CHALMERS
UNIVERSITY OF TECHNOLOGY

Orders of magnitude increased accuracy for quantum many-body problems on quantum computers via an exact transcorrelated method

Downloaded from: https://research.chalmers.se, 2024-10-12 02:50 UTC

Citation for the original published paper (version of record):
Sokolov, I., Barucha-Dobrautz, W., Luo, H. et al (2023). Orders of magnitude increased accuracy for quantum many-body problems on quantum computers via
an exact transcorrelated method. Physical Review Research, 5(2).
http://dx.doi.org/10.1103/PhysRevResearch.5.023174

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

Orders of magnitude increased accuracy for quantum many-body problems on quantum computers via an exact transcorrelated method

Igor O. Sokolov,[1,*,†] Werner Dobrautz,[2,3,*,‡] Hongjun Luo,[3] Ali Alavi,[3,4,§] and Ivano Tavernelli[1,¶]
[1]*IBM Quantum, IBM Research Zurich, Säumerstrasse 4, 8803 Rüschlikon, Switzerland*
[2]*Department of Chemistry and Chemical Engineering, Chalmers University of Technology, 41296 Gothenburg, Sweden*
[3]*Max Planck Institute for Solid State Research, Heisenbergstr. 1, 70569 Stuttgart, Germany*
[4]*Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom*

(Received 28 January 2022; accepted 9 May 2023; published 20 June 2023)

Transcorrelated methods provide an efficient way of partially transferring the description of electronic correlations from the ground-state wave function directly into the underlying Hamiltonian. In particular, Dobrautz *et al.* [Phys. Rev. B **99**, 075119 (2019)] have demonstrated that the use of momentum-space representation, combined with a nonunitary similarity transformation, results in a Hubbard Hamiltonian that possesses a significantly more "compact" ground-state wave function, dominated by a single Slater determinant. This compactness/*single-reference* character greatly facilitates electronic structure calculations. As a consequence, however, the Hamiltonian becomes non-Hermitian, posing problems for quantum algorithms based on the variational principle. We overcome these limitations with the *Ansatz*-based quantum imaginary-time evolution algorithm and apply the transcorrelated method in the context of digital quantum computing. We demonstrate that this approach enables up to four orders of magnitude more accurate and compact solutions in various instances of the Hubbard model at intermediate interaction strength ($U/t = 4$), enabling the use of shallower quantum circuits for wave-function *Ansätzes*. In addition, we propose a more efficient implementation of the quantum imaginary-time evolution algorithm in quantum circuits that is tailored to non-Hermitian problems. To validate our approach, we perform hardware experiments on the `ibmq_lima` quantum computer. Our work paves the way for the use of exact transcorrelated methods for the simulations of *ab initio* systems on quantum computers.

DOI: 10.1103/PhysRevResearch.5.023174

I. INTRODUCTION

Understanding and predicting the properties of materials and chemical systems is of paramount importance for the development of natural sciences and technology. To achieve this goal, classical computers are used to solve, at least approximately, the corresponding quantum mechanical equations and extract the quantities of interest. However, performing this type of calculation is a notoriously hard problem since the dimension of the many-body wave function scales exponentially with the number of degrees of freedom (e.g., number of electrons) [1]. This poses an important limitation on the size of accurately simulatable physical systems and makes the majority of them inaccessible on classical computers.

Quantum computing, on the other hand, is emerging as a new computational paradigm for the solution of many classically hard problems, including the solution of the many-body Schrödinger equation of strongly correlated systems. Nowadays, quantum computers are at the forefront of scientific research thanks to ground-breaking hardware demonstrations including (real-time) error mitigation [2] and error-correction schemes [3], paving the way for large near-term quantum calculations on noisy quantum hardware (with physical qubits), followed, in the near future, by fault-tolerant calculations with logical qubits [4,5]. From an algorithmic perspective, quantum computers can provide scaling advantages in the computation of the ground- and excited-state properties (of isolated and periodic systems) [6–10], vibrational structure calculations [11–13], configuration space sampling (such as protein folding) [14,15], molecular and quantum dynamics [16–18], and lattice gauge theory [19,20], just to mention a few.

The currently most popular quantum optimization algorithm for electronic structure calculations is the variational quantum eigensolver (VQE) [7,8,21–23]. It is a well-tested and well-developed hybrid quantum-classical approach: Quantum hardware is used to efficiently represent an arbitrary wave-function *Ansatz* $|\Psi(\boldsymbol{\theta})\rangle$, and measure the expectation value of a chosen observable $\langle\Psi(\boldsymbol{\theta})|\,\hat{O}\,|\Psi(\boldsymbol{\theta})\rangle$, in conjunction with a classical computer which performs the optimization of the (quantum gate) parameters $\boldsymbol{\theta}$ until a chosen cost function of the observable $\hat{O}$ (e.g., the energy in electronic structure calculations) is minimized. However, most

*These authors contributed equally to this work.
†sokolov.igor.ch@gmail.com
‡dobrautz@chalmers.se
§a.alavi@fkf.mpg.de
¶ita@zurich.ibm.com

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

importantly for the scope of this work, the VQE algorithm is applicable only in the case where the cost function that drives the optimization of the parameters is Hermitian.

As an alternative to VQE, imaginary-time evolution [24] can be used to drive an initial wave-function guess towards the optimal ground-state solution. In particular, the variational *Ansatz-based* quantum imaginary-time evolution (VarQITE) algorithm is a powerful method for the calculation of the ground state of quantum systems by simulating the nonunitary dynamics on quantum computers [25]. Using the McLachlan variational principle [26] it is possible to describe time evolution by means of a fixed-length variational circuit, whose parameters evolve according to a well-defined equation of motion which is solved classically. For a detailed analysis of this approach we refer to the original literature [27]. Applications of the VarQITE algorithm include, for instance, the determination of ground and excited states [24,28], the training of quantum machine-learning models [29], the simulation of quantum field theories [30], the solution of linear systems of equations [31,32] and combinatorial optimization problems [33], and the pricing of financial options [34]. Recent algorithmic developments include the addition of adaptive *Ansätzes* [35], hardware-efficient approaches [36], and derivation of robust error bounds [37].

Both the VQE and VarQITE approaches require sufficiently expressive *Ansätzes* for the representation of the targeted ground-state wave function [6,25,27,38,39], which would need exponentially many parameters for an exact solution. However, it was shown that a polynomial number of variational parameters (single-qubit rotations) is sufficient to achieve accurate results within a given error threshold [40,41]. In this framework, we can therefore aim at solving interesting electronic structure problems using state-of-the-art noisy quantum devices (with limited coherence times and uncorrected gate operations) with relatively shallow circuits.

Despite these advancements, one of the major limitations of the application of current quantum algorithms to model systems, like the Hubbard model and the electronic structure problem, is the severely limited number of qubits. To map such problems to quantum computers, the corresponding second-quantized Hamiltonian is typically transformed into a qubit representation using a fermion-to-qubit transformation, such as the Jordan-Wigner transformation [42]. This transformation relates one qubit to each single-particle orbital associated with the system of interest. In the field of quantum chemistry, this is amplified by the need for large basis set expansion (and consequently qubits) to capture dynamic correlation effects induced by the divergence of the Coulomb potential and, therefore, to deal with the nondifferentiable behavior of the electronic wave function at electron coalescence, known as Kato's cusp condition [43,44]. As a consequence, various theories and approaches [45–48] that aim to explicitly capture these dynamic correlation effects and obtain more accurate results in smaller basis sets exist in the field of computational chemistry, the most notable being the explicitly correlated R12 and F12 methods [49–60]. It is a very active field of research to exploit similar approaches in the field of quantum computing to reduce the quantum resources necessary to obtain accurate results for realistic systems on near-term quantum devices [61–65].

In this work we use the so-called *transcorrelated* (TC) method, introduced by Hirschfelder [66] and Boys and Handy [67–69], and apply it to a lattice Hamiltonian in the form of the Hubbard model. The TC method was originally conceived in the field of *ab initio* quantum chemistry to exactly incorporate electronic correlation effects, via a correlated Jastrow *Ansatz* [70] for the wave function, directly into the Hamiltonian by a similarity transformation (ST). In molecular and *ab initio* TC approaches, one typically works with incomplete bases (finite basis sets), where the main benefit of TC is to reach the complete basis set limit quicker (i.e., higher accuracy is reached with a smaller basis set) [71,72]. However, the same basic concept can be applied to any problem, thus facilitating the subsequent solution. In the case of the Hubbard model in conjunction with a Gutzwiller correlator [73,74], the ST introduces higher-order interaction terms and renders the Hamiltonian non-Hermitian. Nevertheless, in contrast to *ab initio* systems, the Hubbard model (in its original one-band formulation) is defined as a lattice model with one orbital per lattice site. In such a complete basis, the ST does not change the spectrum of the Hamiltonian. In this case, the benefit of the TC method is to make the solution more compact [75], and thus easier to represent with shallow quantum circuits, while yielding the same accuracy. As a consequence of the non-Hermiticity, the variational principle does not apply anymore, and we have to rely on methods like VarQITE to solve the problem on a quantum computer. (Another possible approach is to use VQE based on the variance cost function [76,77], but this would require to square the Hamiltonian and hence presents a significant overhead in terms of measurements [78].)

Following the approach of Dobrautz *et al.* [75], McArdle and Tew [78] recently investigated the exact non-Hermitian TC formulation of the Hubbard Hamiltonian in the real-space representation in the context of quantum computing. They investigated the beneficial effect of the TC approach on the quantum footprint that is caused by a more compact/single-reference right eigenvector. The advantages were demonstrated in numerical simulations using the VarQITE algorithm, however, without the complete implementation of the corresponding quantum circuits. Additionally, McArdle and Tew used the TC approach in the real-space representation of the Hubbard model, which, however, does not display the same level of compactification of the ground-state wave function as in a momentum-space representation [75].

In this work, we expand and complement the study of McArdle and Tew [78] by implementing the full algorithm as would be executed on a quantum computer. To validate our approach, we perform experiments on the IBM quantum computer `ibmq_lima`. In addition, following Dobrautz *et al.* [75], we investigate the formulation of the TC Hubbard Hamiltonian in the momentum space, for which we expect an increased efficiency of our method (compared to real space) while approaching the complexity of *ab initio* chemical problems (i.e., the presence of a three-body term). Thus, this study on the TC Hubbard model in the momentum space will pave the way for future extensions to more general *ab initio* Hamiltonians.

The paper is structured as follows. In Sec. II A, we summarize the general theory of the exact TC method. The (TC)

Hubbard Hamiltonians in real and reciprocal spaces are defined in Sec. II B. In Sec. II C, we discuss the application of the VarQITE algorithm to a non-Hermitian problem. The methods, including the details of the implementation in quantum computers, are given in Sec. III. We discuss the results of our experiments and simulations in Sec. IV. Finally, in Sec. V, we present our conclusions on the advantages and limitations of the exact TC method and present our views on future developments.

II. THEORY

In this section, we review the transcorrelated approach in the classical and quantum frameworks and introduce methods for the optimization of the ground-state wave function.

A. Transcorrelated method

The transcorrelated method was introduced by Boys and Handy [67–69], who suggested incorporating the effect of a correlated wave-function *Ansatz*, in the form of a Jastrow *Ansatz* [70]

$$|\Psi\rangle = e^{\hat{g}}\,|\Phi\rangle\,, \tag{1}$$

directly into the many-body fermionic Hamiltonian via a similarity transformation

$$\hat{H} \rightarrow e^{-\hat{g}}\hat{H}e^{\hat{g}} = \hat{H}_{\mathrm{TC}}. \tag{2}$$

In the work of Boys and Handy, $\hat{g}(\boldsymbol{r})$ represents a pairwise symmetric real function dependent on the interelectronic distances with n electrons located at $\boldsymbol{r} = (r_1, r_2, \ldots, r_n)$ coordinates, which is able to exactly incorporate the electronic cusp condition [43]. In the original work, Boys and Handy used a single Slater determinant (SD), $|\Phi\rangle = |\phi_0\rangle$, and optimized both the single-particle orbitals comprising $|\phi_0\rangle$ as well as the terms in the Jastrow factor $\hat{g}(\mathbf{r})$. In this work, we follow the approach of Dobrautz *et al.* [75] and use a previously optimized fixed Jastrow factor, but allow complete flexibility to the wave-function expansion $|\Phi\rangle = \sum c_i\,|\phi_i\rangle$. Using a fixed Jastrow factor, but allowing a full flexibility to the SD expansion of the fermionic many-body wave function in the TC approach was for the first time studied by Luo and Alavi [79] for the homogeneous electron gas, Dobrautz *et al.* [75] for the Hubbard model, Cohen *et al.* [71] for the *ab initio* treatment of the first row atoms, and Guther *et al.* [80] for the binding curve of the beryllium dimer by combination with the full configuration quantum Monte Carlo (FCIQMC) method [81–85]. In a complete basis, the ST, Eq. (2), does not change the spectrum of $\hat{H}$:

$$e^{-\hat{g}}E\,|\Psi\rangle = e^{-\hat{g}}\hat{H}\,|\Psi\rangle = e^{-\hat{g}}\hat{H}e^{\hat{g}}\,|\Phi\rangle = \hat{H}_{\mathrm{TC}}\,|\Phi\rangle = E\,|\Phi\rangle\,. \tag{3}$$

However, as the transformed $\hat{H}_{\mathrm{TC}}$ is not Hermitian anymore (since $\hat{g}^{\dagger} = \hat{g}$) it possesses different left and right eigenvectors, which form a biorthogonal basis with $\langle\Psi_i^L|\Psi_j^R\rangle = 0$ for $i \neq j$. The loss of unitarity and the variational principle seems like a high price to pay, as standard methods in conventional computational chemistry and physics, as well as in the field of quantum computing, like the phase estimation algorithm [86,87] and VQE [21,22], are not applicable anymore. However, Dobrautz *et al.* [75], found that the TC approach leads to more compact and single-reference right eigenvectors, with dramatic positive effects on projective methods. Consequently, Motta *et al.* [88], and recently Schleich *et al.* [63] and Kumar *et al.* [89], were able to show the benefits of similar approaches on a quantum device, by reducing the problem complexity to achieve a desired accuracy. However, these studies targeted *ab initio* systems and, more importantly, they used an approximated transcorrelated approach, which overcomes the problems associated to a non-Hermitian Hamiltonian with three-body terms, at the cost of a reduction in accuracy. As a representative example, we study the exact TC approach applied to the Hubbard model and show the benefits of a correlated wave-function *Ansatz* to achieve accurate results with fewer quantum resources and introduce an efficient approach to study non-Hermitian problems with the VarQITE algorithm in general.

In the next section, we define the Hubbard model Hamiltionian in the real- and momentum-space representations, including their exact TC versions. Then, the VarQITE algorithm is presented with the corresponding quantum circuits.

B. Hubbard Hamiltonian and Gutzwiller Ansatz

The fermionic Hubbard model [73,90–92] is an extensively studied minimal model of itinerant strongly correlated electrons. Despite its simplicity, it possesses a rich phase diagram and is used to study the physics of high-temperature cuprate superconductors [93–96]. Exact solutions only exist in the limit of one [97,98] and infinite dimensions [99–101], while the study of the two-dimensional model is a very active field of research [102–105]. The real-space representation (r superscript) of the Hubbard Hamiltonian for a two-dimensional lattice is given by

$$\hat{H}^r = -t\sum_{\langle \boldsymbol{i},\boldsymbol{j}\rangle}\sum_{\sigma}\hat{a}_{\boldsymbol{i},\sigma}^{\dagger}\hat{a}_{\boldsymbol{j},\sigma} + U\sum_{\boldsymbol{i}}^{N}\hat{n}_{\boldsymbol{i},\uparrow}\hat{n}_{\boldsymbol{i},\downarrow}, \tag{4}$$

where the indices $\boldsymbol{i} = (i_x, i_y)$ and $\boldsymbol{j} = (j_x, j_y)$ indicate the real-space lattice positions, $\langle \boldsymbol{i}, \boldsymbol{j}\rangle$ denotes a summation over nearest neighbors, and N is the number of lattice sites. $\hat{a}_{\boldsymbol{i},\sigma}^{\dagger}$ is the creation operator of an electron on site $\boldsymbol{i}$ with spin $\sigma \in \{\uparrow, \downarrow\}$, while $\hat{a}_{\boldsymbol{i},\sigma}$ and $\hat{n}_{\boldsymbol{i},\sigma} = \hat{a}_{\boldsymbol{i},\sigma}^{\dagger}\hat{a}_{\boldsymbol{i},\sigma}$ are the corresponding electronic annihilation and number operators. The first term in Eq. (4) represents the electron hopping while the second one denotes the electron interaction with associated parameters $t \geqslant 0$ and $U \geqslant 0$, for the fermionic Hubbard model. The U/t ratio defines their relative strength and also the character of the ground state (single reference or multireference) with intermediate values, $4 \lesssim U/t \lesssim 12$, corresponding to the strongly correlated regime with a multireference ground state. Following the general convention, energies are given in units of t, and thus the Coulomb repulsion strength U remains the sole parameter of the model.

