Gutzwiller wave function on a digital quantum computer

Bruno Murta
QuantaLab, International Iberian Nanotechnology Laboratory (INL), 4715-330 Braga, Portugal
Departamento de Física, Universidade do Minho, 4710-057 Braga, Portugal

J. Fernández-Rossier
QuantaLab, International Iberian Nanotechnology Laboratory (INL), 4715-330 Braga, Portugal
Departamento de Física Aplicada, Universidad de Alicante, San Vicente del Raspeig 03080, Spain

(Dated: March 30, 2021)

The determination of the ground state of quantum many-body systems via digital quantum computers rests upon the initialization of a sufficiently educated guess. This requirement becomes more stringent the greater the system. Preparing physically-motivated ansätze on quantum hardware is therefore important to achieve quantum advantage in the simulation of correlated electrons. In this spirit, we introduce the Gutzwiller Wave Function (GWF) within the context of the digital quantum simulation of the Fermi-Hubbard model. We present a quantum routine to initialize the GWF that comprises two parts. In the first, the noninteracting state associated with the $U = 0$ limit of the model is prepared. In the second, the non-unitary Gutzwiller projection that selectively removes states with doubly-occupied sites from the wave function is performed by adding to every lattice site an ancilla qubit, the measurement of which in the $|0\rangle$ state confirms the projection was made. Due to its non-deterministic nature, we estimate the success rate of the algorithm in generating the GWF as a function of the lattice size and the interaction strength $U/t$. The scaling of the quantum circuit metrics and its integration in general quantum simulation algorithms are also discussed.

The quantum many-body problem permeates a wide range of fields of research within condensed matter physics, quantum chemistry, and materials science. In particular, it is the cornerstone of the electronic structure problem. Conventional-hardware-based numerical methods have played a pivotal role in unravelling the electronic structure of materials, but not without shortcomings. Indeed, although the low-energy properties of weakly interacting materials are well described by Density-Functional Theory with approximate functionals based on the Local Density Approximation, this approach often fails when strong electron-electron interactions prevail. Quantum Monte Carlo methods are a leading alternative, though often plagued by the sign problem.

In principle, a more accurate description of correlated fermions could be achieved via wave-function-based methods, but the storage and manipulation of the wave function in classical hardware is hampered by the exponential wall problem. This is, however, not the case in quantum hardware, thanks to the principle of superposition and the natural encoding of entanglement. Quantum computers have thus been proposed as a platform to simulate quantum many-body models that encapsulate the electronic structure of materials when their understanding demands an explicit representation of the wave function, either because of the presence of strong correlations or a high accuracy requirement.

A number of quantum algorithms to determine the ground state of a given Hamiltonian $\mathcal{H}$ have been put forth. The most prominent example is Quantum Phase Estimation (QPE), whereby an initial state with non-negligible overlap with the exact ground state undergoes time evolution under the action of the propagator $e^{-i\mathcal{H}t/\hbar}$ subject to the control of ancilla qubits, from which the eigenspectrum can be extracted after the application of the inverse quantum Fourier transform. In particular, once the ground state energy is read out from the ancilla qubits, the state collapses into the exact ground state.

Because the resources required to implement QPE are far beyond the capacity of near-term quantum processors, a leaner class of hybrid variational algorithms have been developed, of which the Variational Quantum Eigensolver (VQE) is the reference. In VQE, a parameterized state is prepared on a quantum computer, which is used to compute the expectation value of $\mathcal{H}$. This energy is then provided to a classical computer that performs the optimization routine to find the updated parameter values, which are then fed back to the quantum computer to begin the next iteration. Alternatively, the ground state can be found via Quantum Imaginary Time Evolution (QITE), which can be implemented in quantum hardware by casting it into a variational problem or by finding the unitary operation that transforms the state at the current step, $|\psi(\tau)\rangle$, into the (normalized) state at the next step, $|\psi(\tau + \Delta \tau)\rangle = \mathcal{N} e^{-i\mathcal{H}\Delta \tau/\hbar} |\psi(\tau)\rangle$.

Despite the development of the aforementioned quantum simulation algorithms, it is well established that the problem of finding the ground state of many Hamiltonians involving only local interactions is QMA-complete. Regarding QPE and QITE, the key challenge lies in the preparation of the initial state: the scaling of both algorithms is polynomial with respect to the inverse of the overlap between the initial state and the exact ground state, so the state preparation routine should lead to a polynomially decreasing overlap as the system size increases for the overall algorithm to be efficient. Nevertheless, conventional choices of ini-
tial states, such as noninteracting or mean-field ground states, produce an exponentially vanishing probability of collapsing into the ground state due to the orthogonality catastrophe. This overlap can in principle be enhanced via adiabatic evolution, but its success depends on the gap between the ground and first excited states throughout the adiabatic path, which is generally unknown.

VQE, in turn, is a heuristic, so, contrary to QPE and QITE, there is no theoretical guarantee of its success given an appropriate initial state. In fact, the challenge in VQE is not quite preparing the initial state, but rather formulating a parameterized ansatz such that its manifold includes a path connecting the initial state to the exact ground state without incurring in an exponential number of parameters. In any case, preparing an initial state with a greater overlap with the exact ground state could simplify the variational procedure by shortening the path in parameter space that needs to be covered, potentially reducing the number of optimization steps or the layers of the ansatz. Perhaps even more important is the possible avoidance of barren plateaus, which can be linked to an uninformed initialization of the ansatz.

