This is the accepted manuscript made available via CHORUS. The article has been published as:

**Solving strongly correlated electron models on a quantum computer**

Dave Wecker, Matthew B. Hastings, Nathan Wiebe, Bryan K. Clark, Chetan Nayak, and Matthias Troyer

Phys. Rev. A 92, 062318 — Published 10 December 2015

DOI: [10.1103/PhysRevA.92.062318](https://doi.org/10.1103/PhysRevA.92.062318)
Solving strongly correlated electron models on a quantum computer

Dave Wecker,1 Matthew B. Hastings,2,1 Nathan Wiebe,1 Bryan K. Clark,2,3 Chetan Nayak,2 and Matthias Troyer4
1 Quantum Architectures and Computation Group, Microsoft Research, Redmond, WA 98052, USA
2 Station Q, Microsoft Research, Santa Barbara, CA 93106-6105, USA
3 UIUC
4 Theoretische Physik, ETH Zurich, 8092 Zurich, Switzerland

One of the main applications of future quantum computers will be the simulation of quantum models. While the evolution of a quantum state under a Hamiltonian is straightforward (if sometimes expensive), using quantum computers to determine the ground state phase diagram of a quantum model and the properties of its phases is more involved. Using the Hubbard model as a prototypical example, we here show all the steps necessary to determine its phase diagram and ground state properties on a quantum computer. In particular, we discuss strategies for efficiently determining and preparing the ground state of the Hubbard model starting from various mean-field states with broken symmetry. We present an efficient procedure to prepare arbitrary Slater determinants as initial states and present the complete set of quantum circuits needed to evolve from these to the ground state of the Hubbard model. We show that, using efficient nesting of the various terms each time step in the evolution can be performed with just $O(N)$ gates and $O(\log N)$ circuit depth. We give explicit circuits to measure arbitrary local observables and static and dynamic correlation functions, both in the time and frequency domain. We further present efficient non-destructive approaches to measurement that avoid the need to re-prepate the ground state after each measurement and that quadratically reduce the measurement error.

I. INTRODUCTION

Feynman envisioned that quantum computers would enable the simulation of quantum systems: an initial quantum state can be unitarily evolved with a quantum computer using resources that are polynomial in the size of the system and the evolution time.1 However, it is not clear that this would enable us to determine the answers to the questions of greatest interest to condensed matter physicists. The properties most readily measured in experiments do not appear to be simple to determine in a quantum simulation, and known algorithms for quantum simulation do not return the quantities of greatest physical interest. In experiments in a solid, one can rarely, if ever, prepare a known simple initial state, nor does one know precisely the Hamiltonian with which it would evolve. Meanwhile, the ground state energy (or the energy of any eigenstate yielded by quantum phase estimation [2–4]) is not a particularly enlightening quantity and, furthermore, given a quantum state it is not clear how to determine its long-wavelength ‘universal’ properties. Consider, for instance, the following questions: given a model Hamiltonian for the electrons in a solid, such as the Hubbard model, can we determine if its ground state is superconducting? Can we determine why it is superconducting? In this paper, we show that it is possible, in principle, to answer the first question and, if it has a clear answer (in a sense to be discussed), the second question as well. Moreover, we show that it is not only possible in principle but, in fact, feasible to answer such questions with a quantum computer of moderate size.

The first step in the solution of a quantum Hamiltonian is to map its Hilbert space to the states of a quantum computer. While the Hilbert spaces of systems of bosons or spins are mapped most naturally to the states of a general-purpose quantum computer, fermionic systems can also be simulated by using a Jordan-Wigner transformation to represent them as spin systems [5, 6]; the cost can be reduced to $O(N)$ using the techniques of [7], with additional qubits [8], or using appropriate ordering techniques discussed below. In order to evolve the system on a general purpose quantum computer, we decompose the time-evolution operator according to the Trotter-Suzuki decomposition [9, 10]; the number of time intervals in such a decomposition is determined by the desired accuracy. The evolution for each time interval is expressed in terms of the available gates. If the Hamiltonian can be broken into $k$ non-commuting terms, then the gates can be broken into $k$ sets; within each set, the gates can be applied in parallel, and only $O(k)$ time steps are needed for each interval. Time evolution can be used in conjunction with the quantum phase estimation [2–4] algorithm to find an approximate energy eigenvalue and eigenstate of the Hamiltonian.