Substituting the Fourier transform of the electronic creation and annihilation operators $\hat{c}_{\boldsymbol{k},\sigma}^{\dagger} = 1/\sqrt{N}\sum_{\boldsymbol{r}} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\hat{a}_{\boldsymbol{r},\sigma}^{\dagger}$ and $\hat{c}_{\boldsymbol{k},\sigma} = 1/\sqrt{N}\sum_{\boldsymbol{r}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\hat{a}_{\boldsymbol{r},\sigma}$ into Eq. (4) yields the momentum-space representation (m superscript) of the

Hubbard model

$$\hat{H}^m = \sum_{\boldsymbol{k},\sigma} \epsilon_{\boldsymbol{k}} \hat{n}_{\boldsymbol{k},\sigma} + \frac{U}{2N} \sum_{\boldsymbol{p},\boldsymbol{q},\boldsymbol{k},\sigma} \hat{c}^\dagger_{\boldsymbol{p}-\boldsymbol{k},\sigma} \hat{c}^\dagger_{\boldsymbol{q}+\boldsymbol{k},\bar{\sigma}} \hat{c}_{\boldsymbol{q},\bar{\sigma}} \hat{c}_{\boldsymbol{p},\sigma}, \tag{5}$$

where $\hat{c}^\dagger_{\boldsymbol{k},\sigma}$ and $\hat{c}_{\boldsymbol{k},\sigma}$ operators, respectively, create and annihilate an electron with momentum $\boldsymbol{k} = (k_x, k_y)$ and spin σ; the opposite spin to σ is denoted as $\bar{\sigma}$. For a two-dimensional square lattice, the dispersion relation is given by $\epsilon_{\boldsymbol{k}} = -2t[\cos(k_x) + \cos(k_y)]$. In one dimension, it is given by $\epsilon_k = -2t\ \cos(k_x)$. The Hamiltonians in the real and momentum space contain up to two-body interactions. Due to the diagonal two-body part in Eq. (4) and the restriction to nearest-neighbor hopping, the number of terms in the real-space Hubbard model scale as $O(N)$. On the other hand, because of momentum conservation, the terms of the k-space Hubbard model (5) scale as $O(N^3)$.

Next, we present the TC version of the real-space Hubbard Hamiltonian defined in Eq. (4). Following Tsuneyuki [77] and Dobrautz *et al.* [75], we use a Gutzwiller correlator [73,74,106]

$$\hat{g} = J \sum_{\mathbf{i}} \hat{n}_{\mathbf{i},\uparrow} \hat{n}_{\mathbf{i},\downarrow} \tag{6}$$

for our correlated wave-function *Ansatz*. The action of Eq. (6) is the same as the two-body part of the Hubbard Hamiltonian in the real space [see Eq. (4)] and counts the number of doubly occupied sites in a state $|\phi_i\rangle$, weighted with an optimizable parameter J. The Gutzwiller *Ansatz* is a widely studied approach to solve the Hubbard model [93,99,107,108], where the parameter J is usually optimized to minimize the energy with variational Monte Carlo (VMC) methods [109,110]. Although it misses important correlations, especially in the large-U regime [111–113], it does provide good energy estimates for the low- to intermediate-interaction strengths. In this parameter regime, the use of the momentum-space formulation of the Hubbard model is preferable, as the Fermi-sea (Hartree-Fock) determinant provides a good (single-)reference state for the ground-state wave function.

With the Gutwiller *Ansatz* (6), the corresponding TC Hamiltonian (2) can be expressed in closed form, using the Baker-Campbell-Hausdorff formula exactly resummed to all orders. The resulting TC Hamiltonian is derived in Refs. [75,77,114] and given by

$$\begin{aligned}\hat{H}^r_{\mathrm{TC}} = \hat{H}^r - t \sum_{\langle i,j\rangle,\sigma} \hat{a}^\dagger_{i,\sigma} \hat{a}_{j,\sigma} [(e^J - 1)\hat{n}_{j,\bar{\sigma}} \\ + (e^{-J} - 1)\hat{n}_{i,\bar{\sigma}} - 2[\cosh(J) - 1]\hat{n}_{i,\bar{\sigma}} \hat{n}_{j,\bar{\sigma}}],\end{aligned} \tag{7}$$

with $\hat{H}^r$ being the original real-space Hubbard Hamiltonian (4). In contrast to the approximate unitary version of the TC approach [88], this transformation (7) is exact. An equivalent Hamiltonian can be written in the momentum space by applying the Fourier transform of the fermionic operators as was done for Eq. (5) with details given in Ref. [75]. The Hamiltonian defined in Eq. (7) reads in the momentum space as

$$\begin{aligned}\hat{H}^m_{\mathrm{TC}} = \hat{H}^m - \sum_{\boldsymbol{p},\boldsymbol{q},\boldsymbol{k},\sigma} D_{\mathbf{p},\mathbf{q},\mathbf{k}} \hat{c}^\dagger_{\boldsymbol{p}-\boldsymbol{k},\sigma} \hat{c}^\dagger_{\boldsymbol{q}+\boldsymbol{k},\bar{\sigma}} \hat{c}_{\boldsymbol{q},\bar{\sigma}} \hat{c}_{\boldsymbol{p},\sigma} \\ + T \sum_{\substack{\boldsymbol{p},\boldsymbol{q},\boldsymbol{s},\boldsymbol{k},\boldsymbol{k}',\sigma \\ \boldsymbol{p}'=\boldsymbol{p}-\boldsymbol{k}+\boldsymbol{k}'}} \epsilon_{\boldsymbol{p}'} \hat{c}^\dagger_{\boldsymbol{p}-\boldsymbol{k},\sigma} \hat{c}^\dagger_{\boldsymbol{q}+\boldsymbol{k}',\bar{\sigma}} \hat{c}^\dagger_{\boldsymbol{s}+\boldsymbol{k}-\boldsymbol{k}',\bar{\sigma}} \hat{c}_{\boldsymbol{s},\bar{\sigma}} \hat{c}_{\boldsymbol{q},\bar{\sigma}} \hat{c}_{\boldsymbol{p},\sigma}\end{aligned} \tag{8}$$

with $D_{\mathbf{p},\mathbf{q},\mathbf{k}} = \frac{t}{N}[(e^J - 1)\epsilon_{\boldsymbol{p}-\boldsymbol{k}} + (e^{-J} - 1)\epsilon_{\boldsymbol{p}}]$, $T = 2t\,\frac{\cosh(J)-1}{N^2}$, and $\hat{H}^m$ being the original momentum-space Hubbard Hamiltonian (5). The much more compact right eigenvector of the Hamiltonian $\hat{H}^m_{\mathrm{TC}}$ [75] [Eq. (8)] allowed the limited applicability of FCIQMC to be extended to lattice models [115,116]. Both the TC real- and momentum-space Hubbard Hamiltonians (7) and (8) are non-Hermitian, due to the modified two-body term, and have up to three-body interactions. The number of terms in the real-space TC Hubbard Hamiltonian (7) still scales linearly with the number of sites N, while in the TC momentum-space case, Eq. (8), we end up with $O(N^5)$ terms.

C. Quantum imaginary-time evolution

The (normalized) imaginary-time evolution is defined as

$$|\Phi(\tau)\rangle = \frac{e^{-\hat{H}\tau}|\Phi(0)\rangle}{\sqrt{\langle\Phi(0)|e^{-2\hat{H}\tau}|\Phi(0)\rangle}}, \tag{9}$$

where $|\Phi(0)\rangle$ is some initial state. In the infinite-time limit, the ground state of $\hat{H}$ is obtained only if $|\Phi(0)\rangle$ and that ground state have a nonzero overlap. Note that this is also valid for the non-Hermitian Hamiltonians [78]. To implement the nonunitary evolution defined in Eq. (9) on a quantum computer, the Wick-rotated Schrödinger equation can be written as

$$\frac{\partial\,|\Phi(\tau)\rangle}{\partial\tau} = -(\hat{H} - E)\,|\Phi(\tau)\rangle\,, \tag{10}$$

where $\tau = it$ is the imaginary time and $E = \mathrm{Re}[\langle\Phi(\tau)|\hat{H}|\Phi(\tau)\rangle]$ is the energy of the system. McLachlan's variational principle applied to Eq. (10) yields

$$\delta\|(\partial/\partial\tau + \hat{H} - E)\,|\Phi(\tau)\rangle\,\| = 0. \tag{11}$$

This equation can be defined for each variational parameter θ_i, $\frac{\partial}{\partial\theta_i}\|(\partial/\partial\tau + \hat{H} - E)\,|\Phi(\tau)\rangle\,\| = 0$, where we assume the dependence $\theta_i(\tau)$ for $i = \{0, \ldots, N_p - 1\}$ and N_p is the number of variational parameters. Its solution leads to a system of equations

$$\boldsymbol{A}\dot{\boldsymbol{\theta}} = -\boldsymbol{C}, \tag{12}$$

with the matrix $\boldsymbol{A}$ given by its elements

$$\begin{aligned}A_{ij} &= \frac{1}{2}\left(\frac{\partial\,\langle\Phi(\tau)|}{\partial\theta_i}\frac{\partial\,|\Phi(\tau)\rangle}{\partial\theta_j} + \frac{\partial\,\langle\Phi(\tau)|}{\partial\theta_j}\frac{\partial\,|\Phi(\tau)\rangle}{\partial\theta_i}\right) \\ &= \mathrm{Re}\left(\frac{\partial\,\langle\Phi(\tau)|}{\partial\theta_i}\frac{\partial\,|\Phi(\tau)\rangle}{\partial\theta_j}\right),\end{aligned} \tag{13}$$

and the gradient $\boldsymbol{C}$, with its elements given by

$$\begin{aligned} C_i &= \frac{1}{2}\left(\frac{\partial \langle\Phi(\tau)|}{\partial\theta_i}\hat{H}\,|\Phi(\tau)\rangle + \langle\Phi(\tau)|\,\hat{H}^\dagger\frac{\partial\,|\Phi(\tau)\rangle}{\partial\theta_i}\right) \\ &\quad - \frac{1}{2}\left(E\frac{\partial \langle\Phi(\tau)|}{\partial\theta_i}\,|\Phi(\tau)\rangle + E^*\,\langle\Phi(\tau)|\,\frac{\partial\,|\Phi(\tau)\rangle}{\partial\theta_i}\right) \\ &= \mathrm{Re}\left(\frac{\partial \langle\Phi(\tau)|}{\partial\theta_i}\hat{H}\,|\Phi(\tau)\rangle\right). \end{aligned} \tag{14}$$

We stress once more that although the similarity transformed Hamiltonian is non-Hermitian, it has an unchanged spectrum in a complete basis. Additionally, since all operators and coefficients in the Hamiltonians (4) and (5) and the Gutzwiller *Ansatz* (6) are real, the energy expectation value of any *real-valued* wave-function *Ansatz*, $\langle\Phi(\tau)|\hat{H}^{r/m}_{(\mathrm{TC})}|\Phi(\tau)\rangle$, for both the original and transcorrelated versions, remains real. Thus, there are no contributions from the terms in the second line of Eq. (14) as shown in Ref. [78]. Equation (12) defines the imaginary evolution of the wave function projected onto the space of all possible states that can be represented by a given *Ansatz*, the so-called *Ansatz* space [25]. The state evolution is guided not only by the gradients $\boldsymbol{C}$ but also the metric in the parameter space $\boldsymbol{A}$, which takes into account the structure of the *Ansatz* [27]. The Euler method is then employed to update the variational parameters $\boldsymbol{\theta}(k)$ at iteration k as

$$\boldsymbol{\theta}(k+1) = \boldsymbol{\theta}(k) - \Delta t\boldsymbol{A}^{-1}\boldsymbol{C}. \tag{15}$$

The scaling of the algorithm in terms of measurements is $O(N_C N_p N_H + N_A N_p^2)$ where N_C, N_A are the number of measurements to obtain a required accuracy for $\boldsymbol{C}$ and $\boldsymbol{A}$ matrix elements, respectively, and N_H is the number of terms in the Hamiltonian. Despite the large number of measurements, VarQITE guarantees the convergence to the ground state of non-Hermitian Hamiltonians where VQE algorithms would require the use of the variance as the cost function, requiring to square the Hamiltonian. The VarQITE algorithm requires the inversion of matrix $\boldsymbol{A}$ [or the solution of the linear system in Eq. (15)] with, for instance, the Tikhonov regularization [25] that stabilizes the evolution of variational parameters. These steps present potential sources of instabilities for the simulation. Recently, inversion- and regularization-free approaches were also proposed [37] by formulating the equation of VarQITE as a quadratic optimization problem. Despite possessing an error-prone classical optimization step, the advantage resides in quantifiable error bounds. However, for our systems, the standard VarQITE with Tikhonov regularization performed best, thus it is used in the rest of this work. See Appendix B for additional details.

III. METHODS

The TC method necessitates the determination of the optimal value of the parameter J associated to the Gutzwiller *Ansatz*. For its optimization we can use two independent methods: (1) an efficient representation-independent VMC procedure (polynomially scaling in time) as described in Refs. [117–123] and (2) an even cheaper projection method of the TC Hamiltonian, inspired from the coupled cluster amplitude equations [75,114] (with a single amplitude J in this case), in the momentum-space formulation (see Appendix C for more details). Throughout the work, we optimize J for the half-filled ground state (see Fig. 1 for a sketch). In this case, both methods yield similar values of J for which the right eigenvector of the momentum-space TC Hamiltonian, $\hat{H}^m_{\mathrm{TC}}(J)$, is most "compact" [75], meaning that largest component of the wave function is represented by the Hartree-Fock/Fermi-sea state. As the VMC procedure is independent of the basis, we also use the same value of J for the real-space TC calculation, where the right eigenvector has a similar, albeit less pronounced, compact character [77,78].

All quantum simulations are performed with QISKIT [125]. Hamiltonians are mapped to the qubit space using the Jordan-Wigner transformation [42], that allows to express them as $\hat{H} = \sum_i l_i\hat{P}_i$ where $\hat{P}_i$ denotes a Pauli string (tensor product of Pauli operators), and l_i is the associated (complex) coefficient. Calculations are performed using the matrix and state-vector representations (SV) for the Hamiltonian and the wave-function *Ansatz*. They represent the idealistic simulations that could be obtained without the hardware noise and in the infinite number of measurements limit. In addition, we perform simulations that include the realistic noise model of the `ibmq_lima` quantum chip. We use the quantum assembly language (QASM) description of the operators (represented by a sum of Pauli strings) as well as the wave functions (represented by quantum circuits). For additional details about the device and its noise model, see Appendix G. For both hardware and QASM simulations, we employ the readout error mitigation [126] as implemented in QISKIT.

In the VQE simulations, the optimization of variational parameters is performed by means of a classical optimization algorithm: limited-memory Broyden-Fletcher-Goldfarb-Shanno with boundary constraints (L-BFGS-B) [127] with the convergence criterion set to 10^{-7}. In the VarQITE and SV simulations, the derivatives of wave functions with respect to variational parameters are obtained using the forward finite-differences method [128] with the step size of 10^{-9}.