Given the crucial role played by the initial state in the success of digital quantum simulation schemes, developing routines to prepare on quantum hardware physically-motivated ansätze is of great importance. It is within this context that, in this paper, we propose an algorithm to prepare on quantum hardware the Gutzwiller Wave Function (GWF). This is a conceptually simple, physically-motivated ansatz that contains a single parameter. It can be applied to the wide class of lattice models of correlated electrons where the electron-electron Coulomb repulsion is described, to leading order, in terms of the on-site Hubbard interaction $U \hat{n}_{i, \uparrow} \hat{n}_{i, \downarrow}$, where $\hat{n}_{i, \sigma} = \hat{a}_{i, \sigma}^{\dagger} \hat{a}_{i, \sigma}$ is the number operator acting on site $i$ for electrons of spin $\sigma = \uparrow, \downarrow$, with $\hat{a}_{i, \sigma}$ the annihilation operator in second quantization.

The Gutzwiller Wave Function (GWF) is defined as

$$|\psi_G\rangle = \prod_{i=1}^{N} (1 - g \hat{n}_{i, \uparrow} \hat{n}_{i, \downarrow}) |\psi_0\rangle,$$

where $g \in [0, 1]$ is a free parameter, $N$ is the number of lattice sites, and $|\psi_0\rangle$ is the noninteracting ground state. In words, the GWF is prepared by reducing the amplitude by which the amplitude is decreased is set by $g$. The optimal value of $g$ for a given Hubbard parameter $U$ is found by minimizing the energy. The greater the magnitude of the on-site Hubbard interactions, the more unfavorable the doubly-occupied states are, and hence the greater $g$ is. In particular, when $U = 0$, $g = 0$, and when $U \to \infty$, $g = 1$. Importantly, the Gutzwiller projection does not break any symmetry of the Hamiltonian.

In spite of its apparent simplicity, the GWF is a complex state that captures some correlations between the electrons. This complexity can be understood as follows. The noninteracting ground state

$$|\psi_0\rangle = \prod_{\alpha, \sigma} a_{\alpha, \sigma}^{\dagger} |0\rangle,$$

is a single Slater determinant, or a Fock state, when the single-particle basis $\{|\alpha\rangle\}$ is chosen to be the eigenbasis of the noninteracting tight-binding Hamiltonian. However, the Gutzwiller projection is carried out in the site basis, therefore the operators $\{a_{\alpha, \sigma}^{\dagger}\}$ have to be expanded as a linear combination of site operators $\{a_{i, \alpha}^{\dagger}\}$,

$$|\psi_0\rangle = \prod_{\alpha, \sigma} \left( \sum_{i} \phi_{\alpha}(i) a_{i, \sigma}^{\dagger} \right) |0\rangle,$$

in which case $|\psi_0\rangle$ appears as multi-determinant state.

As a result, the determination of expectation values of the GWF is a many-body problem that cannot be solved exactly except in the special cases of one and infinite dimensions. Hence, in two and three dimensions, numerical methods such as Variational Monte Carlo have been employed to compute expectation values of the GWF. The GWF has been used to model a variety of correlated fermion problems, such as the metal-insulator transition, the low-temperature behaviour of $^{3}$H$_{3}$ and the superconductivity in the cuprate.

In order to appreciate the significance of the Gutzwiller projection as an improved starting point for general quantum simulation algorithms, we consider the Fermi-Hubbard Model (FHM),

$$\hat{H} = -t \sum_{i=1}^{N} \sum_{\sigma=\uparrow, \downarrow} (\hat{a}_{i, \sigma}^{\dagger} \hat{a}_{i+1, \sigma} + \text{H.c.}) + U \sum_{i=1}^{N} \hat{n}_{i, \uparrow} \hat{n}_{i, \downarrow},$$

at half-filling in a one-dimensional lattice with open boundary conditions and with up to $N = 12$ sites. Specifically, we compare the fidelity with respect to the exact ground state of three trial states: the GWF, the noninteracting and the self-consistent mean-field ground states. The exact ground state is obtained via exact diagonalization using the numerical package QuSpin. The self-consistent mean-field theory amounts to a direct Hartree-Fock decoupling of the quartic Hubbard term, which gives rise to coefficients that depend on the expectation values of occupation numbers. Self-consistency is attained when single-particle occupations of the mean-field ground state coincide (to a given precision) with the input values. Random initial conditions are used, and the lowest-energy state out of all trials is selected.