The challenge when applying this approach to electronic systems is that the Coulomb Hamiltonian in second quantized form

$$H = \sum_{p,q=1}^{N} h_{pq} c_{p}^\dagger c_{q} + \frac{1}{2} \sum_{p,q,r,s=1}^{N} h_{pqrs} c_{p}^\dagger c_{q}^\dagger c_{r} c_{s}. \tag{1}$$

has $k = O(N)$ terms for $N$ orbitals. Here the operators $c_{p}^\dagger$ and $c_{p}$ create and annihilate an electron with spin orbital $p$ (combining orbital index $i$ and spin index $\sigma$). The quadratic $h_{pq}$-terms arise from the kinetic energy of the electrons and the Coulomb potential due to the nuclei (which are assumed to be classical in the usual Born-Oppenheimer approximation) and the quartic $h_{pqrs}$-terms encode the Coulomb repulsion. In order to estimate the ground state energy of a system of inter-
acting electrons in a large molecule of $N$ orbitals, naive estimates indicated that $O(N^8)$ operations are needed [11], due to the large number of non-commuting terms in the Hamiltonian, but more recent analyses indicate that $\sim N^5$ operations may be sufficient.[12, 13] While this makes the simulation of small classically-intractable molecules feasible, the scaling is still too demanding to perform electronic structure calculations for more than a few hundred orbitals. This makes the brute-force simulation of large molecules and crystalline solids impractical. One approach to reduce the scaling complexity for the simulation of crystalline solids is to focus on effective models that capture only the relevant bands close to the Fermi energy, and simplify the $O(N^4)$ terms of the Coulomb interaction to just a few dominant terms. One of the paradigmatic effective models for strongly correlated fermions, discussed in detail in Sec. II, is the square lattice Hubbard model, which is a radical simplification of the full electronic structure Hamiltonian (1) to arguably the simplest interacting single-band model, described by the Hamiltonian

$$H_{\text{Hub}} = -\sum_{\langle i,j \rangle} \sum_{\sigma} t_{ij} \left( c_{i,\sigma}^\dagger c_{j,\sigma} + c_{j,\sigma}^\dagger c_{i,\sigma} \right) + U \sum_i n_{i,\uparrow} n_{i,\downarrow} + \sum_i \epsilon_i n_i,$$

(2)

where $n_{i,\sigma} = c_{i,\sigma}^\dagger c_{i,\sigma}$ is the local spin density and $n_i = \sum_{\sigma} n_{i,\sigma}$ is the total local density. Of all the orbitals in a unit cell, we only keep a single orbital (describing e.g. the Cu $d_{x^2-y^2}$ in a cuprate superconductor), reduce the long-range Coulomb repulsion to just the local repulsion $U$ within this orbital, and the hybridizations $t_{ij} = -\hbar p_{ij}$ to nearest neighbors on the lattice. The on-site energies will be called $\epsilon_i = \hbar \omega_i$. Usually one can consider a translationally-invariant model and then drop the indices in $t_{ij}$ and $\epsilon_i$ since all sites are equivalent. Here we will keep them explicitly since they may not all be the same during the adiabatic evolution.

The computational complexity of the Hubbard model (2) is much reduced compared to the full electronic structure problem (1), since only a single orbital is used per unit cell instead of dozens needed for the full problem and the number of terms is linear in $N$ instead of scaling with the fourth power $N^4$. On a lattice of $20 \times 20$ unit cells this means a reduction of the total number of terms from about $\sim 10^{12}$ to $\sim 10^3$, which makes such a simulation feasible on a quantum computer. Furthermore, since the 2D Hubbard Hamiltonian can be expressed as a sum of five non-commuting terms (the on-site interaction terms and four sets of hopping terms, one for each nearest neighbor to a given site) each time step in the Trotter-Suzuki decomposition requires only $\sim \log N$ parallel circuit depth (in fact, constant, if it were not for the Jordan-Wigner strings). To achieve this optimal scaling we use a Jordan-Wigner transform combined with optimal term ordering and the “nesting” and Jordan-Wigner cancellation technique of Ref. 12.