To demonstrate the potential of our algorithm, we will use the quantum unitary coupled-cluster singles doubles (qUCCSD) *Ansatz* to approximate the ground state in SV simulations (see Appendix D for additional details). As we will show below, the benefit of the TC method consists in a reduction of the required circuit depth due to a more compact ground-state wave function, which is independent from the nature of the chosen *Ansatz*. In the following examples, we will apply the qUCCSD *Ansatz* as it is a widely used wave-function form in current quantum computing literature, especially in solid-state physics and electronic structure theory. Additionally, the UCCSD *Ansatz* is very suited for the momentum-space Hubbard model, particularly in the case of small-U values, as the ground state is dominated by a single SD. On the other hand, other recently developed *Ansätze* like the variational Hamiltonian *Ansatz* [129] could also be employed within the same approach. The qUCCSD cluster operator is first written as a quantum circuit (see Ref. [130]), subsequently transformed into a unitary matrix, and finally applied on an initial state vector. For the latter, the ground state of the noninteracting Hubbard model ($t = 1$, $U = 0$) is

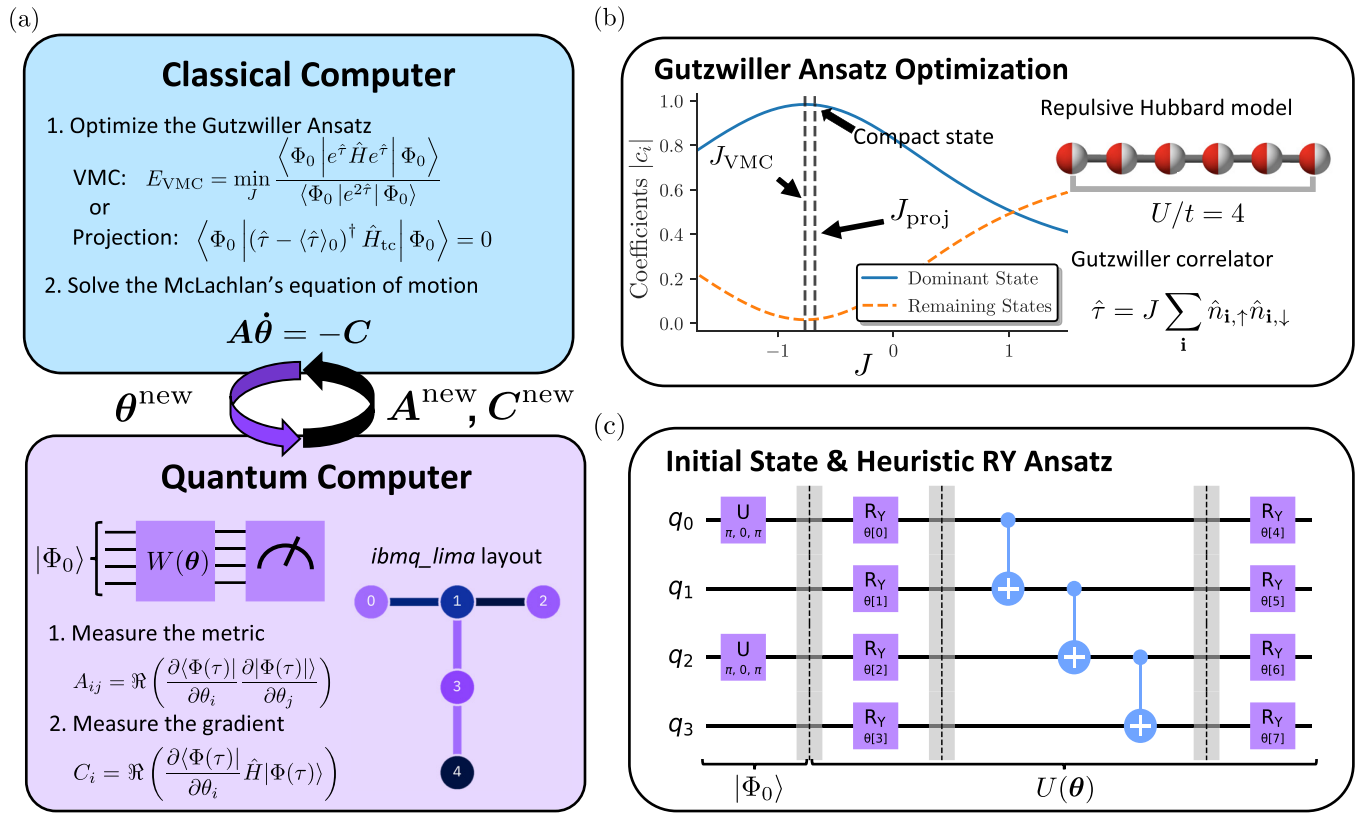

FIG. 1. (a) Hybrid quantum-classical procedure employed to perform VarQITE. First, the Gutzwiller *Ansatz* is optimized by VMC or projection on a classical computer. The resulting TC Hubbard Hamiltonian is represented as a sum of Pauli strings by a fermion-to-qubit mapping (e.g., Jordan-Wigner). The initial parameter values $\boldsymbol{\theta}$ and the list of necessary Pauli-string measurements are sent to the quantum computer. Given a wave-function *Ansatz* $|\Phi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})|\Phi_0\rangle$, the matrix elements of the metric $A_{ij}^{\rm new}$ and the gradient $C_i^{\rm new}$ are measured using the differentiation of general gates via a linear combination of unitaries represented as $W(\boldsymbol{\theta})$ [124]. Finally, the linear system derived from McLachlan's variational principle is solved and by applying the Euler method, we obtain the parameters $\boldsymbol{\theta}^{\rm new}$ of the next time step. The whole procedure is repeated until convergence (i.e., Euclidean norm of the gradient is below a set threshold). (b) Optimization of the Jastrow parameter J. We show that when $J \approx J_{\rm proj}$ or $J \approx J_{\rm VMC}$, the ground state of six-site TC momentum-space Hubbard model is "compact," meaning almost the entire weight of the wave function is concentrated in the Hartree-Fock (HF) state. $J_{\rm VMC}$ and $J_{\rm proj}$ denote the results of a VMC simulation and the solution of the projective scheme, respectively. (c) Heuristic RY unitary operator $U(\boldsymbol{\theta})$ applied on the HF initial state $|\Psi_0\rangle$ of the two-site Hubbard model. In purple are the single-qubit rotations and in blue are the CNOTs. The definitions of the single-qubit gates are given in Appendix A. Exactly this *Ansatz* and initial state are employed in the hardware experiments on the `ibmq_lima` chip [see its layout in (a) where numbers denote the qubits].

chosen as the starting state. It provides a good initial guess and it can be efficiently obtained classically using methods described in Ref. [75].

Due to the high gate number and large circuit depth the qUCCSD *Ansatz* is too "costly" for current quantum hardware limitations. Thus, for the QASM simulations and real hardware experiments (HW), we will use a hardware-efficient Ansatz [131] composed of single-qubit rotation gates around the y-axis (RY gates) with controlled-NOT (CNOT) entangling layers, optimized for the particular topology of the hardware. The complete circuit is given in Fig. 1(c) and the definitions of quantum gates in Appendix A. The detailed discussion about our choices of *Ansätzes* is reported in Appendix D.

Particular care is needed for the evaluation of the matrix elements A_{ij} [Eq. (13)] and C_i [Eq. (14)] required for the optimization of the parameters $\boldsymbol{\theta}$ according to Eq. (15). Quantum circuits for the evaluation of the matrix elements containing partial derivatives of the state wave function with respect to the parameters are well known only for Hermitian operators. A typical circuit $W_2(\boldsymbol{\theta})$ for the calculation of the term

$$2C_i = 2\,\mathrm{Re}\langle\partial_{\theta_i}\Phi|\hat H|\Phi\rangle = \langle\partial_{\theta_i}\Phi|\hat H|\Phi\rangle + \langle\Phi|\hat H^\dagger|\partial_{\theta_i}\Phi\rangle \quad (16)$$

is given in Fig. 2 with $V = H$, a Hadamard gate.

However, in the TC case, $\hat H_{\rm TC}$ is non-Hermitian and therefore such approach is not applicable. Recently, McArdle *et al.* [78] proposed a method for the evaluation of the matrix elements A_{ij} and C_i. To compute C_i elements, they make use of independent circuits $W_1(\boldsymbol{\theta})$, which include control operations associated to each term of the system Hamiltonian (see Fig. 3). Unfortunately, the costs associated to the implementation of the corresponding circuits in hardware calculations on current noisy quantum processors are prohibitively large and therefore not applicable in practice.

In this work, we designed instead a different strategy based on the decomposition of the non-Hermitian TC Hamiltonian

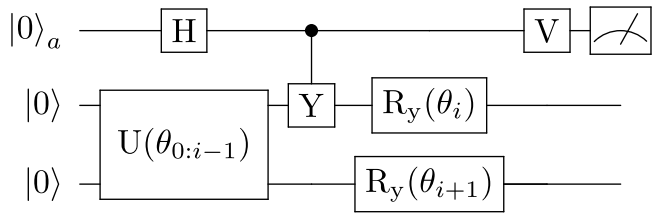

FIG. 2. Quantum circuit $W_2(\boldsymbol{\theta})$ used to calculate the $2C_i$ term in the non-Hermitian case by separation of the TC Hamiltonian into Hermitian and anti-Hermitian parts. We define that $V = H$ and $V = R_x(\frac{\pi}{2})$ for the Hermitian and anti-Hermitian parts of a TC Hamiltonian, respectively. This circuit should be repeated for every term of the Hamiltonian. Note that no controlled Hamiltonian term is present.

into its Hermitian and anti-Hermitian components. We first define the two (Hermitian and anti-Hermitian) operators

$$\hat{H}^{+}_{\mathrm{TC}} = \hat{H}_{\mathrm{TC}} + \hat{H}^{\dagger}_{\mathrm{TC}}, \tag{17}$$

$$\hat{H}^{-}_{\mathrm{TC}} = \hat{H}_{\mathrm{TC}} - \hat{H}^{\dagger}_{\mathrm{TC}}. \tag{18}$$

We then compute the coefficients C_i as

$$C_i = \frac{1}{2}\big(\langle\partial_{\theta_i}\Phi|\hat{H}_{\mathrm{TC}}|\Phi\rangle + \langle\Phi|\hat{H}^{\dagger}_{\mathrm{TC}}|\partial_{\theta_i}\Phi\rangle\big) = \frac{C_i^{+} + C_i^{-}}{4}, \tag{19}$$

where

$$C_i^{+} = \langle\partial_{\theta_i}\Phi|\hat{H}^{+}_{\mathrm{TC}}|\Phi\rangle + \langle\Phi|\hat{H}^{+}_{\mathrm{TC}}|\partial_{\theta_i}\Phi\rangle = 2\mathrm{Re}\langle\partial_{\theta_i}\Phi|\hat{H}^{+}_{\mathrm{TC}}|\Phi\rangle \tag{20}$$

and

$$C_i^{-} = \langle\partial_{\theta_i}\Phi|\hat{H}^{-}_{\mathrm{TC}}|\Phi\rangle - \langle\Phi|\hat{H}^{-}_{\mathrm{TC}}|\partial_{\theta_i}\Phi\rangle = 2\,\mathrm{Re}\langle\partial_{\theta_i}\Phi|\hat{H}^{-}_{\mathrm{TC}}|\Phi\rangle. \tag{21}$$

Following this strategy, we can now implement the calculation of vector elements C_i using two circuits of the form $W_2(\boldsymbol{\theta})$ given in Fig. 2, one for the Hermitian (20) and one for the anti-Hermitian (21) component of the transcorrelated operator $\hat{H}_{\mathrm{TC}}$. For detailed derivations, see Appendixes E and F. The measurements of the A_{ij} matrix elements are performed in the standard way (since they are independent of the Hamiltonian) and can be found in Refs. [8,27].

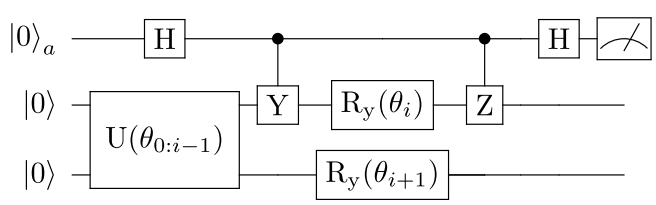

FIG. 3. Quantum circuit $W_1(\boldsymbol{\theta})$ used to calculate the $2C_i$ term in the non-Hermitian case when, for simplicity, the Hamiltonian is given by a single Pauli on the ith qubit, $\hat{H} = 1^{(0)} \otimes \cdots \otimes 1^{(i-1)} Z^{(i)} 1^{(i+1)} \otimes \cdots \otimes 1^{(N_q-1)}$ (controlled-Z gate) where N_q denotes the number of qubits. This circuit should be repeated and adapted for every term of the Hamiltonian (controlled Pauli string) leading to $O(N^6)$ different circuits to be measured in TC cases.

IV. RESULTS AND DISCUSSION

A. Simulations

In this section, we demonstrate the advantages of using the transcorrelated versions of the Hubbard model both in the real and momentum space. As first accessible test cases, we considered the two-, four-, and six-site 1D Hubbard models. It is important to mention that due to hardware and software limitations, validation of quantum computing algorithms is currently restricted to rather small system sizes. For this reason, in this work we decided to restrict ourselves to the study of the 1D Hubbard model, even though it is analytically solvable. Furthermore, accessible 2D systems like the 2×2 and the 2×3 lattice models are anyway dominated by finite-size effects and can be recast into a folded 1D chain.

For each system, we perform VarQITE simulations to obtain a ground-state estimate $|\Phi\rangle$ and quantify the performance in terms of the *absolute energy error*, $|\Delta E| = |E - E_{\mathrm{exact}}|$, and the *infidelity*, $I = 1 - |\langle\Phi|\Phi_{\mathrm{exact}}\rangle|^2$, with respect to the targeted exact ground states at half-filling. We denote the latter with $|\Phi^r_{\mathrm{exact}}\rangle$ and $|\Phi^m_{\mathrm{exact}}\rangle$ for the real- and the momentum-space representations, respectively, and compute them by exactly diagonalizing the corresponding Hamiltonian. Throughout the rest of this work, we assume that our results correspond to SV-type simulations unless specified otherwise and specify energies in units of the Hubbard parameter t. For each system, we initialize the VarQITE algorithm at the mean-field solution of the noninteracting ($U/t = 0$) non-TC Hubbard model (Fermi-sea/Hartree-Fock solution). More specifically, the initial states of the calculations using the TC Hamiltonians are taken to be the same as for the non-TC cases and correspond to the $U/t = 0$ solutions of the Hubbard Hamiltonians in the real space $|\Phi^r_0\rangle$ and the momentum space $|\Phi^m_0\rangle$. For hardware experiments, an inexpensive short-depth VQE calculation can be used for state initialization [i.e., $|\Phi^{r/m}_0(\boldsymbol{\theta}_{\min})\rangle$ from $\min_{\boldsymbol{\theta}}\langle\Phi(\boldsymbol{\theta})|\hat{H}^{r/m}_{U=0}|\Phi(\boldsymbol{\theta})\rangle$] with a suitable *Ansatz* and initial state [see Fig. 1(c)]. To assess the optimal time step t_s for the VarQITE algorithm to reach the required accuracy, we performed series of SV test calculations, which led to a choice of $\Delta t_s = 10^{-1}$ valid for all investigated systems. Unlike VQE approaches, the global phase of an *Ansatz* can potentially affect the end result in variational *Ansatz*-based time evolution (VarQTE) algorithms [27]. The reason is that the derivatives of variational parameters can differ when the global phase is omitted. Hence, similarly to the work of McArdle *et al.* [78], we included the parametrized global phase $e^{i\theta}$ by adding a sequence of single-qubit gates, as reported in Appendix D. Other technical details are summarized in Sec. III.

In Fig. 4, we show the results of VarQITE simulations in SV formulation of a two-, four-, and six-site repulsive Hubbard models with periodic boundary conditions at intermediate interaction strength $U/t = 4$.