Fig. 1(a) presents the fidelity of these three states with respect to the exact ground state for a chain of $N = 12$ sites; similar profiles are observed for chains of different sizes. For concreteness, Fig. 1(b) shows how the fidelity of the three states decays with the size of the chain for the specific value of $U/t = 10$. The decrease of the overlap between the GWF and the exact ground state is significantly slower, thus rendering it a far better
 FIG. 1. Comparison of Gutzwiller Wave Function (GWF) to noninteracting and self-consistent mean-field ground states for Fermi-Hubbard Model (FHM) in chain with open boundary conditions at half-filling. (a) Fidelity of three reference states with respect to exact ground state for a chain of $N = 12$ sites against normalized Hubbard parameter $U/t$. (b) Same as (a), but now size of chain $N$ is varied between 2 and 12 sites, while $U/t$ takes the fixed value of 10, as highlighted in (a) by the vertical dotted line. Numerical results were fitted to exponential decay $c_1 e^{-c_2 N}$.

 FIG. 2. (a) High-level scheme of routine to initialize GWF on a quantum computer, exemplified for a system of $N = 4$ sites. After the noninteracting ground state $|\psi_0\rangle$ is prepared, a single-qubit rotation is applied to each auxiliary qubit if the two qubits that encode the occupations of the corresponding site are in state $|11\rangle$. Measuring the auxiliary qubits and retrieving only the trials that yield 0000 confirms the Gutzwiller projection was executed. (b) Detailed scheme of routine to initialize GWF assuming linear qubit connectivity. Solid red box corresponds to preparation of $|\psi_0\rangle$, which is first initialized in the diagonal basis and then transformed back to the original one via a Givens rotation decomposition, followed by the reordering of the qubits by site instead of spin, which involves fermionic SWAPs. Dashed blue box is the Gutzwiller projection: network of SWAPs places auxiliary qubits next to control-qubits. A detailed description can be found in the Supplemental Material.

Henceforth, the initialization of the GWF on a digital quantum computer will be discussed. To the best of our knowledge, no routine has been proposed to accomplish this. The Jordan-Wigner mapping will be assumed, in which case the computational basis states $|0\rangle$ and $|1\rangle$ of each qubit encode the occupation of a spin orbital at a given site (unoccupied and occupied, respectively), yielding a total of $2N$ qubits to represent the wave function, where the factor of 2 is due to the spin degeneracy.

Unsurprisingly, the algorithm proceeds in two stages, as shown in Fig. 2: first, the noninteracting ground state $|\psi_0\rangle$ is prepared, then the Gutzwiller projection is applied. For the preparation of $|\psi_0\rangle$, we follow previous works based on the exploitation of the Thouless theorem and the expression of the resulting unitary operation as a quantum circuit using a QR decomposition via Givens rotations. As for the implementation of the Gutzwiller projector, the challenge associated with its non-unitary character needs to be overcome. This can be accomplished by embedding the projector in a larger unitary operation.

Since the Gutzwiller projection acts on each site separately, let us consider a single site, which is represented by two qubits, one for each spin. Let us add an auxiliary
qubit, to which the single-qubit gate
\[
U(g) = \left( \frac{1 - g}{\sqrt{2g - g^2}} \right) \left( 1 - g \right) - \frac{\sqrt{2g - g^2}}{1 - g}
\]
is applied if and only if the two qubits that encode the occupations of the site are in state \(|11\rangle\). Hence, for a given arbitrary two-qubit state \(|\Phi\rangle = c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle\), the state after the \(ccU(g)\) reads:
\[
ccU(|\Phi\rangle \otimes |0\rangle) = (c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle) \otimes |0\rangle + c_{11}|11\rangle \otimes \left( (1 - g)|0\rangle + \sqrt{2g - g^2}|1\rangle \right)
\]
Once the ancilla is measured in the computational basis, the state of the main register is collapsed into either
\[
|\Phi_0\rangle = c_{00}|00\rangle + c_{10}|01\rangle + c_{10}|01\rangle + c_{11}(1 - g)|11\rangle
\]
or
\[
|\Phi_1\rangle = c_{11}\sqrt{2g - g^2}|11\rangle,
\]
where the subscript denotes the outcome of the readout. Notice that \(|\Phi_0\rangle\) coincides with the action of the Gutzwiller projection on the initial state \(|\Phi\rangle\). Hence, to guarantee that the Gutzwiller projection is applied to this site, the auxiliary qubit must be initialized in \(|0\rangle\) and measured in \(|0\rangle\) after the application of \(ccU(g)\). The projection method is thus non-deterministic.