The simplifying features of the Hubbard model that afford such advantageous scaling with $N$ are the restriction to nearest-neighbor hopping and on-site interactions. Of course, a real solid will have long-ranged Coulomb interactions and longer-ranged hopping terms. However, the point of a model such as the Hubbard model is not to give a quantitatively accurate description of a solid, but rather to capture some essential features of strongly-correlated electrons in transition-metal oxides. For these purposes, a simplified model that does not include all of the complexity of a real solid should be sufficient (although, for the case of the cuprate high-$T_c$ superconductors, the required simplified model may be a bit more complicated than the Hubbard model; we focus here on the Hubbard model for illustrative purposes).

A central question in the study of a model of strongly correlated electrons is: what is its phase diagram as a function of the parameters in the model? The Hubbard model has the following interesting subquestion: is there a superconducting phase somewhere in the phase diagram? Once the phase diagram has been determined, we wish to characterize the phases in it by computing key quantities such as the quasiparticle energy gap and phase stiffness (superfluid density) in a superconducting phase. To answer these questions we will need to determine the ground state wave function for a range of interaction parameters and densities and then measure ground state correlation functions.

We take the following approach to solve the Hubbard model (or related models) on a quantum computer:

1. Adiabatically prepare an approximate ground state $|\tilde{\Psi}_0\rangle$ of the Hubbard model starting from different initial states.

2. Perform a quantum phase estimation to project the true ground state wave function $|\Psi_0\rangle$ from the approximate state $|\tilde{\Psi}_0\rangle$, and measure the ground state energy $E_0 = \langle \tilde{\Psi}_0 | H | \tilde{\Psi}_0 \rangle$.

3. Measure local observables and static and dynamic correlation functions of interest.

The general strategy for simulating quantum models has already been discussed previously, starting from Feynman’s original proposal.[1] Abrams and Lloyd [14] were the first to discuss time evolution under the Hubbard Hamiltonian, but without giving explicit quantum circuits; in Ref. 15 the same authors discuss how quantum phase estimation can be used to project from a trial state into an energy eigenstate. Refs. 5 and 16 provided more details of how to simulate quantum lattice models, proposed an algorithm to construct arbitrary Slater determinants and discussed measurements of various observables of interest. The idea of adiabatic state preparation for interacting fermion systems was proposed in Ref. 5 and adiabatic preparation has been considered by a number of authors[17–19].

In this paper we go beyond the earlier work in several crucial aspects. Besides presenting full details of the
necessary quantum circuits for time evolution (Sec. IV) and measurements (Secs. VI and VII) we address crucial issues that have not received much attention so far. In particular, in Sec. III we discuss in detail how adiabatically prepare approximate ground states \(|\Psi_0\rangle\) of the Hubbard model starting from different initial states, capturing various proposed broken symmetries, and use this to determine the phase diagram of the Hubbard model. Having thereby deduced the correct initial state, we can cheaply prepare multiple copies of the ground state of the system.

Due to the benign scaling of circuit depth with \(\log N\), the ground state of the Hubbard model on large lattices can be prepared in very short time of less than a second, even assuming logical gate times in the \(\mu s\) range as we discuss in Sec. V.

We then present, in Sec. VI, details of how to measure various quantities of interest, including both equal time and dynamic correlation functions. For the latter, we present a new approach, based on importance sampling in the frequency domain to complement the usual approach of measuring in the time domain.

Since each measurement only gives limited information, the ground state has to be prepared many times, which leads to a significant total runtime of a hypothetical quantum computer. In Sec. VII, we present new approaches to non-destructive measurements of arbitrary observables that substantially reduce the runtime. By using a relatively cheap quantum phase estimation to recreate the ground state after a measurement, we reduce the number of times the ground state needs to be prepared from scratch. Furthermore their runtime scales as \(O(\epsilon^{-1})\) with the desired accuracy \(\epsilon\) instead of the \(O(\epsilon^{-2})\) scaling of the naïve approach to measurements.