After a first inspection, we can already advance the following general main observations: First, the VarQITE algorithm can be efficiently used to optimize the ground state of the Hubbard model, both in its original Hermitian formulation, as well as in the non-Hermitian TC form, in the real and momentum space. Second, we observe a clear advantage of the momentum-space representation of the Hubbard model

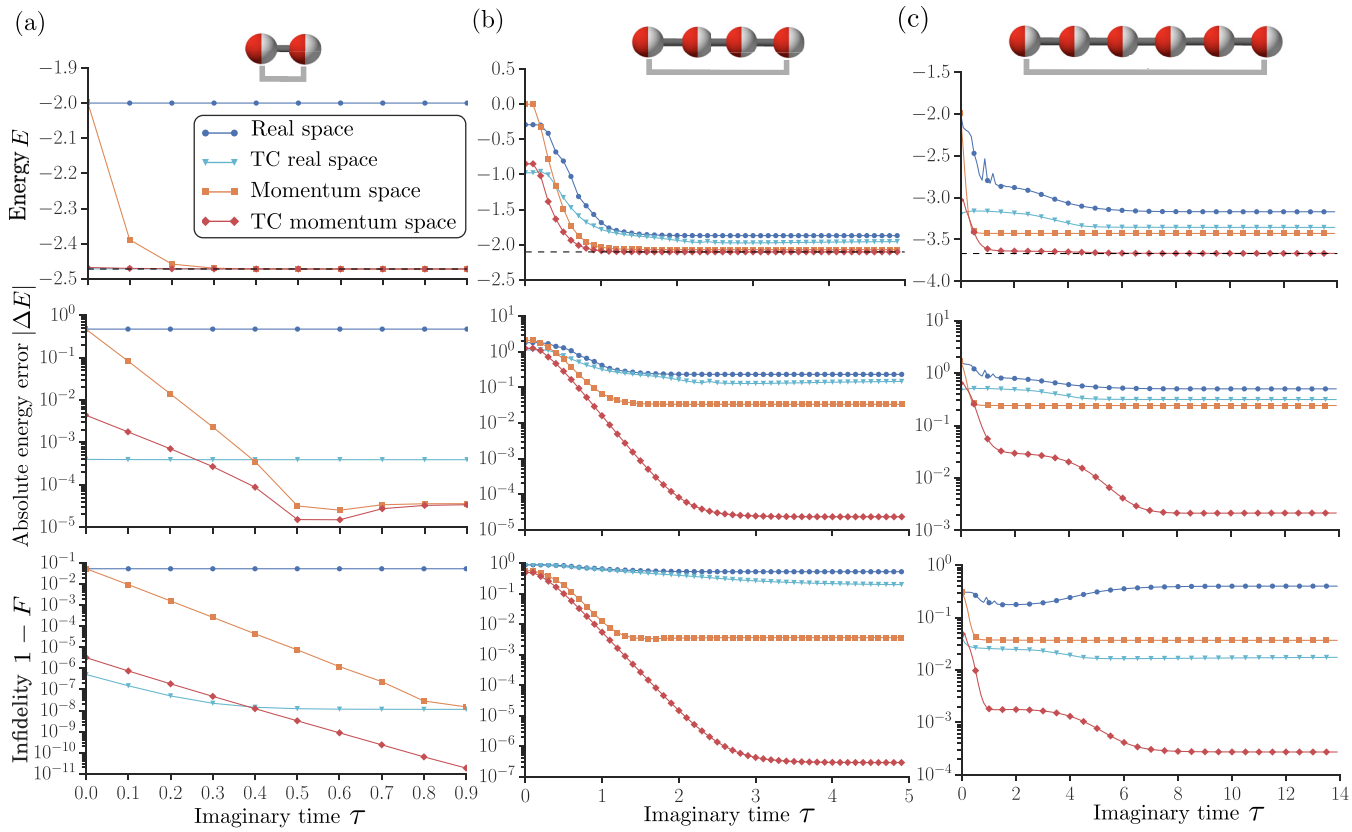

FIG. 4. Results of VarQITE state-vector simulations of repulsive Hubbard models with (a) two, (b) four, and (c) six sites at half-filling with $t = 1$ and $U = 4$, using the qUCCSD *Ansatz* with one layer. For every system, we report the evolution of the energy E (top), absolute energy error $|\Delta E| = |E - E_{\text{exact}}|$ (middle), and infidelity $1 - F$ (bottom) where the fidelity $F = |\langle\Phi|\Phi_{\text{exact}}\rangle|^2$ is computed with respect to the exact ground state at half-filling $|\Phi_{\text{exact}}\rangle$ of the corresponding Hamiltonian. The dashed lines represent the exact ground-state energies E_{exact} (-2.472, -2.103, -3.669, respectively) for the three systems. (For the six-site system, we only show every fifth data point for improved readability.) In most cases, the TC momentum-space method presents orders of magnitude improvement both in energy error and infidelity.

in conjunction with the VarQITE algorithm (at least for this critical intermediate interaction strength regime). Finally, the transcorrelated formulation of both the real-space and, more strikingly, momentum-space Hubbard model leads to a faster and tighter convergence of the ground-state energies and corresponding state fidelities.

1. Advantages of the momentum representation

In all systems investigated (Fig. 4), we observe a fast relaxation from the initial state towards the optimized ground states, with the exception of the real-space representations (in blue), which remain stacked at higher energy values due to the limitations of the wave-function *Ansatz* for this particular description of the problem. The reason for this behavior resides in the fact that the momentum representation allows for a more compact form of the wave function and therefore requires a shallower quantum circuit to describe the ground-state wave function. In fact, in the real space, the qUCCSD *Ansatz* is not expressive enough to span the portion of the Hilbert space that contains the ground-state wave function. This behavior is confirmed by VQE simulations, which reproduce equivalent results [see Fig. 5(a)].

The advantage of the momentum-space representation in the low-to-intermediate interaction strength regime is demonstrated by the fact that the final energy errors (middle row in Fig. 4), both with and without the TC method, are lower than the corresponding real-space results for all lattice sizes. Except for the six-site case, where the infidelity of the TC real-space result is lower than the non-TC momentum-space result (lower right panel of Fig. 4), all the state infidelities (bottom row) of the approximate ground state are lower in the momentum-space than in the real-space formulation. This exception shows that a larger infidelity of the ground state does not directly correspond to a large error in energy, as the momentum-space energy error in the six-site case is still lower than the corresponding TC real-space result. Additionally, in all momentum-space simulations, with and without applying transcorrelation, the same qUCCSD *Ansatz* used in the real space is capable of representing the ground state by improving on the initial state.

2. Advantages of the transcorrelated formulation

As discussed in Sec. II, the use of the transcorrelated transformation can further simplify the structure of the many-electron wave function, making the mapping to a quantum circuit more efficient. As a consequence, with the TC Hamiltonian, the optimization converges to the approximate ground states with less resources (shallower circuits) than using the

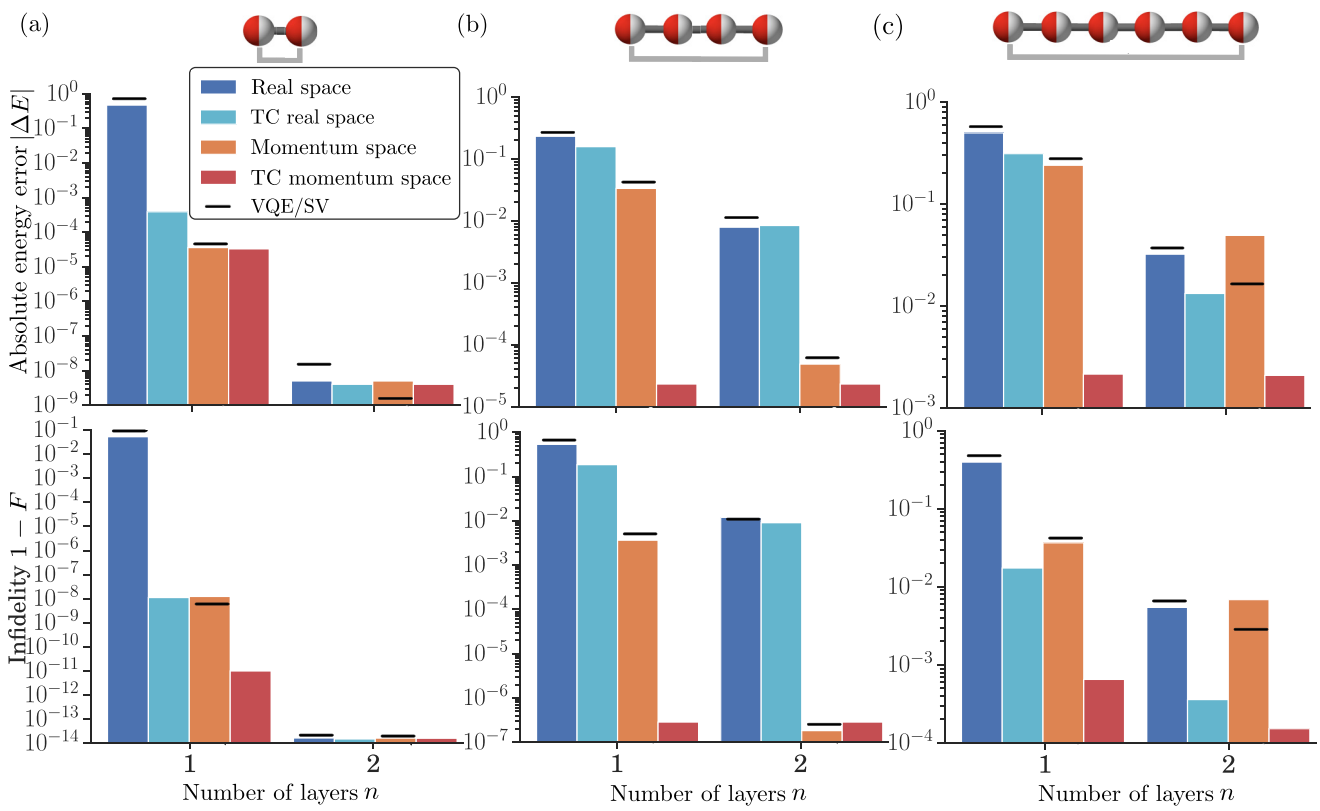

FIG. 5. Results of VarQITE state-vector simulations of repulsive Hubbard models with (a) two, (b) four, and (c) six sites at half-filling with $t = 1$ and $U = 4$, using the qUCCSD *Ansatz* with n layers. For every system, we report the absolute energy error $|\Delta E| = |E - E_{\text{exact}}|$ (top) and the infidelity $1 - F$ (bottom) at convergence, where the fidelity $F = |\langle\Phi|\Phi_{\text{exact}}\rangle|^2$ is computed with respect to the exact ground state at half-filling $|\Phi_{\text{exact}}\rangle$ of the corresponding Hamiltonian. The solid lines mark the converged outcomes of VQE state-vector simulations (VQE/SV) for the real and momentum-space Hamiltonians. The benefits of the transcorrelated approach can be seen, for instance, in the result of the six-site TC momentum-space Hamiltonian with $n = 1$, which is significantly more accurate than the plain momentum-space outcome with $n = 2$, especially in terms of fidelity.

usual (non-TC) approach. Figures 4 and 5 summarize all results for the optimization of the Hubbard models with two, four, and six sites. The optimization dynamics in Fig. 4 are given for a fixed depth ($n = 1$) qUCCSD wave-function *Ansatz*, while Fig. 5 shows the converged values for the energy deviations and state infidelities as a function of the circuit depth ($n = 1, 2$), for the different Hamiltonian representations. In all cases, a maximum of two repetition layers was sufficient to achieve a tight convergence (i.e., high fidelities) at least for the TC cases.

The efficiency of the TC approach is a consequence of the extremely "compact" form of the exact right eigenvector (i.e., almost single reference) of the TC Hamiltonians, which are dominated by the ground states with the interaction term $U = 0$, $|\Phi_0^{r/m}\rangle$, that were used as initial state in all VarQITE calculations (TC and non-TC). This effect is most pronounced for two sites [see Fig. 4(a)], where for the TC Hamiltonians both in the real and momentum space, the starting states $|\Phi_0^{r/m}\rangle$ are already a reasonable approximation of the exact ground state characterized by energy deviations $|\Delta E| < 10^{-3}$ and infidelities $I < 10^{-5}$. Compared to the original real-space results, the energy error for the TC real-space case is reduced by three orders of magnitude from about 4×10^{-1} to 3×10^{-4} and, as in the purely real-space formulation, the value of the energy or infidelity is not improved over the whole duration of the VarQITE dynamics.

Also, for the four-site model [see Fig. 4(b)] the TC formulation of the Hamiltonian in the momentum space is the best method. The initial states $|\Phi_0^{r/m}\rangle$ provide an improved starting point upon non-TC methods (see initial energy values) and we observe significant improvements of the energies and infidelities due to VarQITE for all Hamiltonians. Most importantly, the Hamiltonian in the TC momentum-space representation offers approximately up to 4 and 7 orders of magnitude improvement in absolute energy error and infidelity, respectively, in comparison to the TC real-space representation.

In Fig. 4(c), we show the results for the largest six-site system we studied in this work, presenting a challenge for VarQITE. All simulations, except the TC momentum-space ones, have significant residual energy errors $|\Delta E| > 10^{-1}$ when using the standard qUCCSD *Ansatz*. The presence of kinks at the beginning of the real-space simulation (blue circles) is due to the errors in the inversion of the linear system [Eq. (12)] and are suppressed by means of the Tikhonov regularization at future time steps. The TC momentum-space approach allows for at least 2 orders of magnitude improvement in energy error and infidelity with respect to all other approaches. Due to the inclusion of correlation directly into the TC

Hamiltonian, it is the only approach which allows to resolve the exact ground state with a limited (in terms of expressibility) one-layer qUCCSD *Ansatz*. A similar behavior was found in Dobrautz *et al.* [75], where it was shown that a momentum-space TC Hubbard model can be accurately solved with a limited restricted configuration interaction approach that only includes up to quadruple excitation for an 18-site system.

The results in Fig. 4 confirm that there exists a clear advantage in the use the TC momentum-space formulation of the Hubbard model, while the TC real-space approach presents only minor improvements in comparison. The momentum-space TC results suggest that a less expressive *Ansatz*, and thus a shallower quantum circuit, is required to obtain accurate results for the Hubbard model. For this reason, in Fig. 5 we report, as anticipated above, the converged results of VarQITE simulations when we double the number of layers in the qUCCSD *Ansatz* by repeating it with independent variational parameters (doubling the number of parameters, denoted as 2-qUCCSD), and compare the results to the ones obtained with a single qUCCSD layer. In addition, to validate our results and highlight the potential advantage of the proposed TC approach combined with the VarQITE algorithm, we also perform VQE optimizations with the original real- and momentum-space Hubbard Hamiltonians (since the standard VQE is applicable only to Hermitian case).

As for the two-site system [Fig. 5(a)], inclusion of a second layer in the qUCCSD *Ansatz* improves the circuit expressibility leading to a drastic improvement of the real-space results (dark blue), where the energy error drops from 4×10^{-1} for one layer to below 10^{-4} for two layers. Similarly, the energy error of the TC real space (light blue) and the TC momentum space (red) is reduced by more than 2 orders of magnitude upon inclusion of a second qUCCSD layer and all approaches achieve a staggeringly small infidelity of 10^{-14}. The original momentum-space results (orange) do not improve upon adding an extra quantum circuit layer. However, this is consistent with the benchmark VQE calculations (black lines in Fig. 5) that we performed for the real- and momentum-space Hubbard models.

The same trend is also confirmed for larger systems [see Figs. 5(b) and 5(c)]. In fact, for all Hamiltonian representations with the exception of the TC momentum-space one, we observe a reduction of absolute energy error (about 1–3 orders of magnitude) and of the infidelities (about 2–4 orders of magnitude) when a second layer is added to the wave-function *Ansatz*. Note that the quality of the momentum-space TC results (in red) with a single qUCCSD layer is significantly better ($|\Delta E| \lesssim 10^{-3}$) than the one obtained with all other approaches, even when in these cases two layers of the qUCCSD *Ansatz* are used. The same is true when we compared the converged momentum-space TC values with the results obtained with VQE using the real- and momentum-space (non-TC) Hamiltonians (black lines). This is a very important result in view of future applications of this approach in near-term quantum computing.

To summarize, the inclusion of correlation directly into the TC Hamiltonian via the similarity transformation based on a Gutzwiller *Ansatz* allows us to obtain highly accurate results using the VarQITE algorithm with very shallow circuits (one-layer qUCCSD), in particular when the momentum-space representation of the Hubbard model is used.

B. Hardware calculations

Due to the limited number of available qubits, we only performed hardware (HW) experiments for the two-site Hubbard model. Consequently, based on the results of Fig. 5(a), we opted for the study of the real-space Hubbard model since it shows a noticeable effect on the accuracy (see blue and teal bars) when the TC method is applied for a single UCCSD *Ansatz* layer. On the other hand, the momentum-space results for the two-site model show similar accuracy for both TC and non-TC approaches. Before moving to HW experiments, we performed QASM simulations (see Ref. [16] for some examples) of the VarQITE algorithm, which include statistical measurement noise as well as a noise model tuned for the particular IBM quantum computer used in this paper, namely, `ibmq_lima` [see Fig. 1(a)]. Details about the quantum device and the noise model used are given in Appendix G.