The application of the Gutzwiller projector to the whole wave function merely amounts to repeating this procedure for every site, as illustrated in Fig. 2(a). For the sake of clarity, let us compute its action on the following four-site wave function explicitly:
\[
|\psi\rangle = a|\uparrow, \uparrow, \downarrow, \downarrow\rangle + b|\uparrow, \uparrow, \downarrow, \downarrow\rangle + c|\downarrow, \downarrow, \uparrow, \uparrow\rangle + (1 - g)\sqrt{2g - g^2}|\downarrow, \downarrow, \uparrow, \uparrow\rangle
\]
First, let us obtain the expected outcome by applying the Gutzwiller projector \(\hat{P}_G(g) = \prod_{i=1}^N \mathbb{1} - g \hat{n}_i \hat{\bar{n}}_i\) to \(|\psi\rangle\):
\[
\hat{P}_G(g)|\psi\rangle = a|\uparrow, \uparrow, \downarrow, \downarrow\rangle + b|\uparrow, \uparrow, \downarrow, \downarrow\rangle + d|\uparrow, \downarrow, \uparrow, \downarrow\rangle + (1 - g)^2|\uparrow, \downarrow, \uparrow, \downarrow\rangle + (1 - g)^2|\uparrow, \downarrow, \uparrow, \downarrow\rangle.
\]
In words, the amplitude of the basis states with \(n\) doubly-occupied sites is reduced by a factor of \((1 - g)^n\). After normalization, the basis states with no doubly-occupied sites have a greater amplitude than originally. Let us now compare this to the action of the proposed quantum routine on \(|\psi\rangle\). Adding the four ancillas initially in \(|0000\rangle\), the wave function after the application of the four \(ccU\) but before the measurement of the ancillas is
\[
|0000\rangle \otimes \left( (1 - g)|0\rangle + \sqrt{2g - g^2}|1\rangle \right) = (1 - g)|0\rangle + \sqrt{2g - g^2}|1\rangle.
\]
The part of the wave function associated with the state \(|0000\rangle\) in the auxiliary register coincides with \(\hat{P}_G(g)|\psi\rangle\). Hence, all ancillas have to be measured in \(|0\rangle\) to confirm the Gutzwiller projection was applied to the full state.

Given this method to apply the Gutzwiller projection on quantum hardware, the only piece left in the initialization of the GWF is the determination of the optimal value of \(g\). This can be accomplished by minimizing the energy explicitly on the quantum computer, computing the expectation value of the Hamiltonian either via QPE or by decomposing it in the Pauli basis, as in VQE. Nevertheless, one can deduce the optimal value of \(g\) for a large system by extrapolating from small system simulations carried out on a conventional computer. This is indeed the case for the one-dimensional FHM at half-filling, for which minor variations in the \(g(U)\) relation are observed as the size of the chain varies. The same extrapolation should be possible for rectangular and square lattices.

In the remainder of this paper, we will discuss the scalability of the quantum routine herein proposed to initialize the GWF. In particular, the scaling of the relevant quantum circuit metrics (depth, width and number of CNOTs) and of the number of repetitions due to the non-deterministic nature of the method will be detailed.

Regarding the implementation of the quantum circuit for the Gutzwiller projection, the resulting overhead is found to be analogous to that associated with the preparation of the noninteracting ground state. Given \(N\) lattice sites, \(N\) additional qubits are required to perform the Gutzwiller projection, so the circuit width is \(3N\) rather than \(2N\). As for the circuit depth, although the controlled-rotations all act on independent trios of
TABLE I. Average number of repetitions required to prepare GWF on a digital quantum computer for one-dimensional FHM at half-filling with open boundary conditions. Values corresponding to $N = 10$ sites were obtained directly from simulation, while remaining ones were extrapolated from simulations of chains with $N = 2, 4, 6, 8, 10, 12$ sites. Cf. Supplementary Material for further details.

| $N$  | $U/t = 1$ | $N = 20$ | $N = 30$ | $N = 40$ |
|------|-----------|----------|----------|----------|
| 5    | 2.7       | 6.5      | 16       | 39       |
| 10   | 29        | 940      | 30,700   | 1,000,000|
| 30   | 63        | 5,900    | 550,000  | 51,000,000|
| 50   | 77        | 9,000    | 9,000    | 9,000    |

TABLE II. Estimate of average number of repetitions due to the choice of initial state $|\psi_{\text{initial}}\rangle$ that are required to find exact ground state $|\psi_{\text{exact}}\rangle$ of Fermi-Hubbard chain with $U/t = 10$ at half-filling via QPE or QITE. The initial states considered are the noninteracting ground state $|\psi_0\rangle$, the self-consistent mean-field ground state $|\psi_{\text{MF}}\rangle$ and the GWF, the latter under two assumptions: excluding repetition overhead to initialize it $|\psi_G\rangle$, and including this overhead $|\psi_G^{**}\rangle$. The number of trials due to the choice of initial state in QPE and QITE is estimated as $1/|\langle \psi_{\text{exact}}|\psi_{\text{initial}}\rangle|^2$ (cf. Supplementary Material for a detailed explanation). Values corresponding to $N = 10$ sites were obtained directly from simulation, while remaining ones were extrapolated from simulations of chains with $N = 2, 4, 6, 8, 10, 12$ sites.

| $N$  | 10  | 20  | 30  | 40  |
|------|-----|-----|-----|-----|
| $|\psi_0\rangle$ | 11  | 120 | 1,500 | 17,000 |
| $|\psi_{\text{MF}}\rangle$ | 7.8 | 70  | 580  | 5,000  |
| $|\psi_G\rangle$ | 1.1 | 1.2 | 1.4  | 1.6    |
| $|\psi_G^{**}\rangle$ | 69  | 7,100 | 770,000 | 82,000,000 |

Successful preparation of the GWF. This quantity depends both on $N$ and the value of $g$. In turn, $g$ depends on $U$. We have carried out this estimate numerically for chains with $N = 2, 4, 6, 8, 10, 12$ sites, and then extrapolated to larger chains. Table I shows the number of repetitions required to initialize the GWF for the FHM at half-filling for multiple chain sizes. Even for as many as $N = 40$ sites, the number of repetitions, though large, is clearly feasible for all $U/t$ values. We note that, in a similar spirit to VQE, this Gutzwiller projection scheme corresponds to a relatively shallow circuit that must be repeated multiple times. However, whereas in VQE these repetitions serve to explore the parameter space in search of the global minimum, in the Gutzwiller projection their cause is the non-deterministic nature of the routine.