For the aficionado of quantum algorithms, the paper contains the following new algorithmic ideas. In Sec. V, we introduce a variant of quantum phase estimation that gives a two-fold reduction in the number of time steps and a four-fold reduction in rotation gates compared to the standard approach. Section III presents an approach to speeding up adiabatic state preparation by adiabatically changing the Trotter time step. In Sec. VI, we present an algorithm for measuring dynamical correlation functions by importance sampling of the spectral function in the frequency domain. Finally, Sec. VII presents two approaches to non-destructive measurements.

II. THE HUBBARD MODEL

A. Derivation of the model

The square lattice Hubbard model is one of the paradigmatic models for strongly interacting fermions. Although it was originally introduced as a model for itinerant-electron ferromagnetism [20, 21], it is now studied primarily as a model for the Mott transition and broken-symmetry (and, perhaps, even topological) ordering phenomena in the vicinity of this transition (for recent reviews, see, for instance, Refs. 22 and 23). In particular, it has been hypothesized that the square lattice Hubbard model captures the important ingredients of the cuprate high temperature superconductors. It is a radical simplification of the full Hamiltonian of a solid.

A complete description of a solid takes the form:

\[
H = \sum_i \frac{p_i^2}{2m_e} + \sum_a \frac{P_a^2}{2M_a} + \sum_{i,j} \epsilon_{ij} c_{i,a}^\dagger c_{j,a} + \frac{1}{2} \sum_{i,a} \epsilon a n_{i,a} + \sum_i U_a n_{i,a}(n_{i,a} - 1)/2 + \sum_{i,j,k,l,a,b,c,d} V_{ijkl}^{abcd} c_{i,a}^\dagger c_{j,b}^\dagger c_{k,c} c_{l,d} + \ldots
\]

where \(i, j\) label unit cells and \(a, b\) label orbitals within a unit cell. The local density in orbital \(a\) is \(n_{i,a} = \sum_i n_{i,a}\). The coupling constants have the following meanings: \(t_{ij}^{ab}\) is the matrix element for an electron to hop from orbital \(b\) in unit cell \(j\) to orbital \(a\) in unit cell \(i\) (possibly two different orbitals within the same unit cell); \(\epsilon_a\) is the energy of orbital \(a\) in isolation; \(U_a\) is the energy penalty for two electrons to occupy the same orbital in the same unit cell. This term can be rewritten equivalently as \(\sum_{i,j} U_a n_{i,a} n_{j,a}\). The prime in the summation in the final line indicates that we have omitted the term with \(i = j = k = l\) and \(a = b = c = d\), which has been separately included as the \(U_a\) term. They include, for instance, the terms with \(i = k \neq j = l\), \(a = c \neq b = d\), which is the density-density interaction \(V_{ijkl}^{abcb} n_{i,a} n_{j,b}\), and the terms with \(j = l\), \(b = d\), which is correlated hopping term \(V_{ijkl}^{abcb} n_{i,a} c_{k,c} n_{j,b}\). These terms will be assumed to be much smaller than \(t_{ij}^{ab}\), \(U_a\) and will be dropped.

If the differences between the \(\epsilon_a\)s are the largest energy scales in the problem, then there will be, at most, a single
partially-filled orbital and we can ignore all of the other orbitals when studying the physics of the model at temperatures less than $|e_i - e_j|$. These energy differences can be on the order of eV and, therefore, such an approximation will be valid over a very large range of temperatures and energies. Similarly, if $t_{\alpha\beta}^{ab}$ is a large energy scale, it may be possible to reduce the model to one with a single orbital per unit cell (which, in this case, will be a linear combination of the orbitals in Eq. (4)). If the on-site interactions $U$ in this orbital (we drop the subscript $a$ since there is only a single orbital under consideration) are much larger than the interactions with nearby unit cells (nearest-neighbor, next-nearest neighbor, etc.), then we can drop the latter. Similarly, if the nearest-neighbor hopping is much larger than more distant hopping matrix elements, then we can drop the latter and write down the simplified model of Eq. (2). This is the Hubbard model.