As mentioned above, due to its large circuit depth the qUCCSD *Ansatz* is not usable for current HW experiments, due to limited coherence time and gate errors. Thus, in both QASM and HW simulations, we used the hardware-efficient heuristic RY ansatz [130,131] applied to the Fermi-sea/Hartree-Fock initial state $|\Phi_0^r\rangle$. For more details on the nature of the wave-function *Ansatz*, see Appendix D. Furthermore, these new calculations confirm that the proposed TC approach is not restricted to a specific wave-function *Ansatz*. (On the other hand, this is also the reason why the initial state and the corresponding initial energies are different compared to the ones seen in the SV simulations.) The quantum circuit used for the HW calculations is given in Fig. 1(c).

Figure 6 shows the evolution of the total energies and absolute energy errors for the QASM simulations (pink and magenta) and the HW experiments (light and dark blue) as a function of imaginary time. Both the ordinary and TC VarQITE/QASM results are in qualitative agreement with SV results reported in Fig. 4(a), even though, as expected, the accuracy is reduced by the presence of a realistic noise model. Surprisingly, the HW calculations converge faster than the corresponding QASM noisy simulations, demonstrating that noise is not necessarily deteriorating the results. In all cases, the TC approaches converge faster than the corresponding Hermitian cases. On the other hand, achieving a tight convergence (with energy errors less than 10^{-2}) in the presence of noise is harder and, therefore, to limit the costs of the calculations, we stopped all HW experiments when no further significant improvement of the energy was noticed (hence the different duration of the simulations). The qualitative good match between the converged solutions of QASM and HW experiments demonstrate that the TC approach is not only compatible with a simpler wave-function *Ansatz* but also leads to a noise resilient implementation of the VarQITE algorithm.

V. CONCLUSIONS AND OUTLOOK

In this paper, we demonstrated the advantages of using the transcorrelated (TC) formulation of the Hubbard Hamiltonian both in real- and momentum-space representations. One

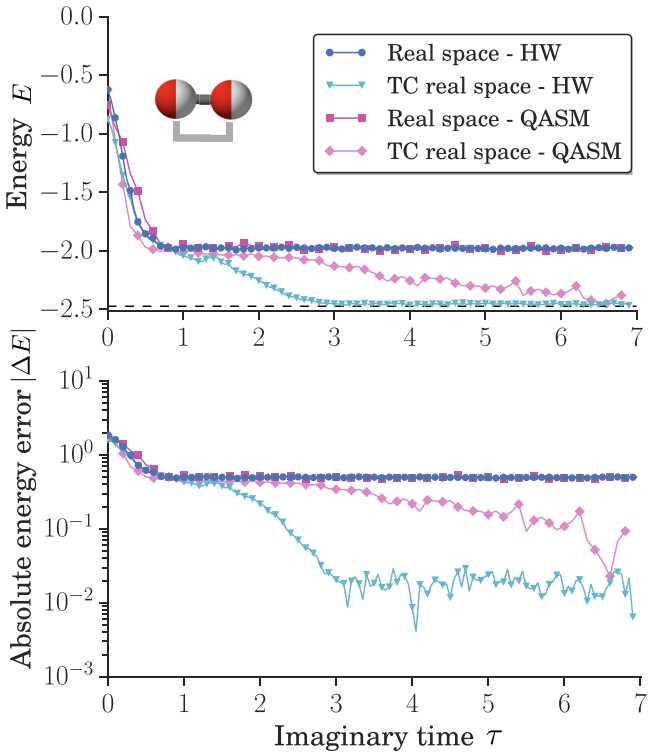

FIG. 6. Results of VarQITE simulations of the repulsive two-site Hubbard model at half-filling with $U/t = 4$. The RY heuristic *Ansatz* (applied to the Hartree-Fock state), consisting of one layer with linear entanglement, is used [see Fig. 1(c)]. We report the evolution of the energy E (top) and absolute energy error $|\Delta E| = |E - E_{\text{exact}}|$ (bottom) computed with respect to the exact ground state at half-filling $|\Phi_{\text{exact}}\rangle$ of the corresponding Hamiltonian. The dashed line represents the exact ground-state energy $E_{\text{exact}} = -2.472$. HW denotes the simulations performed on the `ibmq_lima` quantum computer. QASM marks the classical noisy simulations with the noise model of the `ibmq_lima` chip. (Only every second data point is shown to improve the readability.)

of the strengths of our approach resides in the absence of approximations in the derivation and the implementation of the TC Hamiltonian together with the efficient classical optimization of the Gutzwiller factor. The difficulties posed by the non-Hermiticity of TC Hamiltonians are overcome by using the variational quantum imaginary-time evolution (VarQITE) algorithm. In particular, we performed state-vector VarQITE simulations (without statistical and hardware noise) of two-, four-, and six-site Hubbard models, showing that the TC method in the momentum space offers up to 4 orders of magnitude improvement of the absolute energy error for a fixed *Ansatz* in comparison to the non-TC approaches.

To demonstrate the validity of our approach on a quantum computer, we propose a hardware-efficient implementation of the VarQITE algorithm in quantum circuits tailored to non-Hermitian Hamiltonians. We showed that VarQITE in the non-Hermitian case can be performed using the standard approach, differentiation of general gates via a linear combination of unitaries [124], by separating the Hamiltonian into Hermitian and non-Hermitian parts. This is in contrast with the suggestion of McArdle *et al.* [78] where different quantum circuits for each term in the Hamiltonian (i.e., controlled Hamiltonian terms) are required. Our implementation is tested by performing realistic quantum circuit (QASM) simulations for two-site Hubbard model, including the statistical error and noise sources modeled after the `ibmq_lima` quantum computer. Moreover, we further confirm our methodology by performing the same experiments on the actual `ibmq_lima` chip. The converged results are in qualitative agreement with QASM and SV simulations.

Concerning the scaling of the TC methods of this work, the presence of three-body interactions in TC Hamiltonians increases the number of required measurements on quantum hardware from a $O(N^4)$ scaling for non-TC Hamiltonians to a $O(N^6)$ scaling, with N being the total number of sites. This poses a potential challenge for applying such TC methods in near-term noisy quantum computers. However, Pauli grouping [132–135] and positive operator-valued measure [136] methods could optimize the measurement process and further reduce the number of measurements. More detailed investigations are required in order to better assess the validity of these approaches, which will be addressed in future works.

In conclusion, despite the increased number of measurements due to the presence of three-body terms, the TC approaches studied in this work significantly improve the accuracy of our calculations by making the ground state extremely "compact" and enable the use of shallower quantum circuits as wave-function *Ansätzes*, compatible with near-term noisy quantum computers. The established methodology paves the way for applications to *ab initio* Hamiltonians [71,72,137], bringing closer the first relevant demonstration of quantum advantage in a relevant use case in the field of quantum chemistry.

ACKNOWLEDGMENTS

I.O.S. and I.T. gratefully acknowledge the financial support from the Swiss National Science Foundation (SNF) through the Grant No. 200021-179312. W.D., H.L., and A.A. gratefully acknowledge financial support from the Max Planck society. W.D. acknowledges funding from the Horizon Europe research and innovation program of the European Union under the Marie Skłodowska-Curie Grant Agreement No. 101062864. We thank Jürg Hutter, Julien Gacon, Christa Zoufal, Max Rossmanek, Guglielmo Mazzola, and Daniel Miller for fruitful discussions.

APPENDIX A: QUANTUM GATES

The unitary matrix representation of the most general single-qubit gate, which allows us to obtain any quantum state on the Bloch sphere, can be written as

$$U(\theta,\phi,\lambda) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -e^{i\lambda}\sin\left(\frac{\theta}{2}\right) \\ e^{i\phi}\sin\left(\frac{\theta}{2}\right) & e^{i(\phi+\lambda)}\cos\left(\frac{\theta}{2}\right) \end{pmatrix}. \tag{A1}$$

Frequently, the gates that perform the rotations around the x, y, and z axes on the Bloch sphere are particularly useful in

heuristic *Ansätzes* [see Fig. 1(c)] and are given by

$$R_x(\theta) = U\left(\theta, -\frac{\pi}{2}, \frac{\pi}{2}\right),$$
$$R_y(\theta) = U(\theta, 0, 0),$$
$$R_z(\lambda) = e^{-i\lambda/2}U(0, 0, \lambda). \tag{A2}$$

The X gate allows us to construct the initial state $|\Phi_0\rangle = |0101\rangle$ in Fig. 1(c), and is given by $U(\pi, 0, \pi)$.

APPENDIX B: TIKHONOV REGULARIZATION

In this work, we combined the Tikhonov regularization approach with the implementation of the VarQITE algorithm, as suggested in Ref. [25]. The solution of the linear system $\boldsymbol{A}\dot{\boldsymbol{\theta}} = -\boldsymbol{C}$ at each time step of VarQITE requires the inversion of the matrix $\boldsymbol{A}$ [see Eq. (15)] which is prone to be ill conditioned. In addition, problems can occur due to the presence of hardware noise and statistical error originated from a finite number of measurements in the computation of the A_{ij} matrix elements [see Eq. (13)]. Instead, we use the aforementioned regularization to update the parameters $\boldsymbol{\theta}$, which minimizes

$$\|\boldsymbol{C} + \boldsymbol{A}\dot{\boldsymbol{\theta}}\|^2 + \lambda\|\dot{\boldsymbol{\theta}}\|^2. \tag{B1}$$

The Tikhonov parameter $\lambda \in \mathbb{R}$ can be tuned to provide a smoother evolution of the parameters $\boldsymbol{\theta}$ (i.e., when λ is large) in detriment of the accuracy (i.e., when λ is small). The optimal regularization parameter λ_{opt} can be efficiently found at each time step by finding the "corner" of an L curve in certain interval for λ [138]. For all simulations and experiments, we use $\lambda \in [10^{-3}, 1]$ and the termination threshold of the L-curve corner search set to 10^{-8}. In our experience, those parameters provided the best results.

APPENDIX C: OPTIMIZATION OF J

As mentioned in Sec. III, the optimization of the Gutzwiller parameter J based on a projection method is similar to the solution of a coupled-cluster amplitudes equation [139]. We start from a general single determinant eigenvalue equation

$$\hat{H}_{\text{TC}}(J)\,|\Phi_0\rangle = E\,|\Phi_0\rangle\,, \tag{C1}$$

where the explicit dependence of $\hat{H}_{\text{TC}}$ on the parameter J is indicated and $|\Phi_0\rangle$ denotes the HF/Fermi-sea determinant. If we project Eq. (C1) onto $\langle\Phi_0|$,

$$\langle\Phi_0|\hat{H}_{\text{TC}}|\Phi_0\rangle = E_0(J), \tag{C2}$$

we obtain an expression of the TC "Hartree-Fock" energy, which depends on the parameter J. Projecting Eq. (C1) onto $\langle\Phi_0|\,\hat{g}(J)$ yields

$$\langle\Phi_0|\hat{g}\hat{H}_{\text{TC}}(J)|\Phi_0\rangle = \langle\Phi_0|\hat{g}E|\Phi_0\rangle$$
$$= E_0(J)\langle\Phi_0|\hat{g}|\Phi_0\rangle. \tag{C3}$$

Then, combining Eqs. (C2) and (C3) yields

$$\langle\Phi_0|(\hat{g} - \langle\hat{g}\rangle_0)\hat{H}_{\text{TC}}(J)|\Phi_0\rangle = 0, \tag{C4}$$

TABLE I. Optimized Gutzwiller parameters obtained by projection, J_{proj}, and VMC optimization, J_{VMC}, for the two-, four-, and six-site Hubbard model at half-filling and $U/t = 4$. J_{VarQITE} denotes the Gutzwiller parameters that are used in our VarQITE simulations and experiments.

Number of sites	2	4	6
J_{proj}	−0.48	−0.88	−0.67
J_{VMC}	−0.47	−1.00	−0.76
J_{VarQITE}	−0.48	−0.73	−0.59

where $\langle\hat{g}\rangle_0 = \langle\Phi_0|\hat{g}|\Phi_0\rangle$ and $\hat{g}$ is expressed in the momentum space

$$\hat{g} = \frac{J}{N}\sum_{\mathbf{p},\mathbf{q},\mathbf{k},\sigma} c^\dagger_{\mathbf{p}-\mathbf{k},\sigma}c^\dagger_{\mathbf{q}+\mathbf{k},\bar{\sigma}}c_{\mathbf{q},\bar{\sigma}}c_{\mathbf{p},\sigma}. \tag{C5}$$

Equation (C4) can be efficiently solved for J with a mean-field level computational cost. To see the connection to coupled-cluster theory: Eq. (C4) can also be interpreted as a projection of the eigenvalue equation $(\hat{H}_{\text{TC}} - E)\,|\Phi_0\rangle = 0$ on the single basis of the correlation factor $\hat{g}$; the parameter J can be interpreted as the sole and uniform amplitude of a coupled-cluster *Ansatz* [Eq. (C5)]. The specific values obtained by solving Eq. (C4), J_{proj}, and VMC optimized results, J_{VMC}, for the lattice sizes, fillings, and U/t values used in this study are listed in Table I. As already studied in Ref. [75], too large values of J can cause instabilities in the imaginary-time evolution due to a resulting wide span in magnitude of the off-diagonal matrix elements after the similarity transformation. For this reason we chose slightly smaller values than J_{proj} would suggest (see J_{VarQITE} values in Table I). This still causes a more "compact" right eigenvector, while having a positive influence on the stability of the VarQITE algorithm.

APPENDIX D: QUCCSD AND HEURISTIC RY *ANSÄTZES*

The qUCCSD wave-function *Ansatz* [21]

$$|\Phi(\boldsymbol{\theta})\rangle = \hat{U}^{(n)}_{\text{ucc}}(\boldsymbol{\theta})|\Phi_0\rangle \tag{D1}$$

has seen great success in providing accurate results when embedded in variational quantum algorithms, such as VQE and VarQITE, to prepare the ground state of molecular [130,140–142] and condensed-matter Hamiltonians [143–145]. The corresponding cluster operator is given by a Trotterized version of the UCCSD operator [146]

$$\hat{U}^{(n)}_{\text{ucc}}(\boldsymbol{\theta}) = \prod_n \left(\prod_{ij}\exp\left(\theta^n_{ij}(\hat{a}^\dagger_i\hat{a}_j - \hat{a}^\dagger_j\hat{a}_i)\right)\right.$$
$$\left.\times\prod_{ijkl}\exp\left(\theta^n_{ijkl}(\hat{a}^\dagger_i\hat{a}^\dagger_j\hat{a}_k\hat{a}_l - \hat{a}^\dagger_l\hat{a}^\dagger_k\hat{a}_j\hat{a}_i)\right)\right), \tag{D2}$$

where, for this work, we also consider n-independent layers (referred to as n-qUCCSD) and all singles and doubles excitations from the Fermi-sea/Hartree-Fock state. By mapping the fermionic operators to qubits using the Jordan-Wigner transformation, corresponding quantum circuits [130] can be

derived with the number of parameters scaling as $O(nN_q^2N_e^2)$ and number of gates scaling as $O(nN_q^3N_e^2)$ where N_e and N_q are, respectively, the numbers of electrons and qubits [140]. Despite a significant number of gates for current noisy quantum processors, the attractive quality of the qUCCSD *Ansatz* resides in the guarantee of a reasonable approximation to the solution of quantum many-body problems.

To perform the experiments on quantum computers, we also employ a variant of the heuristic hardware-efficient wavefunction *Ansätzes*, which were introduced in [131], that can be written as

$$|\Phi(\boldsymbol{\theta})\rangle = \hat{U}_{\text{heu}}^{(n)}(\boldsymbol{\theta})|\Phi_0\rangle = \prod_n (\hat{U}_{\text{ent}}\hat{U}_{\text{rot}}(\boldsymbol{\theta}))|\Phi_0\rangle, \tag{D3}$$

where multiple layers are combined, consisting of alternating blocks of arbitrary parametrized single-qubit rotations $\hat{U}_{\text{rot}}(\boldsymbol{\theta})$ and entangling blocks $\hat{U}_{\text{ent}}$ composed of arbitrary arrangement of two-qubit gates. Note that the vector $\boldsymbol{\theta}$ includes the parameters from all n layers. The selection of single-qubit gates and entangling gates is typically performed to make the quantum circuit shallow enough to fit in the limits of the corresponding quantum computer. An example of such *Ansatz* is given in Fig. 1(c), where we employ two rotation layers $\hat{U}_{\text{rot}}(\boldsymbol{\theta}) = \bigotimes_{i=0}^{N_q-1} R_y(\theta_i)$ composed of $R_y(\theta_i)$ rotations on each ith qubit, separated by an entangling layer, $\hat{U}_{\text{ent}} = \prod_{i=0}^{N_q-2} \text{CNOT}(i, i+1)$, where the first parameter denotes the control qubit and the second denotes the target qubit. The initial state is a single reference state $|0101\rangle$ constructed by applying $U(\pi, 0, \pi) = X$, the Pauli-X gate (see Appendix A).

