Booth, G. H. & Chan, G. K.-L. Communication: excited states, dynamic correlation functions and spectral properties from full configuration interaction quantum Monte Carlo. *J. Chem. Phys.* **137**, 191102 (2012).

49. Ten-no, S. Stochastic determination of effective Hamiltonian for the full configuration interaction solution of quasi-degenerate electronic states. *J. Chem. Phys.* **138**, 164126 (2013).

50. Blunt, N. S., Smart, S. D., Booth, G. H. & Alavi, A. An excited-state approach within full configuration interaction quantum Monte Carlo. *J. Chem. Phys.* **143**, 134117 (2015).

51. Ten-no, S. L. Multi-state effective Hamiltonian and size-consistency corrections in stochastic configuration interactions. *J. Chem. Phys.* **147**, 244107 (2017).

52. Liu, J.-G., Zhang, Y.-H., Wan, Y. & Wang, L. Variational quantum eigensolver with fewer qubits. *Phys. Rev. Res.* **1**, 023025 (2019).

53. Gard, B. T. et al. Efficient symmetry-preserving state preparation circuits for the variational quantum eigensolver algorithm. *npj Quantum Inf.* **6**, 10 (2020).

54. Seki, K., Shirakawa, T. & Yunoki, S. Symmetry-adapted variational quantum eigensolver. *Phys. Rev. A* **101**, 052340 (2020).

55. Yen, T.-C., Lang, R. A. & Izmaylov, A. F. Exact and approximate symmetry projectors for the electronic structure problem on a quantum computer. *J. Chem. Phys.* **151**, 164111 (2019).

56. Tsuchimochi, T., Mori, Y. & Ten-no, S. L. Spin-projection for quantum computation: a low-depth approach to strong correlation. *Phys. Rev. Res.* **2**, 043142 (2020).

57. Tsuchimochi, T., Taii, M., Nishimaki, T. & Ten-no, S. L. Adaptive construction of shallower quantum circuits with quantum spin projection for fermionic systems. *Phys. Rev. Res.* **4**, 033100 (2022).

58. Bravyi, S., Gambetta, J. M., Mezzacapo, A. & Temme, K. Tapering off qubits to simulate fermionic Hamiltonians. Preprint at https://arxiv.org/abs/1701.08213 (2017).

59. Setia, K. et al. Reducing qubit requirements for quantum simulations using molecular point group symmetries. *J. Chem. Theory Comput.* **16**, 6091–6097 (2020).

60. Löwdin, P. On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *J. Chem. Phys.* **18**, 365–375 (1950).

61. Szabo, A. & Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover Publications, Mineola, NY, 1996).

62. Carlson, B. C. & Keller, J. M. Orthogonalization procedures and the localization of wannier functions. *Phys. Rev.* **105**, 102–103 (1957).

63. Mayer, I. On löwdin's method of symmetric orthogonalization. *Int. J. Quantum Chem.* **90**, 63–65 (2002).

64. Ibe, Y. et al. Calculating transition amplitudes by variational quantum deflation. *Phys. Rev. Res.* **4**, 013173 (2022).

65. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).

66. Barkoutsos, P. K. et al. Quantum algorithms for electronic structure calculations: particle-hole Hamiltonian and optimized wave-function expansions. *Phys. Rev. A* **98**, 022322 (2018).

67. Romero, J. et al. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Sci. Technol.* **4**, 014008 (2019).

68. Stair, N. H., Huang, R. & Evangelista, F. A. A multireference quantum Krylov algorithm for strongly correlated electrons. *J. Chem. Theory Comput.* **16**, 2236–2245 (2020).

69. Seki, K. & Yunoki, S. Quantum power method by a superposition of time-evolved states. *PRX Quantum* **2**, 010333 (2021).

70. Klymko, K. et al. Real-time evolution for ultracompact Hamiltonian eigenstates on quantum hardware. *PRX Quantum* **3**, 020323 (2022).

71. Lee, J., Huggins, W. J., Head-Gordon, M. & Whaley, K. B. Generalized unitary coupled cluster wave functions for quantum computation. *J. Chem. Theory Comput.* **15**, 311–324 (2019).

72. Andrews, J. S., Jayatilaka, D., Bone, R. G., Handy, N. C. & Amos, R. D. Spin contamination in single-determinant wavefunctions. *Chem. Phys. Lett.* **183**, 423 – 431 (1991).

73. Ryabinkin, I. G., Genin, S. N. & Izmaylov, A. F. Constrained variational quantum eigensolver: Quantum computer search engine in the Fock space. *J. Chem. Theory Comput.* **15**, 249–255 (2019).

74. Yalouz, S. et al. A state-averaged orbital-optimized hybrid quantum–classical algorithm for a democratic description of ground and excited states. *Quantum Sci. Technol.* **6**, 024004 (2021).

75. Kandala, A. et al. Error mitigation extends the computational reach of a noisy quantum processor. *Nature* **567**, 491–495 (2019).

76. Kim, Y. et al. Scalable error mitigation for noisy quantum circuits produces competitive expectation values. *Nat. Phys.* **19**, 752–759 (2023).

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

78. Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.* **119**, 180509 (2017).

79. Endo, S., Benjamin, S. C. & Li, Y. Practical quantum error mitigation for near-future applications. *Phys. Rev. X* **8**, 031027 (2018).

80. Tsuchimochi, T. et al. Quket: the comprehensive quantum simulator for quantum chemistry https://github.com/quket/quket (2022). Date of access: 2023/9/29.

81. McClean, J. R. et al. OpenFermion: the electronic structure package for quantum computers. *Quantum Sci. Technol.* **5**, 034014 (2020).

82. Sun, Q. et al. Pyscf: the python-based simulations of chemistry framework. *WIREs Comput. Mol. Sci.* **8**, e1340 (2018).

83. Suzuki, Y. et al. Qulacs: a fast and versatile quantum circuit simulator for research purpose. *Quantum* **5**, 559 (2021).

84. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

85. Qiskit contributors. Qiskit: an open-source framework for quantum computing https://doi.org/10.5281/zenodo.2573505 (2023).

AUTHOR CONTRIBUTIONS

T.T. conceived the idea and wrote the paper. Y.R., S.C.T., and T.T. implemented the algorithms and performed numerical simulations. T.T., Y.R., S.C.T., and S.L.T. all participated in discussions that developed the theory and shaped the project.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41534-023-00780-y.

Correspondence and requests for materials should be addressed to Takashi Tsuchimochi.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.