In the case of the copper-oxide superconductors, the only orbital that is retained is the Cu $d_{x^2−y^2}$ orbital. (However, it has also been argued that a 3-band model including O $p_x$ and $p_y$ orbitals is necessary to capture the physics of the cuprates.[24]) Usually one can drop the indices in $t_{ij}, U_i$ and $\epsilon_i$ since all sites are equivalent, but we will keep them here since during the adiabatic limit, where both scales are comparable. [27] When the subleading $O(t^3/U^2)$ terms are kept, the system remains insulating, but the spin physics may change considerably.

The situation is even less clear when the system is “doped”, i.e. when the density is changed from half-filling. Then, an effective model can be derived to lowest-order in $t/U$. The model is governed by the $t − J$ Hamiltonian:

$$H = - \sum_{i,j} t_{ij} \sum_\sigma \left( c_i^\dagger \sigma c_{j,\sigma} + \text{h.c.} \right) + J \sum_{\langle i,j \rangle} S_i \cdot S_j + O(t^3/U^2)$$

which together with a no-double-occupancy constraint. As a result of this constraint, the $t − J$ model has a smaller per site Hilbert space than the Hubbard model, which makes it a much more attractive target for exact diagonalization. However, the $t − J$ model does not capture all of the physics of the Hubbard model (especially in the regime in which $t/U$ is small but not so small that higher-order in $t/U$ terms can be neglected). Any deviation from half-filling will cause the conductivity to be non-zero in a completely clean system, in which $e_i$ is independent of $i$ in Eq. (2). However, in any real material, there will be some disorder due to impurities, and the system will remain insulating for densities very close to half-filling. As the density deviates from half-filling, antiferromagnetic order is suppressed and eventually disappears. If the Hubbard model captures the physics of the cuprate superconductors, then superconducting order[30] with $d_{x^2−y^2}$ pairing symmetry must occur for a density of $1 − x$ electrons per unit cell, with $x$ in the range $0.05 \leq x \leq 0.25$. Moreover, on the under-doped side of the phase diagram, $0.05 \leq x \leq 0.15$, other symmetry-breaking order parameters, such as stripe order, $d$-density wave order, or antiferromagnetic order may also develop (for recent reviews, see, for instance, Refs. 22 and 23; see, also, Ref. 31 for a recent calculation finding several competing orders). On the other hand, if the electrons on neighboring sites. Since an electron can only hop to a neighboring site if its spin forms a singlet with the spin of the electron there (due to the Pauli principle), the energy is lowered by such processes if neighboring spins are antiferromagnetically-correlated. We can derive an effective Hamiltonian at energies much lower than $U$ by taking into account these virtual hopping processes. It takes the form:

$$H = J \sum_{\langle i,j \rangle} S_i \cdot S_j + O(t^3/U^2)$$

B. The physics of the Hubbard model

At half-filling (one electron per unit cell) the Hubbard model gives a simple account of Mott insulating behavior. [25] For $U = 0$, the ground state is metallic. There is a single band and it is half-filled; the Fermi surface is the diamond $|k_x \pm k_y| = \pi/a$, where $a$ is the lattice constant. However, for $U > 0$, the ground state is insulating. While this is true for all $U > 0$, the physics is different in the small $U \ll t$ and $U \gg t$ limits. For $U \ll t$, the system lowers its energy by developing antiferromagnetic order, as a result of which the unit cell doubles in size. There are now two electrons per unit cell and the system is effectively a band insulator. The antiferromagnetic moment $N$ and the charge gap $\Delta_{\text{charge}}$ are both exponentially-small $N \sim \Delta_{\text{charge}} \sim t e^{-c t/U}$ for some constant $c$. For large $U$, i.e. $U \gg t$, on the other hand $\Delta_{\text{charge}} \sim U$. In this limit, it is instructive to expand about $t = 0$. Then, the ground state is necessarily insulating; there is one electron on each site and an energy penalty of $U$ to move any electron since some site would have to be doubly-occupied. For small but non-zero $t/U$, this picture is dressed by small fluctuations in which an electron makes virtual hops to a neighboring site and then returns. As a result of these fluctuations, there is an effective interaction between the spins of the