In general, for heuristic *Ansätzes* it is unclear what number of layers n is required to achieve highly accurate results. Progresses to alleviate this problem have been made with additions to variational quantum frameworks of adaptive [147–149] and evolutionary methods [150,151] for the ground-state preparation, to name only a few. Moreover, Hamiltonian-inspired *Ansätzes* were shown to provide benefits in comparison to the qUCCSD *Ansatz* for Hubbard models in terms of reduced number of variational parameters but also requiring the tuning of the number of layers as for heuristic *Ansätzes* [78,129]. In the context of transcorrelated Hamiltonians, an *Ansatz* based on the repeated layers of a Trotterized decomposition of the time-evolution operator $\exp[-i\hat{H}t]$ has been used, where a variational parameter θ is associated to every Hamiltonian term [78]. However, this *Ansatz* is inappropriate for current noisy quantum computers in comparison to the qUCCSD *Ansatz* due to the worse scaling of the number of variational parameters in $O(N^6)$ (i.e., the number of terms in TC Hamiltonians, see Sec. II B). Therefore, in this work, we focus on the most simple and validated approaches (e.g., the qUCCSD and hardware-efficient *Ansätzes*) that allow us to showcase the benefits of exact TC methods. The specific number of parameters of the different *Ansätzes* used in this work are shown in Table II. In this work, we included the parametrized global phase explicitly by means of the unitary

$$U_{\text{global}}(\theta) = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix},$$

TABLE II. Numbers of variational parameters in the wavefunction *Ansätzes* used for different N-site Hubbard models. An additional variational parameter is used to track the global phase in VarQITE simulations.

Number of sites	2	4	6
qUCCSD	3	26	117
2-qUCCSD	6	52	234
RY	8		

added to the first qubit at the end of every circuit *Ansatz*. The $U_{\text{global}}(\theta)$ gate can be decomposed as

$$\begin{aligned} U_{\text{global}}(\theta) &= U_1(\theta)XU_1(\theta)X \\ &= \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ e^{i\theta} & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ e^{i\theta} & 0 \end{pmatrix} = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix}, \end{aligned}$$

where the gates X and U_1 are native on the IBM Quantum devices.

APPENDIX E: COMPUTATION OF C-VECTOR ELEMENTS

We first define the Hermitian $\hat{H}_{\text{TC}}^{+}$ and the anti-Hermitian $\hat{H}_{\text{TC}}^{-}$ operators derived from the non-Hermitian TC Hamiltonian $\hat{H}_{\text{TC}} \in \{\hat{H}_{\text{tc}}^{r}, \hat{H}_{\text{tc}}^{m}\}$:

$$\hat{H}_{\text{TC}}^{+} = \hat{H}_{\text{TC}} + \hat{H}_{\text{TC}}^{\dagger}, \tag{E1}$$

$$\hat{H}_{\text{TC}}^{-} = \hat{H}_{\text{TC}} - \hat{H}_{\text{TC}}^{\dagger}, \tag{E2}$$

with $\hat{H}_{\text{TC}}^{+\dagger} = \hat{H}_{\text{TC}}^{+}$ and $\hat{H}_{\text{TC}}^{-\dagger} = -\hat{H}_{\text{tc}}^{-}$.

Then, we can write

$$\begin{aligned} &\langle\partial_\theta\Phi|\hat{H}_{\text{TC}}^{+}|\Phi\rangle + \langle\Phi|\hat{H}_{\text{TC}}^{+\dagger}|\partial_\theta\Phi\rangle \\ &\quad = \langle\partial_\theta\Phi|\hat{H}_{\text{TC}}^{+}|\Phi\rangle + (\langle\partial_\theta\Phi|\hat{H}_{\text{TC}}^{+}|\Phi\rangle)^* \\ &\quad = 2\,\text{Re}\langle\partial_\theta\Phi|\hat{H}_{\text{TC}}^{+}|\Phi\rangle \end{aligned} \tag{E3}$$

and

$$\begin{aligned} &\langle\partial_\theta\Phi|\hat{H}_{\text{TC}}^{-}|\Phi\rangle + \langle\Phi|\hat{H}_{\text{TC}}^{-\dagger}|\partial_\theta\Phi\rangle \\ &\quad = \langle\partial_\theta\Phi|\hat{H}_{\text{TC}}^{-}|\Phi\rangle + (\langle\partial_\theta\Phi|\hat{H}_{\text{TC}}^{-}|\Phi\rangle)^* \\ &\quad = 2\,\text{Re}\langle\partial_\theta\Phi|\hat{H}_{\text{TC}}^{-}|\Phi\rangle. \end{aligned} \tag{E4}$$

Using the (anti-)Hermiticity, we obtain

$$\langle\partial_\theta\Phi|\hat{H}_{\text{TC}}^{+}|\phi\rangle + \langle\Phi|\hat{H}_{\text{TC}}^{+}|\partial_\theta\Phi\rangle = 2\,\text{Re}\langle\partial_\theta\Phi|\hat{H}_{\text{TC}}^{+}|\Phi\rangle, \tag{E5}$$

$$\langle\partial_\theta\Phi|\hat{H}_{\text{TC}}^{-}|\Phi\rangle - \langle\Phi|\hat{H}_{\text{TC}}^{-}|\partial_\theta\Phi\rangle = 2\,\text{Re}\langle\partial_\theta\Phi|\hat{H}_{\text{tc}}^{-}|\Phi\rangle. \tag{E6}$$

Combining Eqs. (E5) and (E6), one gets

$$\begin{aligned} &\langle\partial_\theta\Phi|\hat{H}_{\text{TC}}^{+} + \hat{H}_{\text{TC}}^{-}|\Phi\rangle + \langle\Phi|\hat{H}_{\text{TC}}^{+} - \hat{H}_{\text{tc}}^{-}|\partial_\theta\phi\rangle \\ &\quad = 2\,\text{Re}\langle\partial_\theta\Phi|\hat{H}_{\text{TC}}^{+}|\Phi\rangle + 2\,\text{Re}\langle\partial_\theta\Phi|\hat{H}_{\text{TC}}^{-}|\Phi\rangle \end{aligned} \tag{E7}$$

which implies

$$\langle\partial_\theta\Phi|2\hat{H}_{\rm TC}|\Phi\rangle + \langle\Phi|2\hat{H}^\dagger_{\rm TC}|\partial_\theta\Phi\rangle = 2\,{\rm Re}\langle\partial_\theta\Phi|\hat{H}^+_{\rm TC}|\Phi\rangle + 2\,{\rm Re}\langle\partial_\theta\Phi|\hat{H}^-_{\rm TC}|\Phi\rangle. \tag{E8}$$

Inserting the definitions in Eqs. (17) and (18) and dividing by 2, we finally get

$$\langle\partial_\theta\Phi|\hat{H}_{\rm TC}|\Phi\rangle + \langle\Phi|\hat{H}^\dagger_{\rm TC}|\partial_\theta\Phi\rangle = {\rm Re}\langle\partial_\theta\Phi|\hat{H}^+_{\rm TC}|\Phi\rangle + {\rm Re}\langle\partial_\theta\Phi|\hat{H}^-_{\rm TC}|\Phi\rangle, \tag{E9}$$

which proves Eq. (19).

APPENDIX F: QUANTUM CIRCUIT FOR THE C-VECTOR ELEMENTS

As mentioned in Sec. III, we derive the quantum circuits that are proposed to measure the elements of the gradient vector $\boldsymbol{C}$ in the VarQITE algorithm (see Fig. 2). In particular, we explain how to construct the quantum circuits that are compatible with the measurements of the (anti-)Hermitian terms contained in $(\hat{H}^-_{\rm TC})\,\hat{H}^+_{\rm TC}$. To this end, we follow closely the derivations made for the Hermitian case in Ref. [124] using the linear combination of unitaries approach. Consider the case of the heuristic RY *Ansatz* as used in our QASM and hardware experiments. To compute a derivative with respect to some gate parameter θ_i of the RY *Ansatz*, we make use of a single ancilla qubit. We show how the quantum circuit for the anti-Hermitian case [see Fig. 2 with $V = R_x(\frac{\pi}{2})$] can be derived. First, we write the initial state of our quantum register as

$$|0\rangle_a \otimes |0\rangle. \tag{F1}$$

By applying a Hadamard gate on the ancilla qubit, with its unitary matrix

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}, \tag{F2}$$

we obtain the state

$$\frac{1}{\sqrt{2}}(|0\rangle_a + |1\rangle_a) \otimes |0\rangle. \tag{F3}$$

We add a part of the *Ansatz* circuit that comes before the differentiated gate, obtaining the state

$$\frac{1}{\sqrt{2}}(|0\rangle_a + |1\rangle_a) \otimes U(\theta_{0:i-1})|0\rangle. \tag{F4}$$

According to Schuld *et al.* [124], a derivative of the gate $\bar{G}$ can be decomposed into a linear combination of unitary gates Q_1 and Q_2 as

$$\partial_{\theta_i}\bar{G} = \frac{\alpha}{2}[(Q_1 + Q_1^\dagger) + i(Q_2 + Q_2^\dagger)], \tag{F5}$$

with a parameter $\alpha \in \mathbb{R}$. For instance, for a gate $R_y(\theta_i)$ of the RY *Ansatz*, as in Fig. 2, $\partial_{\theta_i}\bar{G} = \beta R_y(\theta_i)C_Y$ with $\beta = -\frac{1}{2}i$. C_Y is a controlled-Y gate where Y stands for the Pauli operator $\hat{\sigma}_y$. This decomposition is performed automatically in QISKIT. The state of the circuit then becomes

$$\frac{1}{\sqrt{2}}[|0\rangle_a \otimes U(\theta_{0:i-1})|0\rangle + |1\rangle_a \otimes \beta Y U(\theta_{0:i-1})|0\rangle]. \tag{F6}$$

Then, we add the $R_y(\theta_i)$ and $R_y(\theta_{i+1})$ gates to the circuit, as in Fig. 2:

$$\frac{1}{\sqrt{2}}[|0\rangle_a \otimes R_y(\theta_{i+1})R_y(\theta_i)U(\theta_{0:i-1})|0\rangle + |1\rangle_a \otimes R_y(\theta_{i+1})R_y(\theta_i)\beta Y U(\theta_{0:i-1})|0\rangle]. \tag{F7}$$

For compactness, we rewrite $\partial_{\theta_i}G = \beta R_y(\theta_i)Y$ and $U(\theta_{0:i}) = R_y(\theta_{i+1})R_y(\theta_i)U(\theta_{0:i-1})$ and obtain

$$\frac{1}{\sqrt{2}}[|0\rangle_a \otimes U(\theta_{0:i})|0\rangle + |1\rangle_a \otimes R_y(\theta_{i+1})\partial_{\theta_i}GU(\theta_{0:i-1})|0\rangle]. \tag{F8}$$

Next, we apply the $R_x(\frac{\pi}{2})$ gate, with its unitary matrix

$$R_x\left(\frac{\pi}{2}\right) = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -i\\ -i & 1\end{pmatrix}, \tag{F9}$$

on the ancilla qubit, obtaining

$$\tfrac{1}{2}[(|0\rangle_a - i|1\rangle_a) \otimes U(\theta_{0:i})|0\rangle + (-i|0\rangle_a + |1\rangle_a) \otimes R_y(\theta_{i+1})\partial_{\theta_i}GU(\theta_{0:i-1})|0\rangle], \tag{F10}$$

which can be written as

$$\tfrac{1}{2}[|0\rangle_a \otimes [U(\theta_{0:i}) - iR_y(\theta_{i+1})\partial_{\theta_i}GU(\theta_{0:i-1})]|0\rangle + |1\rangle_a \otimes [-iU(\theta_{0:i}) + R_y(\theta_{i+1})\partial_{\theta_i}GU(\theta_{0:i-1})]|0\rangle]. \tag{F11}$$

Therefore, if the ancilla qubit is measured in the state $|0\rangle_a$, we obtain the state

$$|\Psi_0\rangle = \frac{1}{2\sqrt{p_0}}\underbrace{[U(\theta_{0:i}) - iR_y(\theta_{i+1})\partial_{\theta_i}GU(\theta_{0:i-1})]}_{R_0}|0\rangle, \tag{F12}$$

and if it is measured in the state $|1\rangle_a$, we obtain the state

$$|\Psi_1\rangle = \frac{1}{2\sqrt{p_1}}\underbrace{[-iU(\theta_{0:i}) + R_y(\theta_{i+1})\partial_{\theta_{i+1}}GU(\theta_{0:i-1})]}_{R_1}|0\rangle, \tag{F13}$$

where $p_i = \frac{1}{4}\langle 0|R_i^\dagger R_i|0\rangle$ for $i \in \{0, 1\}$. To obtain the C_i^- vector elements [see Eq. (21)], we can combine such measurements in the following way: If we measure our Hamiltonian and the ancilla is in the state $|0\rangle_a$, we get

$$\langle\Psi_0|\hat{H}^-_{\rm TC}|\Psi_0\rangle = \frac{1}{4p_0}[\langle 0|[U(\theta_{0:i})^\dagger + iU(\theta_{0:i-1})^\dagger\partial_{\theta i}G^\dagger R_y(\theta_{i+1})^\dagger]\hat{H}^-_{tc}[U(\theta_{0:i}) - iR_y(\theta_{i+1})\partial_{\theta_i}GU(\theta_{0:i-1})]|0\rangle] \tag{F14}$$

and if the ancilla is in the state $|1\rangle_a$, we get

$$\langle\Psi_1|\hat{H}^-_{\rm tc}|\Psi_1\rangle = \frac{1}{4p_1}[\langle 0|[iU(\theta_{0:i})^\dagger + U(\theta_{0:i-1})^\dagger\partial_{\theta i}G^\dagger R_y(\theta_{i+1})^\dagger]\hat{H}^-_{tc}[-iU(\theta_{0:i}) + R_y(\theta_{i+1})\partial_{\theta i}GU(\theta_{0:i-1})]|0\rangle]. \tag{F15}$$

As we want to keep the term $\langle 0|[U(\theta_{0:i-1})^\dagger\partial_{\theta i}G^\dagger R_y(\theta_{i+1})^\dagger]\hat{H}^-_{\rm tc}U(\theta_{0:i})|0\rangle = \langle\partial_i\Psi|\hat{H}^-_{\rm tc}|\Psi\rangle$, we can subtract Eq. (F14) from Eq. (F15), including also the probabilities, to

TABLE III. Characteristics of the `ibmq_lima` quantum computer at the time of hardware and noisy QASM simulations. Data for the calibration date of 21/11/2021 obtained at 11:00:00 GMT. The notation "0_1" denotes the CNOT gate between qubits 0 (control) and 1 (target).

Qubit	T1 (μs)	T2 (μs)	Frequency (GHz)	Readout error	Pauli-X error	CNOT error
Q0	119.3	164.63	5.03	5.60×10^{-2}	3.693×10^{-4}	0_1: 7.143×10^{-3}
Q1	55.21	145.47	5.128	1.87×10^{-2}	2.001×10^{-4}	1_0: 7.143×10^{-3}; 1_3: 1.478×10^{-2}; 1_2: 5.417×10^{-3}
Q2	109.17	122.27	5.247	2.21×10^{-2}	2.520×10^{-4}	2_1: 5.417×10^{-3}
Q3	96.83	103	5.302	2.91×10^{-2}	2.978×10^{-4}	3_4: 1.528×10^{-2}; 3_1: 1.478×10^{-2}
Q4	25.19	19.48	5.092	5.01×10^{-2}	6.759×10^{-4}	4_3: 1.528×10^{-2}

obtain

$$\begin{aligned} &p_0\langle\Psi_0|\hat{H}_{\rm tc}^-|\Psi_0\rangle - p_1\langle\Psi_1|\hat{H}_{\rm tc}^-|\Psi_1\rangle \\ &\quad = \frac{i}{2}[-\langle 0|U(\theta_{0:i})^\dagger \hat{H}_{tc}^- R_y(\theta_{i+1})\partial_{\theta i} G U(\theta_{0:i-1})|0\rangle \\ &\qquad + \langle 0|U(\theta_{0:i-1})^\dagger \partial_{\theta i} G^\dagger R_y(\theta_{i+1})^\dagger \hat{H}_{tc}^- U(\theta_{0:i})|0\rangle]. \end{aligned} \tag{F16}$$

Finally, multiplying by the factor $-2i$, we obtain C_i^-:

$$\begin{aligned} &-2i[p_0\langle\Psi_0|\hat{H}_{\rm tc}^-|\Psi_0\rangle - p_1\langle\Psi_1|\hat{H}_{\rm tc}^-|\Psi_1\rangle] \\ &\quad = \langle\partial_{\theta_i}\Phi|\hat{H}_{tc}^-|\Phi\rangle - \langle\Phi|\hat{H}_{tc}^-|\partial_{\theta_i}\Phi\rangle = C_i^-. \end{aligned} \tag{F17}$$

For the Hermitian case, the computation of $C_i^+ = \langle\partial_\theta\Phi|\hat{H}_{\rm tc}^+|\Phi\rangle + \langle\Phi|\hat{H}_{\rm tc}^+|\partial_\theta\Phi\rangle$ is described in detail in Ref. [124], Sec. III B.