However, if the GWF is needed as a starting point to carry out QPE or QITE, the improved fidelity does not offset the additional number of repetitions. This is illustrated in Table II, which compares the average number of repetitions due to the choice of initial state that are required to find the exact ground state, taking the initial state to be the noninteracting ground state $|\psi_0\rangle$, the mean-field ground state $|\psi_{\text{MF}}\rangle$ and the GWF, the latter case under two assumptions: excluding the repetition overhead to prepare the GWF $|\psi_G^{**}\rangle$, and including this overhead $|\psi_G^{**}\rangle$. The repetition overhead overwhelms the savings arising from replacing $|\psi_0\rangle$ or $|\psi_{\text{MF}}\rangle$ by $|\psi_G\rangle$. Likewise, the potential reduction of the number of layers or parameter updates in VQE arising from taking the GWF as the reference state will probably not compensate the increase in the number of repetitions, unless the barren plateau problem is especially acute.

One potentially promising route to tackle the repetition overhead issue is to combine this non-deterministic scheme with quantum amplitude amplification\cite{69-71} for a suitable choice of oracle. This could lead to a trade-off between the number of repetitions and the circuit depth, thus making it a more viable solution for long-term quantum simulation algorithms such as QPE. Another line of research to be explored that could be especially relevant for near-term quantum simulation algorithms such as VQE is the preparation of the GWF via an appropriate parameterized quantum circuit.

In conclusion, the GWF was proposed as a promising trial state for the digital quantum simulation of strongly-correlated electrons. We developed the first routine to prepare the GWF on a gate-based quantum computer, the circuit depth and width requirements of which are similar to those associated with the initialization of the noninteracting ground state. The non-deterministic nature of this scheme leads to a repetition overhead that, though viable for the implementation of the GWF on its own, must be reduced in order to integrate it in general quantum simulation algorithms.

Acknowledgements. B.M. acknowledges support from the FCT PhD scholarship no. SFRH/BD/08444/2020. JFR acknowledges financial support from the Spanish Government (grant no. PID2019-109539GB-C41), and...
Generalitat Valenciana funding Prometeo 2017/139. The authors acknowledge use of the QuSpin\textsuperscript{33} package for the exact diagonalization calculations and of Qiskit\textsuperscript{48} for the in-silico simulations of the quantum circuits.

**Supplemental Material for “Gutzwiller wave function on a digital quantum computer”**

I. BASIS GATE DECOMPOSITION AND SCALING OF QUANTUM CIRCUIT DEPTH AND WIDTH

This section details how to decompose each operation in the quantum circuit shown in Fig. 2 of the main text in terms of basis gates. The reference two-qubit gate assumed is the CNOT\textsuperscript{49}. The initialization of the Gutzwiller Wave Function (GWF) on a digital quantum computer comprises two parts: the preparation of the noninteracting ground state $|\psi_0\rangle$ and the Gutzwiller projection $\hat{P}_G(g)$.

**Preparation of Noninteracting Ground State**

If the noninteracting Hamiltonian is spin-polarized (which is generally the case, since hopping is not normally assumed to mix different spin states), the preparation of $|\psi_0\rangle$ can be treated separately for the two spins $\uparrow$ and $\downarrow$, which effectively leads to two subcircuits that can be executed in parallel, as shown in Fig. 2(b) of the main text for the spin-balanced case $N_\uparrow = N_\downarrow = 2$ (in which case the two subcircuits are equal). We can thus consider only one spin state in the following discussion.

![Fig. S1. Scheme of quantum circuits to initialize spin-polarized noninteracting ground state of a given Hamiltonian for $N_s = N/2$, with $N = 4$ (left) and $N = 6$ (center) sites. In both examples, $N_s = N/2$, where $s = \uparrow, \downarrow$ is the selected spin. For a different number of electrons, the number of NOT gates in the initialization of the reference Slater determinant and the number of Givens rotations $G_i$ in the decomposition of the unitary transformation between the two bases would change. The two-qubit operation $G_i$ can be decomposed in terms of more elementary gates as shown on the right-hand side.](image)

Hence, each $G_i$ operation requires 4 CNOTs. Since a total of $N_s(N - N_s)$ $G_i$ operations are required to perform the basis transformation, this gives a total of $2 \cdot N_s(N - N_s) \cdot 4 = 8N_s(N - N_s)$ CNOTs, where the extra factor of 2 is due to the spin degeneracy, assuming, as usual, that $N_\uparrow = N_\downarrow$. At half-filling, $N_s = N/2$, so the total number of CNOTs in this case is $2N^2$. As for the circuit depth, the unitary transformation that relates the two bases involves $N - 1$ layers of $G_i$ operations regardless of the filling level. Ignoring single-qubit gates, which have execution times and error rates far lower than those of CNOTs, and noting that each $G_i$ involves 4 CNOTs that must be applied sequentially, the total circuit depth for this part is $4(N - 1)$. Note that there is no factor of 2 due to the spin degeneracy, since the two subcircuits, one for each spin, can be executed in parallel.
FIG. S2. Scheme of quantum circuit to reorder qubits by site instead of spin after the initialization of the spin-polarized noninteracting ground states in parallel for \( N = 4, 6 \) sites. After the application of this array of fermionic-SWAPs, the two qubits that encode the occupations of the two spin-orbitals associated with the same site are next to each other.