$$H = J \sum_{\langle i,j \rangle} S_i \cdot S_j + O(t^3/U^2)$$
Hubbard model is not superconducting in this doping range, then superconductivity in the cuprates must be due to effects that are missing in the Hubbard model, such as longer-ranged interactions, longer-ranged hopping, interlayer coupling, and phonons.

Thus, as far as the cuprates are concerned, the principle aim of an analysis of the Hubbard model is to determine if it has a $d_{x^2-y^2}$ superconducting ground state for $U/t \sim 8$ over the range of dopings $0.05 \leq x \leq 0.25$. If the answer is in the affirmative, then various properties of this superconducting ground state can be studied. For instance, the gap and its dependence on the doping, $\Delta(x)$, can be compared to experimental measurements. If the Hubbard model does not superconduct, then different models must be considered, restoring terms that have been dropped in the passage to the Hubbard model.

III. ADIABATIC PREPARATION OF THE GROUND STATE

We start the adiabatic preparation from the ground state of some Hamiltonian whose solution is known. For the sake of concreteness, we consider two examples. The first, in Sec. III A, is the quadratic mean-field Hamiltonian whose exact ground state is the BCS ground state for a $d$-wave superconductor or the analogous Hamiltonian for any other ordered state such as striped [32], antiferromagnetic [33], $d$-density wave [34], etc. states. The second, discussed in Sec. III B is the Hubbard model with hopping matrix elements tuned so that the system breaks into disconnected $2 \times 2$ plaquettes. In either case, the ground state is known and there is a gap to all excited states.

An essential point is that these initial states can be efficiently prepared on a quantum computer. Our approach to preparing a Slater determinant, discussed in Sec. III A 2, is both deterministic and has better scaling than the algorithm previously proposed in Ref. 5.

We then adiabatically evolve the Hamiltonian into the Hamiltonian of the Hubbard model with an additional weak symmetry-breaking term (to give a gap to Goldstone modes, a subtlety that is discussed in Section III A). In Sec. IV we review the quantum circuit used to implement time evolution of the Hubbard model and show how it can be done with $O(N)$ gates and a parallel circuit depth of only $O(\log N)$ operations per time step. This step is formally similar to adiabatic quantum optimization [17], but with a different viewpoint. If no phase transition occurs along this adiabatic evolution, then we conclude that we guessed correctly about the phase of the system (for this value of parameters in the final Hubbard model). However, if a phase transition does occur, then the gap will close and we will know that we guessed incorrectly and can repeat the procedure with a different initial state in a different phase. Through a process of elimination, we determine the phase of the system for a given set of parameters and, by repeating this process for different Hubbard model parameters, we map out the phase diagram. In adiabatic quantum optimization [17], it is assumed that a phase transition occurs and it is hoped that the minimum energy at the transition point scales polynomially in system size rather than exponentially, thereby allowing the evolution to occur in polynomial time. In our protocol for solving the Hubbard model, however, any gap closing, either polynomial or exponential – corresponding, respectively, to a second- or first-order phase transition[35] – is an invitation to repeat the procedure with a different initial state and correspondingly different annealing path until no gap closing occurs. When the initial state is in the same phase as the ground state of the Hubbard model, the adiabatic evolution can be done in constant time up to subpolynomial factors (see V C for a more detailed discussion of errors in adiabatic evolution; while in most of the paper we treat the time required for adiabatic evolution as scaling with inverse gap squared, in that section we provide a more precise discussion of diabatic errors (i.e., those transitions out of the ground state that are due to a finite time scale for the evolution) and further show how it is possible to improve the scaling with gap by using smoother annealing paths. We hope that one of the mean-field states described above will be in the correct phase; if none are, this would be very surprising physical information to learn about the Hubbard model.