APPENDIX G: HARDWARE CHARACTERISTICS AND NOISE MODEL

In Table III, we provide the device characteristics at the time of our hardware experiments, subsequently used for the noise model in our QASM simulations. The necessary information (T1, T2, qubit frequencies, readout errors, error rates for single-qubit and two-qubit gates per qubit) is reported here to enable the reconstruction of our noise model using QISKIT. To build the noise model of the `ibmq_lima` quantum processor, the same procedure as in Ref. [16] is employed, which is summarized below. The error sources considered in QASM simulations (see Fig. 6) are the depolarization, thermalization, and readout errors.

The depolarization error is represented as the decay of the noiseless density matrix $\rho = |\Phi\rangle\langle\Phi|$ to the uncorrelated density matrix $\mathbf{1}/2^{N_q}$:

$$\rho_d = \gamma_1 \mathrm{Tr}[\rho]\mathbf{1}/2^{N_q} + (1-\gamma_1)\rho, \tag{G1}$$

with N_q being the number of qubits and γ_1 representing the decay rate. The latter is estimated using gate fidelities given in Table III. The thermalization error of a qubit, which consists of general amplitude dampening and phase-flip error, is defined as the decay towards the Fermi-Dirac distribution of ground and excited states based on their energy difference ω:

$$\rho_t = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|, \tag{G2}$$

with $p = (e^{\frac{-\omega}{k_b T}} + 1)^{-1}$, T being the temperature and k_B the Boltzmann constant.

The readout error is classically modeled by calibrating the so-called measurement error matrix $\mathcal{M}$. The $\mathcal{M}$ matrix assigns to any N_q-qubit computational basis state $|i\rangle$ (i.e., the correct state that should be obtained) a probability to read out all the states $|j\rangle$ (i.e., the states that are actually obtained due to noise), or concisely $\mathcal{P}(i|j)$ where i, j are N_q-qubit bit strings. In an ideal noiseless situation, this matrix $\mathcal{M}$ would be characterized by its matrix elements $\mathcal{P}(i|j) = 1$ for $i = j$ and $\mathcal{P}(i|j) = 0$ for $i \neq j$.

[1] R. P. Feynman, Simulating physics with computers, Int. J. Theor. Phys. **21**, 467 (1982).

[2] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Error mitigation extends the computational reach of a noisy quantum processor, Nature (London) **567**, 491 (2019).

[3] L. Egan, D. M. Debroy, C. Noel, A. Risinger, D. Zhu, D. Biswas, M. Newman, M. Li, K. R. Brown, M. Cetina *et al.*, Fault-tolerant control of an error-corrected qubit, Nature (London) **598**, 281 (2021).

[4] E. T. Campbell, B. M. Terhal, and C. Vuillot, Roads towards fault-tolerant universal quantum computation, Nature (London) **549**, 172 (2017).

[5] IBM, roadmap for building an open quantum software ecosystem (2021), https://www.ibm.com/quantum/roadmap.

[6] N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn *et al.*, Quantum optimization using variational algorithms on near-term quantum devices, Quantum Sci. Technol. **3**, 030503 (2018).

[7] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. Sawaya *et al.*, Quantum chemistry in the age of quantum computing, Chem. Rev. **119**, 10856 (2019).

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

[9] J. Liu, L. Wan, Z. Li, and J. Yang, Simulating periodic systems on quantum computer using molecular orbitals, J. Chem. Theory Comput. **16**, 6904 (2020).

[10] K. Yamamoto, D. Z. Manrique, I. Khan, H. Sawada, and D. M. Ramo, Quantum hardware calculations of periodic systems: Hydrogen chain and iron crystals, Phys. Rev. Res. **4**, 033110 (2022).

[11] S. McArdle, A. Mayorov, X. Shan, S. Benjamin, and X. Yuan, Digital quantum simulation of molecular vibrations, Chem. Sci. **10**, 5725 (2019).

[12] N. P. Sawaya, T. Menke, T. H. Kyaw, S. Johri, A. Aspuru-Guzik, and G. G. Guerreschi, Resource-efficient digital quantum simulation of d-level systems for photonic, vibrational, and spin-s hamiltonians, npj Quantum Inf. **6**, 49 (2020).

[13] P. J. Ollitrault, A. Baiardi, M. Reiher, and I. Tavernelli, Hardware efficient quantum algorithms for vibrational structure calculations, Chem. Sci. **11**, 6842 (2020).

[14] C. O. Rubiera, M. Strahm, J. Shi, G. M. Morris, S. C. Benjamin, and C. M. Deane, Investigating the potential for a limited quantum speedup on protein lattice problems, New J. Phys. **23**, 103030 (2021).

[15] A. Robert, P. K. Barkoutsos, S. Woerner, and I. Tavernelli, Resource-efficient quantum algorithm for protein folding, npj Quantum Inf. **7**, 38 (2021).

[16] I. O. Sokolov, P. K. Barkoutsos, L. Moeller, P. Suchsland, G. Mazzola, and I. Tavernelli, Microcanonical and finite-temperature *ab initio* molecular dynamics simulations on quantum computers, Phys. Rev. Res. **3**, 013125 (2021).

[17] D. A. Fedorov, M. J. Otten, S. K. Gray, and Y. Alexeev, *Ab initio* molecular dynamics on quantum computers, J. Chem. Phys. **154**, 164103 (2021).

[18] P. J. Ollitrault, A. Miessen, and I. Tavernelli, Molecular quantum dynamics: A quantum computing perspective, Acc. Chem. Res. **54**, 4229 (2021).

[19] S. V. Mathis, G. Mazzola, and I. Tavernelli, Toward scalable simulations of lattice gauge theories on quantum computers, Phys. Rev. D **102**, 094501 (2020).

[20] G. Mazzola, S. V. Mathis, G. Mazzola, and I. Tavernelli, Gauge invariant quantum circuits for $U(1)$ and Yang-Mills lattice gauge theories, Phys. Rev. Res. **3**, 043209 (2021).

[21] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 4213 (2014).

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

[23] B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan, Quantum algorithms for quantum chemistry and quantum materials science, Chem. Rev. **120**, 12685 (2020).

[24] M. Motta, C. Sun, A. T. K. Tan, M. J. O'Rourke, E. Ye, A. J. Minnich, F. G. S. L. Brandão, and G. K.-L. Chan, Publisher correction: Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, Nat. Phys. **16**, 231 (2020).

[25] S. McArdle, T. Jones, S. Endo, Y. Li, S. C. Benjamin, and X. Yuan, Variational ansatz-based quantum simulation of imaginary time evolution, npj Quantum Inf. **5**, 75 (2019).

[26] A. McLachlan, A variational solution of the time-dependent schrodinger equation, Mol. Phys. **8**, 39 (1964).

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

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

[29] C. Zoufal, A. Lucchi, and S. Woerner, Variational quantum boltzmann machines, Quantum Mach. Intell. **3**, 7 (2021).

[30] J. Liu and Y. Xin, Quantum simulation of quantum field theories as quantum chemistry, J. High Energy Phys. 12 (2020) 011.

[31] H.-Y. Huang, K. Bharti, and P. Rebentrost, Near-term quantum algorithms for linear systems of equations, New J. Phys. **23**, 113021 (2021).

[32] X. Xu, J. Sun, S. Endo, Y. Li, S. C. Benjamin, and X. Yuan, Variational algorithms for linear algebra, Sci. Bull. **66**, 2181 (2021).

[33] D. Amaro, C. Modica, M. Rosenkranz, M. Fiorentini, M. Benedetti, and M. Lubasch, Filtering variational quantum algorithms for combinatorial optimization, Quantum Sci. Technol. **7**, 015021 (2022).

[34] F. Fontanela, A. Jacquier, and M. Oumgari, A quantum algorithm for linear pdes arising in finance, SIAM J. Finan. Math. **12**, SC98 (2021).

[35] N. Gomes, A. Mukherjee, F. Zhang, T. Iadecola, C.-Z. Wang, K.-M. Ho, P. P. Orth, and Y.-X. Yao, Adaptive variational quantum imaginary time evolution approach for ground state preparation, Adv. Quantum Technol. **4**, 2100114 (2021).

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

[37] C. Zoufal, D. Sutter, and S. Woerner, Error bounds for variational quantum time evolution, arXiv:2108.00022.

[38] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Simulated quantum computation of molecular energies, Science **309**, 1704 (2005).

[39] R. Babbush, N. Wiebe, J. McClean, J. McClain, H. Neven, and G. K.-L. Chan, Low-Depth Quantum Simulation of Materials, Phys. Rev. X **8**, 011044 (2018).

[40] A. Y. Kitaev, Quantum computations: Algorithms and error correction, Russ. Math. Surv. **52**, 1191 (1997).

[41] A. J. C. Woitzik, P. K. Barkoutsos, F. Wudarski, A. Buchleitner, and I. Tavernelli, Entanglement production and convergence properties of the variational quantum eigensolver, Phys. Rev. A **102**, 042402 (2020).

[42] P. Jordan and E. Wigner, Über das paulische äquivalenzverbot, Z. Phys. **47**, 631 (1928).

[43] T. Kato, On the eigenfunctions of many-particle systems in quantum mechanics, Commun. Pure Appl. Math. **10**, 151 (1957).

[44] R. T. Pack and W. B. Brown, Cusp conditions for molecular wavefunctions, J. Chem. Phys. **45**, 556 (1966).

[45] E. A. Hylleraas, Neue berechnung der energie des heliums im grundzustande, sowie des tiefsten terms von ortho-helium, Z. Phys. **54**, 347 (1929).

[46] K. Szalewicz and B. Jeziorski, Explicitly-correlated gaussian geminals in electronic structure calculations, Mol. Phys. **108**, 3091 (2010).

[47] K. Szalewicz, B. Jeziorski, H. J. Monkhorst, and J. G. Zabolitzky, A new functional for variational calculation of atomic and molecular second-order correlation energies, Chem. Phys. Lett. **91**, 169 (1982).

[48] J. Mitroy, S. Bubin, W. Horiuchi, Y. Suzuki, L. Adamowicz, W. Cencek, K. Szalewicz, J. Komasa, D. Blume, and K. Varga, Theory and application of explicitly correlated gaussians, Rev. Mod. Phys. **85**, 693 (2013).

[49] W. Kutzelnigg, r_{12}-dependent terms in the wave function as closed sums of partial wave amplitudes for large l, Theor. Chim. Acta **68**, 445 (1985).

[50] W. Kutzelnigg and W. Klopper, Wave functions with terms linear in the interelectronic coordinates to take care of the correlation cusp. I. General theory, J. Chem. Phys. **94**, 1985 (1991).

[51] J. Noga and W. Kutzelnigg, Coupled cluster theory that takes care of the correlation cusp by inclusion of linear terms in the interelectronic coordinates, J. Chem. Phys. **101**, 7738 (1994).

[52] S. Ten-no, Initiation of explicitly correlated slater-type geminal theory, Chem. Phys. Lett. **398**, 56 (2004).

[53] S. Ten-no, Explicitly correlated second order perturbation theory: Introduction of a rational generator and numerical quadratures, J. Chem. Phys. **121**, 117 (2004).

[54] E. F. Valeev, Improving on the resolution of the identity in linear R12 *ab initio* theories, Chem. Phys. Lett. **395**, 190 (2004).

[55] H.-J. Werner, T. B. Adler, and F. R. Manby, General orbital invariant MP2-F12 theory, J. Chem. Phys. **126**, 164102 (2007).

[56] A. Grüneis, S. Hirata, Y. ya Ohnishi, and S. Ten-no, Perspective: Explicitly correlated electronic structure theory for complex systems, J. Chem. Phys. **146**, 080901 (2017).

[57] C. Hättig, W. Klopper, A. Köhn, and D. P. Tew, Explicitly correlated electrons in molecules, Chem. Rev. **112**, 4 (2011).

[58] L. Kong, F. A. Bischoff, and E. F. Valeev, Explicitly correlated R12/F12 methods for electronic structure, Chem. Rev. **112**, 75 (2011).

[59] S. Ten-no, Explicitly correlated wave functions: Summary and perspective, Theor. Chem. Acc. **131**, 1070 (2012).

[60] S. Ten-no and J. Noga, Explicitly correlated electronic structure theory from R12/F12 ansätze, WIREs Comput. Mol. Sci. **2**, 114 (2011).

[61] T. Takeshita, N. C. Rubin, Z. Jiang, E. Lee, R. Babbush, and J. R. McClean, Increasing the Representation Accuracy of Quantum Simulations of Chemistry without Extra Quantum Resources, Phys. Rev. X **10**, 011004 (2020).

[62] J. S. Kottmann, P. Schleich, T. Tamayo-Mendoza, and A. Aspuru-Guzik, Reducing qubit requirements while maintaining numerical precision for the variational quantum eigensolver: A basis-set-free approach, J. Phys. Chem. Lett. **12**, 663 (2021).

[63] P. Schleich, J. S. Kottmann, and A. Aspuru-Guzik, Improving the accuracy of the variational quantum eigensolver for molecular systems by the explicitly-correlated perturbative $[2]_{R12}$-correction, Phys. Chem. Chem. Phys. **24**, 13550 (2022).

[64] G. Mazzola, P. J. Ollitrault, P. K. Barkoutsos, and I. Tavernelli, Nonunitary Operations for Ground-State Calculations in Near-Term Quantum Computers, Phys. Rev. Lett. **123**, 130501 (2019).

[65] F. Benfenati, G. Mazzola, C. Capecci, P. K. Barkoutsos, P. J. Ollitrault, I. Tavernelli, and L. Guidoni, Improved accuracy on noisy devices by nonunitary variational quantum eigensolver for chemistry applications, J. Chem. Theory Comput. **17**, 3946 (2021).

[66] J. O. Hirschfelder, Removal of electron-electron poles from many-electron hamiltonians, J. Chem. Phys. **39**, 3145 (1963).

[67] S. Boys and N. Handy, A condition to remove the indeterminacy in interelectronic correlation functions, Proc. R. Soc. London A **309**, 209 (1969).

[68] S. Boys and N. Handy, The determination of energies and wavefunctions with full electronic correlation, Proc. R. Soc. London A **310**, 43 (1969).

[69] S. Boys and N. Handy, A calculation for the energies and wavefunctions for states of neon with full electronic correlation accuracy, Proc. R. Soc. London A **310**, 63 (1969).

[70] R. Jastrow, Many-body problem with strong forces, Phys. Rev. **98**, 1479 (1955).

[71] A. J. Cohen, H. Luo, K. Guther, W. Dobrautz, D. P. Tew, and A. Alavi, Similarity transformation of the electronic schrödinger equation via jastrow factorization, J. Chem. Phys. **151**, 061101 (2019).

[72] J. P. Haupt, S. M. Hosseini, P. López Ríos, W. Dobrautz, A. Cohen, and A. Alavi, Optimizing jastrow factors for the transcorrelated method, J. Chem. Phys. **158**, 224105 (2023).

[73] M. C. Gutzwiller, Effect of Correlation on the Ferromagnetism of Transition Metals, Phys. Rev. Lett. **10**, 159 (1963).

[74] W. F. Brinkman and T. M. Rice, Application of gutzwiller's variational method to the metal-insulator transition, Phys. Rev. B **2**, 4302 (1970).