Before the noninteracting ground state \(|\psi_0\rangle\) is used as the starting point of a quantum simulation algorithm such as Quantum Phase Estimation (QPE)\(^9\) or the Variational Quantum Eigensolver (VQE)\(^{12}\), it is common practice to reorder the qubits by site instead of spin (i.e. place the two qubit that encode the occupations of the two spin states at the same site next to each other) when dealing with lattice models of correlated electrons with on-site Hubbard interaction\(^{44,45}\). This assumes, of course, that the qubit connectivity is limited, namely that qubits are only connected linearly. This qubit reordering therefore allows to reduce the circuit depth of the implementation of the propagator \( e^{-i\mathcal{H}t/\hbar} \), and it can be accomplished via the quantum circuits shown in Fig. S2 for \( N = 4, 6 \) sites. To account for the exchange anti-symmetry, fermionic-SWAPs are required, which are analogous to conventional SWAPs, except for the extra minus sign applied to the basis state \(|11\rangle\). The basis gate decomposition of a fermionic-SWAP is

\[
\begin{array}{cccc}
|1\uparrow\rangle & |2\uparrow\rangle & |3\uparrow\rangle & |4\uparrow\rangle \\
|1\downarrow\rangle & |2\downarrow\rangle & |3\downarrow\rangle & |4\downarrow\rangle \\
\end{array}
\]

Hence, the reordering of the \( 2N \) qubits by site involves a total of \( N(N - 1)/2 \) fermionic-SWAPs, each taking 4 CNOTs, which brings the total CNOT count to \( 2N(N - 1) \). As for the circuit depth, \( N - 1 \) layers of fermionic-SWAPs are needed and each takes 4 CNOTs, so the total circuit depth (in CNOTs) is of \( 4(N - 1) \).

To sum up, the initialization of the noninteracting ground state corresponds to a total circuit depth of \( 8N - 8 \) CNOTs and a total number of CNOTs of \( 8N_s(N - N_s) + 2N(N - 1) \) at any filling, or \( 4N^2 - 2N \) at half-filling.

**Gutzwiller Projection**

Ignoring any qubit connectivity constraints for the moment, all \( N \) controlled-controlled-\( \mathcal{U}(g) \) corresponding to the Gutzwiller projection can be implemented in parallel, since each acts on its own set of three qubits. Since \( \mathcal{U}(g) = R_y(\alpha(g)) \), with \( \alpha(g) = 2 \arctan(\sqrt{2g - g^2/(1 - g)}) \), the basis gate decomposition of the controlled-controlled-\( \mathcal{U}(g) \) can be derived immediately from that of the controlled-\( R_y(\theta) \) (cf. previous page):

\[
\begin{array}{cccc}
|1\uparrow\rangle & |2\uparrow\rangle & |3\uparrow\rangle & |4\uparrow\rangle \\
|1\downarrow\rangle & |2\downarrow\rangle & |3\downarrow\rangle & |4\downarrow\rangle \\
\end{array}
\]

Hence, each controlled-controlled-\( \mathcal{U}(g) \) requires two Toffoli gates\(^{50}\), each of which can be implemented as\(^{49}\).

In summary, assuming all-to-all connectivity, the Gutzwiller projection corresponds to a circuit depth of \( 2 \cdot 6 = 12 \) CNOTs (i.e. independent of the size of the system) and a total number of CNOTs of \( 12N \).

However, in current quantum processors, the assumption of linear qubit connectivity is far more realistic than all-to-all connectivity. In such case, there are two main changes relative to the analysis above: first, the auxiliary qubits must be moved next to the two qubits associated with the respective site; second, in the basis gate decomposition
of the Toffoli gate shown above, the middle and bottom qubits must be swapped in order to execute the first four CNOTs (assuming only the top/middle and middle/bottom pairs of qubits are connected to each other). The latter change, shown in the scheme below, leads to a total of $2 \cdot 3 + 6 = 12$ CNOTs, as opposed to the original 6.

Hence, with linear qubit connectivity, the implementation of the $N$ controlled-controlled-$U(g)$ gates involves $24N$ CNOTs in total and a circuit depth of $24$ CNOTs.

![Diagram](image)

FIG. S3. Scheme of quantum circuit to place auxiliary qubits next to respective control-qubits to implement controlled-controlled-$U(g)$ corresponding to Gutzwiller projection for $N = 4, 6$ sites assuming linear qubit connectivity constraint. The inverse of this network of SWAPs must be applied after the execution of the controlled-controlled-$U(g)$ to return the $2N$ qubits in the main register to their ideal positions for the remainder of the quantum simulation.