A. Adiabatic Evolution from Mean-Field States

1. Mean-Field Hamiltonians

Let us suppose that we wish to test whether the ground state of the 2D Hubbard model is superconducting for some values of $t, U$ and density. We can do this by seeing if the Hubbard model is adiabatically connected to some simple superconducting Hamiltonian. It need not be realistic; it merely needs to be a representative Hamiltonian for a superconductor. We can take the BCS mean-field Hamiltonian:

$$H_{DSC}^{MF} = -\sum_{\langle i,j \rangle} \sum_{\sigma} t_{ij} \left( c_{i,\sigma}^\dagger c_{j,\sigma} + c_{j,\sigma}^\dagger c_{i,\sigma} \right)$$

$$- \sum_{\langle i,j \rangle} \Delta_{ij}^{x^2-y^2} \left( c_{i,\uparrow}^\dagger c_{j,\downarrow} - c_{i,\downarrow}^\dagger c_{j,\uparrow} \right) + \text{h.c.},$$

(7)

where $\Delta_{ij}^{x^2-y^2} = \Delta/2$ for $i = j \pm \hat{x}$ and $\Delta_{ij}^{x^2-y^2} = -\Delta/2$ for $i = j \pm \hat{y}$. This is a mean-field $d_{x^2-y^2}$ superconductor (DSC) with fixed, non-dynamical superconducting gap of magnitude $\Delta$.

The ground state of this Hamiltonian can be written in the form of a Slater determinant by making a staggered particle-hole transformation on the down-spins: $c_{\downarrow i}^\dagger \rightarrow$
Then the Hamiltonian takes the form:
\[
H_{\text{DSC}}^{\text{MF}} = -\sum_{\langle i,j \rangle} t_{ij} \left( c_{i\uparrow}^\dagger c_{j\downarrow}^\sigma + c_{i\downarrow}^\sigma c_{j\uparrow}^\dagger \right) - \sum_{\langle i,j \rangle} (-1)^j \Delta^{x^2-y^2}_{ij} \left( c_{i\uparrow}^\dagger c_{j\downarrow}^\dagger - c_{j\uparrow}^\dagger c_{i\downarrow}^\dagger \right) + \text{h.c.},
\]
Under this transformation, the number of up-spin electrons per unit cell is particle-hole transformed: \( n_{i\downarrow} \rightarrow 1 - n_{i\uparrow} \) while the number of up-spin electrons is unchanged; equivalently, the chemical potential for up-spin electrons is flipped in sign, \( \mu_{\uparrow} = -\mu_{\downarrow} \). Consequently, the number of down-spin electrons is not equal to the number of up-spin electrons. In addition, the superconducting pair field \( \Delta^{x^2-y^2}_{ij} \) has become transformed into a staggered magnetic field in the \( S_z \) direction with a \( d_{x^2-y^2} \) form factor (a \( d \)-wave spin-density wave, in the terminology of Ref. 34). Thus, the ground state of this quadratic Hamiltonian is a Slater determinant of the form:
\[
|\Psi\rangle = \prod_k (u_k c_{k\uparrow}^\dagger + v_k c_{k\downarrow}^\dagger + Q_{k\downarrow})|0\rangle
\]
where the momenta \( k \) range over the Brillouin zone, \( -\frac{\pi}{a} \leq k_x, k_y \leq \frac{\pi}{a} \) where \( a \) is the lattice constant, and the functions \( u_k, v_k \) are given by:
\[
u_k^2 = \frac{1}{2} \left( 1 - \frac{\epsilon_k}{\sqrt{\epsilon_k^2 + |\Delta_k|^2}} \right)
\]
\[
v_k^2 = \frac{1}{2} \left( 1 + \frac{\epsilon_k}{\sqrt{\epsilon_k^2 + |\Delta_k|^2}} \right)
\]
where \( \epsilon_k \equiv -2t(\cos k_x a + \cos k_y a) \) and \( \Delta_k \equiv \Delta(\cos k_x a - \cos k_y a) \). In Sec. III A 2, we show how to prepare this ground state. For now, we will simply assume that this ground state can be prepared efficiently. In computing the properties of this state, it is important to keep in mind that physical operators have been transformed according to \( c_{i\uparrow}^\dagger \rightarrow (-1)^i c_{i\downarrow}^\dagger \) and to use the appropriate transformed operators.