[75] W. Dobrautz, H. Luo, and A. Alavi, Compact numerical solutions to the two-dimensional repulsive hubbard model obtained via nonunitary similarity transformations, Phys. Rev. B **99**, 075119 (2019).

[76] D.-B. Zhang, B.-L. Chen, Z.-H. Yuan, and T. Yin, Variational quantum eigensolvers by variance minimization, Chinese Phys. B **31**, 120301 (2022).

[77] S. Tsuneyuki, Transcorrelated method: Another possible way towards electronic structure calculation of solids, Prog. Theor. Phys. Suppl. **176**, 134 (2008).

[78] S. McArdle and D. P. Tew, Improving the accuracy of quantum computational chemistry using the transcorrelated method, arXiv:2006.11181.

[79] H. Luo and A. Alavi, Combining the transcorrelated method with full configuration interaction quantum monte carlo: Application to the homogeneous electron gas, J. Chem. Theory Comput. **14**, 1403 (2018).

[80] K. Guther, A. J. Cohen, H. Luo, and A. Alavi, Binding curve of the beryllium dimer using similarity-transformed FCIQMC: Spectroscopic accuracy with triple-zeta basis sets, J. Chem. Phys. **155**, 011102 (2021).

[81] K. Guther, R. J. Anderson, N. S. Blunt, N. A. Bogdanov, D. Cleland, N. Dattani, W. Dobrautz, K. Ghanem, P. Jeszenszki, N. Liebermann, G. L. Manni, A. Y. Lozovoi, H. Luo, D. Ma, F. Merz, C. Overy, M. Rampp, P. K. Samanta, L. R. Schwarz, J. J. Shepherd *et al.*, NECI: N-electron configuration interaction with an emphasis on state-of-the-art stochastic methods, J. Chem. Phys. **153**, 034107 (2020).

[82] G. H. Booth, A. J. W. Thom, and A. Alavi, Fermion monte carlo without fixed nodes: A game of life, death, and annihilation in slater determinant space, J. Chem. Phys. **131**, 054106 (2009).

[83] D. Cleland, G. H. Booth, and A. Alavi, Communications: Survival of the fittest: Accelerating convergence in full

configuration-interaction quantum monte carlo, J. Chem. Phys. **132**, 041103 (2010).

[84] W. Dobrautz, S. D. Smart, and A. Alavi, Efficient formulation of full configuration interaction quantum Monte Carlo in a spin eigenbasis via the graphical unitary group approach, J. Chem. Phys. **151**, 094104 (2019).

[85] W. Dobrautz, O. Weser, N. A. Bogdanov, A. Alavi, and G. L. Manni, Spin-pure stochastic-CASSCF via GUGA-FCIQMC applied to iron–sulfur clusters, J. Chem. Theory Comput. **17**, 5684 (2021).

[86] A. Y. Kitaev, Quantum measurements and the Abelian stabilizer problem, arXiv:quant-ph/9511026.

[87] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2009).

[88] M. Motta, T. P. Gujarati, J. E. Rice, A. Kumar, C. Masteran, J. A. Latone, E. Lee, E. F. Valeev, and T. Y. Takeshita, Quantum simulation of electronic structure with a transcorrelated hamiltonian: Improved accuracy with a smaller footprint on the quantum computer, Phys. Chem. Chem. Phys. **22**, 24270 (2020).

[89] A. Kumar, A. Asthana, C. Masteran, E. F. Valeev, Y. Zhang, L. Cincio, S. Tretiak, and P. A. Dub, Quantum simulation of molecular electronic states with a transcorrelated hamiltonian: Higher accuracy with fewer qubits, J. Chem. Theory Comput. **18**, 5312 (2022).

[90] J. Hubbard, Electron correlations in narrow energy bands, Proc. R. Soc. London A **276**, 238 (1963).

[91] J. Hubbard, Electron correlations in narrow energy bands III. An improved solution, Proc. R. Soc. London A **281**, 401 (1964).

[92] J. Kanamori, Electron correlation and ferromagnetism of transition metals, Prog. Theor. Phys. **30**, 275 (1963).

[93] F. C. Zhang, C. Gros, T. M. Rice, and H. Shiba, A renormalised hamiltonian approach to a resonant valence bond wavefunction, Supercond. Sci. Technol. **1**, 36 (1988).

[94] E. Dagotto, Correlated electrons in high-temperature superconductors, Rev. Mod. Phys. **66**, 763 (1994).

[95] P. W. Anderson, Superconductivity in high T_c cuprates: The cause is no longer a mystery, Phys. Scr. **T102**, 10 (2002).

[96] D. J. Scalapino, Numerical studies of the 2d hubbard model, *Handbook of High-Temperature Superconductivity* (Springer, New York, 2007), pp. 495–526.

[97] F. H. L. Essler, H. Frahm, F. Göhmann, A. Klümper, and V. E. Korepin, The bethe ansatz solution, *The One-Dimensional Hubbard Model* (Cambridge University Press, Cambridge, 2005), pp. 50–119.

[98] H. Bethe, Zur theorie der metalle, Z. Phys. **71**, 205 (1931).

[99] W. Metzner and D. Vollhardt, Correlated Lattice Fermions in $d = \infty$ Dimensions, Phys. Rev. Lett. **62**, 324 (1989).

[100] A. Georges and G. Kotliar, Hubbard model in infinite dimensions, Phys. Rev. B **45**, 6479 (1992).

[101] M. Jarrell, Hubbard Model in Infinite Dimensions: A Quantum Monte Carlo Study, Phys. Rev. Lett. **69**, 168 (1992).

[102] T. Schäfer, N. Wentzell, F. Šimkovic, Y.-Y. He, C. Hille, M. Klett, C. J. Eckhardt, B. Arzhang, V. Harkov, F.-M. Le Régent, A. Kirsch, Y. Wang, A. J. Kim, E. Kozik, E. A. Stepanov, A. Kauch, S. Andergassen, P. Hansmann, D. Rohe, Y. M. Vilk *et al.*, Tracking the Footprints of Spin Fluctuations: A MultiMethod, MultiMessenger Study of the Two-Dimensional Hubbard Model, Phys. Rev. X **11**, 011058 (2021).

[103] M. Qin, C.-M. Chung, H. Shi, E. Vitali, C. Hubig, U. Schollwöck, S. R. White, and S. Zhang (Simons Collaboration on the Many-Electron Problem), Absence of Superconductivity in the Pure Two-Dimensional Hubbard Model, Phys. Rev. X **10**, 031016 (2020).

[104] J. P. F. LeBlanc, A. E. Antipov, F. Becca, I. W. Bulik, G. K.-L. Chan, C.-M. Chung, Y. Deng, M. Ferrero, T. M. Henderson, C. A. Jiménez-Hoyos, E. Kozik, X.-W. Liu, A. J. Millis, N. V. Prokof'ev, M. Qin, G. E. Scuseria, H. Shi, B. V. Svistunov, L. F. Tocchio *et al.* (Simons Collaboration on the Many-Electron Problem), Solutions of the Two-Dimensional Hubbard Model: Benchmarks and Results from a Wide Range of Numerical Algorithms, Phys. Rev. X **5**, 041041 (2015).

[105] E. W. Huang, C. B. Mendl, H.-C. Jiang, B. Moritz, and T. P. Devereaux, Stripe order from the perspective of the Hubbard model, npj Quantum Mater. **3**, 22 (2018).

[106] M. C. Gutzwiller, Correlation of electrons in a Narrow s Band, Phys. Rev. **137**, A1726 (1965).

[107] T. Ogawa, K. Kanda, and T. Matsubara, Gutzwiller approximation for antiferromagnetism in hubbard model, Prog. Theor. Phys. **53**, 614 (1975).

[108] D. Vollhardt, Normal ^{3}He: An almost localized fermi liquid, Rev. Mod. Phys. **56**, 99 (1984).

[109] C. Gros, R. Joynt, and T. M. Rice, Antiferromagnetic correlations in almost-localized fermi liquids, Phys. Rev. B **36**, 381 (1987).

[110] P. Horsch and T. A. Kaplan, Exact and monte carlo studies of gutzwiller's state for the localised-electron limit in one dimension, J. Phys. C: Solid State Phys. **16**, L1203 (1983).

[111] T. A. Kaplan, P. Horsch, and P. Fulde, Close Relation between Localized-Electron Magnetism and the Paramagnetic Wave Function of Completely Itinerant Electrons, Phys. Rev. Lett. **49**, 889 (1982).

[112] W. Metzner and D. Vollhardt, Ground-State Properties of Correlated Fermions: Exact Analytic Results for the Gutzwiller Wave Function, Phys. Rev. Lett. **59**, 121 (1987).

[113] F. Gebhard and D. Vollhardt, Correlation Functions for Hubbard-Type Models: The Exact Results for the Gutzwiller Wave Function in One Dimension, Phys. Rev. Lett. **59**, 1472 (1987).

[114] J. M. Wahlen-Strothman, C. A. Jiménez-Hoyos, T. M. Henderson, and G. E. Scuseria, Lie algebraic similarity transformed hamiltonians for lattice model systems, Phys. Rev. B **91**, 041114(R) (2015).

[115] S. Yun, W. Dobrautz, H. Luo, and A. Alavi, Benchmark study of Nagaoka ferromagnetism by spin-adapted full configuration interaction quantum monte carlo, Phys. Rev. B **104**, 235102 (2021).

[116] K. Guther, W. Dobrautz, O. Gunnarsson, and A. Alavi, Time Propagation and Spectroscopy of Fermionic Systems Using a Stochastic Technique, Phys. Rev. Lett. **121**, 056401 (2018).

[117] E. Neuscamman, H. Changlani, J. Kinder, and G. K.-L. Chan, Nonstochastic algorithms for jastrow-slater and correlator product state wave functions, Phys. Rev. B **84**, 205132 (2011).

[118] H. Yokoyama and H. Shiba, Variational Monte-Carlo studies of hubbard model. I, J. Phys. Soc. Jpn. **56**, 1490 (1987).

[119] M. Capello, F. Becca, M. Fabrizio, S. Sorella, and E. Tosatti, Variational Description of Mott Insulators, Phys. Rev. Lett. **94**, 026406 (2005).

[120] H. Yokoyama and H. Shiba, Variational Monte-Carlo studies of hubbard model. II, J. Phys. Soc. Jpn. **56**, 3582 (1987).

[121] H. Luo, Variational transcorrelated method, J. Chem. Phys. **133**, 154109 (2010).

[122] H. Luo, W. Hackbusch, and H.-J. Flad, Quantum monte carlo study of the transcorrelated method for correlation factors, Mol. Phys. **108**, 425 (2010).

[123] H. Luo, Complete optimisation of multi-configuration jastrow wave functions by variational transcorrelated method, J. Chem. Phys. **135**, 024109 (2011).

[124] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, Evaluating analytic gradients on quantum hardware, Phys. Rev. A **99**, 032331 (2019).

[125] G. Aleksandrowicz, T. Alexander, P. Barkoutsos, L. Bello, Y. Ben-Haim, D. Bucher, F. J. Cabrera-Hernández, J. Carballo-Franquis, A. Chen, C.-F. Chen, J. M. Chow, A. D. Córcoles-Gonzales, A. J. Cross, A. Cross, J. Cruz-Benito, C. Culver, S. D. L. P. González, E. D. L. Torre, D. Ding, E. Dumitrescu *et al.*, QISKIT: An Open-source Framework for Quantum Computing (2019).

[126] S. Bravyi, S. Sheldon, A. Kandala, D. C. Mckay, and J. M. Gambetta, Mitigating measurement errors in multiqubit experiments, Phys. Rev. A **103**, 042605 (2021).

[127] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, A limited memory algorithm for bound constrained optimization, SIAM J. Sci. Comput. **16**, 1190 (1995).

[128] L. M. Milne-Thomson, *The Calculus of Finite Differences* (American Mathematical Society, Providence, RI, 2000).

[129] D. Wecker, M. B. Hastings, and M. Troyer, Progress towards practical quantum variational algorithms, Phys. Rev. A **92**, 042303 (2015).

[130] P. K. Barkoutsos, J. F. Gonthier, I. Sokolov, N. Moll, G. Salis, A. Fuhrer, M. Ganzhorn, D. J. Egger, M. Troyer, A. Mezzacapo *et al.*, Quantum algorithms for electronic structure calculations: Particle-hole hamiltonian and optimized wavefunction expansions, Phys. Rev. A **98**, 022322 (2018).

[131] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, Nature (London) **549**, 242 (2017).

[132] P. Gokhale, O. Angiuli, Y. Ding, K. Gui, T. Tomesh, M. Suchara, M. Martonosi, and F. T. Chong, Minimizing state preparations in variational quantum eigensolver by partitioning into commuting families, arXiv:1907.13623.

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

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

[135] E. van den Berg and K. Temme, Circuit optimization of hamiltonian simulation by simultaneous diagonalization of pauli clusters, Quantum **4**, 322 (2020).

[136] G. García-Pérez, M. A. C. Rossi, B. Sokolov, F. Tacchino, P. K. Barkoutsos, G. Mazzola, I. Tavernelli, and S. Maniscalco, Learning to measure: Adaptive informationally complete generalized measurements for quantum algorithms, PRX Quantum **2**, 040342 (2021).

[137] W. Dobrautz, I. O. Sokolov, K. Liao, P. López Ríos, M. Rahm, A. Alavi, and I. Tavernelli, *Ab initio* transcorrelated method enabling accurate quantum chemistry on near-term quantum hardware, arXiv:2303.02007.

[138] A. Cultrera and L. Callegaro, A simple algorithm to find the l-curve corner in the regularisation of ill-posed inverse problems, IOP SciNotes **1**, 025004 (2020).

[139] P. Taylor, *Coupled-cluster Methods in Quantum Chemistry* (Springer, Berlin, 1994).

[140] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz, Quantum Sci. Technol. **4**, 014008 (2018).

[141] I. O. Sokolov, P. K. Barkoutsos, P. J. Ollitrault, D. Greenberg, J. Rice, M. Pistoia, and I. Tavernelli, Quantum orbital-optimized unitary coupled cluster methods in the strongly correlated regime: Can quantum algorithms outperform their classical equivalents? J. Chem. Phys. **152**, 124107 (2020).

[142] N. Gomes, F. Zhang, N. F. Berthusen, C.-Z. Wang, K.-M. Ho, P. P. Orth, and Y. Yao, Efficient step-merged quantum imaginary time evolution algorithm for quantum chemistry, J. Chem. Theory Comput. **16**, 6256 (2020).

[143] C. Cade, L. Mineh, A. Montanaro, and S. Stanisic, Strategies for solving the fermi-hubbard model on near-term quantum computers, Phys. Rev. B **102**, 235122 (2020).

[144] L. Xu, J. T. Lee, and J. Freericks, Test of the unitary coupled-cluster variational quantum eigensolver for a simple strongly correlated condensed-matter system, Mod. Phys. Lett. B **34**, 2040049 (2020).

[145] X. Bonet-Monroig, H. Wang, D. Vermetten, B. Senjean, C. Moussa, T. Bäck, V. Dunjko, and T. E. O'Brien, Performance comparison of optimization methods on variational quantum algorithms, Phys. Rev. A **107**, 032407 (2023).

[146] M. R. Hoffmann and J. Simons, A unitary multiconfigurational coupled-cluster method: Theory and applications, J. Chem. Phys. **88**, 993 (1988).

[147] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, An adaptive variational algorithm for exact molecular simulations on a quantum computer, Nat. Commun. **10**, 3007 (2019).

[148] H. L. Tang, V. O. Shkolnikov, G. S. Barron, H. R. Grimsley, N. J. Mayhall, E. Barnes, and S. E. Economou, Qubit-ADAPT-VQE: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor, PRX Quantum **2**, 020310 (2021).

[149] Y. S. Yordanov, V. Armaos, C. H. Barnes, and D. R. Arvidsson-Shukur, Qubit-excitation-based adaptive variational quantum eigensolver, Commun. Phys. **4**, 228 (2021).

[150] A. G. Rattew, S. Hu, M. Pistoia, R. Chen, and S. Wood, A domain-agnostic, noise-resistant, hardware-efficient evolutionary variational quantum eigensolver, arXiv:1910.09694.

[151] D. Chivilikhin, A. Samarin, V. Ulyantsev, I. Iorsh, A. Oganov, and O. Kyriienko, MoG-VQE: Multiobjective genetic variational quantum eigensolver, arXiv:2007.04424.