The placement of the auxiliary qubits next to the two respective control-qubits requires a network of conventional SWAPs such as the ones shown in Fig. S3 for $N = 4, 6$ sites. Note, however, that this network of SWAPs must be repeated after the execution of the controlled-controlled-$U(g)$ gates so that the auxiliary qubits are returned to their original positions and, more importantly, the $2N$ qubits of the main register are ready for the remaining part of the quantum simulation algorithm to be implemented. In total, this therefore requires $2N(N - 1)$ SWAPs, or $6N(N - 1)$ CNOTs, and a circuit depth of $4(N - 1)$ SWAPs, or $12(N - 1)$ CNOTs. To sum up, the Gutzwiller projection requires a total of $6N^2 + 18N$ CNOTs and a circuit depth of $12N + 12$ CNOTs.

Generalization to Arbitrary Geometries and Dimensions

It should be stressed that this analysis of the scaling of the circuit depth, width and total number of CNOTs of the quantum circuit corresponding to the Gutzwiller projection routine is agnostic to the geometry of the lattice considered. Regarding the preparation of the noninteracting ground state, what determines the number of Givens
rotations required is simply the number of electrons and the number of single-particle orbitals, the latter being set by the number of lattice sites only. As for the Gutzwiller projection itself, being a local operator, it is clearly independent of the geometry. Finally, the SWAP networks exemplified schematically in Figs. S2 and S3 are applicable to lattices of arbitrary geometry and dimension. The only underlying assumption is that, at the moment of the preparation of the spin-polarized noninteracting states, the qubits corresponding to a given spin \( s = \uparrow, \downarrow \) are already ordered according to the convention chosen for the labelling of the lattice sites, whatever that is from a spatial point of view.

### II. SCALING OF REPETITION OVERHEAD

The Gutzwiller projection routine herein proposed is probabilistic, in that the GWF is only initialized on the quantum computer if all auxiliary qubits are measured in \(| 0 \rangle\). In this section, the average number of repetitions required to prepare the GWF is determined. The application of the \( \bar{N} \) controlled-controlled-\( \bar{U}(g) \) gates on \(| 0 \rangle^{\otimes N} \otimes |\psi_0 \rangle \) gives \(| \psi' \rangle \equiv | 0 \rangle^{\otimes N} \otimes |\psi_G \rangle + \ldots \), where \(|\psi_G \rangle = \hat{P}_G(g) |\psi_0 \rangle \) is the GWF and the ellipsis refers to the remaining terms of the wave function for which the auxiliary qubits are in a computational basis state other than \(| 0 \rangle^{\otimes N} \). The probability that all auxiliary qubits are measured in the \(| 0 \rangle\) state is given by

\[
P(N, g) = \sum_{|q_1 q_2 \ldots q_{2N}| \in \{0,1\}^{2N}} \left| \langle (0)^{\otimes N} \otimes \langle q_1 q_2 \ldots q_{2N} | \rangle |\psi' \rangle \right|^2 = \left| \langle \psi_G |\psi_G \rangle \right|^2, \quad (S1)
\]

where \(|q_1 q_2 \ldots q_{2N}| \rangle \) is the main 2\( N \)-qubit register where \(|\psi_0 \rangle \) is initialized and \(|\psi_G \rangle \) is later obtained. Notice that, due to the non-unitarity of \( \hat{P}_G(g) \), \(|\psi_G \rangle \) is not normalized, which is why the probability of preparing it is not generally equal to 1. The noninteracting ground state \(|\psi_0 \rangle\) at half-filling for a lattice of \( N \) sites can be written as

\[
|\psi_0 \rangle = \sum_{n=0}^{N/2} \sum_{i=1}^{M_n} c_i^{(n)} |\phi_i^{(n)} \rangle, \quad (S2)
\]

where \( n \) is the number of doubly-occupied sites, and therefore \(|\phi_i^{(n)} \rangle \) is a basis state with \( n \) doubly-occupied sites, with \( M_n = \binom{N}{n} \binom{N-n}{n} \binom{N-2n}{N/2} \) the number of basis states with \( n \) doubly-occupied sites. The GWF is given by

\[
|\psi_G \rangle = \hat{P}_G(g) |\psi_0 \rangle = \sum_{n=0}^{N/2} \sum_{i=1}^{N_n} (1-g)^n c_i^{(n)} |\phi_i^{(n)} \rangle, \quad (S3)
\]

in which case the probability of measuring the auxiliary qubits in \(|0 \rangle^{\otimes N}\) is

\[
P(N, g) = \sum_{n=0}^{N/2} \sum_{i=1}^{N_n} \left| (1-g)^2 |c_i^{(n)} \rangle \right|^2. \quad (S4)
\]

The average number of repetitions for a given number of lattice sites \( N \) and Gutzwiller parameter \( g \) is simply \( 1/P(N, g) \). As a sanity check, if we set \( g = 0 \), \(|\psi_G \rangle \) is just \(|\psi_0 \rangle\), so \( P = 1 \) and the number of repetitions is 1, as expected.

Fig. S4 shows the average number of repetitions required to prepare the GWF on quantum hardware against the normalized Hubbard parameter \( U/t \) for the Fermi-Hubbard model in chains of varying size. In all cases, the optimal Gutzwiller parameter \( g \) that minimizes the energy is considered. As is clear from the right-hand-side plot in logarithmic scale, the number of repetitions grows exponentially with the system size.

In order to estimate the number of repetitions required for larger lattices where quantum advantage may be attained (cf. Table I in main text), an extrapolation of the results obtained from simulations with chains of \( N = 2, 4, 6, 8, 10, 12 \) sites was carried out for \( U/t = 1, 5, 10, 30, 50 \), as shown in Fig. S5.

### III. SCALING OF REPETITIONS AS A FUNCTION OF TRIAL STATE FOR QPE AND QITE

The replacement of the noninteracting or self-consistent mean-field ground states by the GWF as the reference state for Quantum Phase Estimation (QPE) or Quantum Imaginary Time Evolution (QITE) could lead to significant savings of repetitions. In order to make an estimate of such savings for sufficiently large lattices where quantum advantage may be attained, an extrapolation analysis of the fidelity with respect to the exact ground state for the three reference states considered was carried out, from which the average number of repetitions due to the choice of initial state that are required to obtain the exact ground state was estimated as \( 1/|\langle \psi_{\text{trial}} |\psi_{\text{exact}} \rangle|^2 \).
FIG. S4. Average number of repetitions required to initialize GWF on quantum computer against normalized Hubbard parameter $U/t$ for the Fermi-Hubbard model in chains of varying size.

FIG. S5. Extrapolation of average number of repetitions required to initialize the GWF on quantum hardware from results of simulations for Fermi-Hubbard model in chains of $N = 2, 4, 6, 8, 10, 12$ sites for different values of $U/t$. Numerical results were fitted to exponential function $c_1 e^{c_2 N}$, with free parameters $c_1$ and $c_2$.

- QPE takes the state $|0\rangle^\otimes t \otimes |\psi^{(n)}_{\text{exact}}\rangle$ to $|\tilde{\epsilon}_n\rangle \otimes |\psi^{(n)}_{\text{exact}}\rangle$, where the $t$-qubit register corresponds to the auxiliary qubits that encode the phase, $|\psi^{(n)}_{\text{exact}}\rangle$ is the $n^{th}$ exact eigenstate of the Hamiltonian $H$, $\epsilon_n$ is the respective eigenvalue and $\tilde{\epsilon}_n$ is the estimate of $\epsilon_n$ accurate to $k$ bits of precision with probability of success $1 - s$. It then follows by linearity that, if instead the input happens to be a linear combination of the eigenstates of the Hamiltonian, as in $|0\rangle^\otimes t \otimes \sum_n \langle \psi^{(n)}_{\text{exact}} | \psi_{\text{trial}} \rangle |\psi^{(n)}_{\text{exact}}\rangle$, the probability of measuring $\epsilon_n$ accurate to $k$ bits and thus collapsing the main register onto the corresponding eigenstate is at least $|\langle \psi^{(n)}_{\text{exact}} | \psi_{\text{trial}} \rangle|^2 (1 - s)^2$.

- In QITE, the trial state can also be expanded in the eigenbasis of the Hamiltonian, in which case the action of the (normalized) imaginary-time propagator is $\mathcal{N}(\tau) e^{-\epsilon_0 \tau} \left[ \langle \psi_{\text{exact}} | \psi_{\text{trial}} \rangle |\psi^{(1)}_{\text{exact}}\rangle + e^{-(\epsilon_1 - \epsilon_0) \tau} \langle \psi^{(2)}_{\text{exact}} | \psi_{\text{trial}} \rangle |\psi^{(2)}_{\text{exact}}\rangle + \ldots \right]$, where $\mathcal{N}(\tau) = 1/\sqrt{\langle \psi_{\text{trial}} | e^{-2i H \tau} | \psi_{\text{trial}} \rangle}$ is the normalization factor. The probability of collapsing onto $|\psi_{\text{exact}}\rangle$ after imaginary time $\tau$ is $P(\tau) = \mathcal{N}(\tau)^2 e^{-2\epsilon_0 \tau} |\langle \psi_{\text{exact}} | \psi_{\text{trial}} \rangle|^2$. 
The estimation of the circuit depth only includes CNOT gates, since two-qubit gates have considerably greater execution times and error rates than single-qubit gates. This assumption is particularly relevant for quantum computers based on superconducting circuits, for which the architectures typically only include linear connections for most qubits. As for trapped-ion quantum computers, there is, in principle, the possibility of achieving all-to-all connectivity, but, at the current state of development of the hardware, this comes at the cost of lower gate fidelities and greater execution times. Of course, if these technical limitations are overcome, trapped-ion quantum computers will allow to implement this Gutzwiller projection routine with negligible circuit depth overhead.

Qiskit: an open-source SDK for working with quantum computers at the level of pulses, circuits and application modules,
https://qiskit.org/ (last checked on March 23, 2021).

49 M. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information (Cambridge University Press, 2010).
50 T. Toffoli, in Automata, Languages and Programming (Springer, 1980) pp. 632–644.