We now adiabatically deform this Hamiltonian into a weakly-perturbed Hubbard model. There are two reasons why our final Hamiltonian is a weakly-perturbed Hubbard model, rather than the Hubbard model itself: (1) in the absence of long-ranged Coulomb interactions, \( \langle \hat{n}_{i\downarrow} \rangle \), which are neglected in the Hubbard model) a superconducting ground state will have gapless Goldstone excitations, and (2) a \( d_{x^2-y^2} \) superconductor has gapless fermionic excitations at the nodes of the superconducting gap. Although these gapless excitations are important for the phenomenology of superconductors (e.g. the temperature dependence of the superfluid density and the structure of vortices), they are a nuisance in the present context. Hence, we weakly perturb the Hubbard model in order to give energy gaps to Goldstone modes and nodal fermions. However, this can be done weakly so that we do not bias the tendency towards superconductivity (or lack thereof). If the Hubbard model superconducts, then the scale of the \( d_{x^2-y^2} \) superconducting gap should be \( \sim xJ \), where \( \langle \hat{n}_i \rangle = 1 \) is the average occupancy of a site and \( J = 4t^2/U \) is the superexchange interaction. (On the other hand, if the Hubbard model does not superconduct, then the superconducting gap in the cuprates must be determined by some other energy scale.) Hence, we apply a fixed \( d_{x^2-y^2} \) superconducting gap that is much smaller than \( xJ \) in order to pin the phase of any superconducting order that may develop. In addition, we apply a small imaginary \( d_{xy} \) superconducting gap to give a small gap to the nodal fermions. Neither gap scales with system size.

More concretely, we adiabatically evolve the Hamiltonian according to:
\[
H_{\text{DSC}}(s) = (1 - s)H_{\text{DSC}}^{\text{MF}} + sH_{\text{Hub}} - \sum_{\langle i,j \rangle} u \Delta^{x^2-y^2}_{ij} \left( c_{i\uparrow}^\dagger c_{j\downarrow}^\dagger - c_{j\uparrow}^\dagger c_{i\downarrow}^\dagger \right) + \text{h.c.} - \sum_{\langle (j,k) \rangle} i \Delta^{xy}_{jk} \left( c_{j\uparrow}^\dagger c_{k\downarrow}^\dagger - c_{k\uparrow}^\dagger c_{j\downarrow}^\dagger \right) + \text{h.c.}
\]
A similar strategy can be used to test whether the Hubbard model has any other ordered ground state, such as antiferromagnetic order. In this case, our starting Hamiltonian is, instead:

\[ H_{AF}^{MF} = -\sum_{(i,j)} t_{ij} (c_{i,\sigma}^\dagger c_{j,\sigma} + c_{j,\sigma}^\dagger c_{i,\sigma}) \]
\[-\sum_i (-1)^i N (c_{i\uparrow}^\dagger c_{i\uparrow} - c_{i\downarrow}^\dagger c_{i\downarrow}), \]

(12)

We then evolve along a path analogous to Eq. (11), but with the symmetry-breaking terms in the last two lines replaced by a weak staggered magnetic field. If the Hubbard model has an antiferromagnetic ground state (for some values of \( t, U, x \)) then the gap will not close along this path. However, if the actual ground state of the Hubbard model is superconducting for these values of the couplings, then the gap will become \( O(1/N^{1/2}) \) where \( z \) is the dynamical critical exponent of the transition (i.e. the gap will close, up to finite-size effects) at the point along the evolution at which superconductivity develops. There will be a non-zero staggered magnetization throughout since we never fully remove the staggered magnetic field, but there will be a sharp signature at the point at which superconductivity develops.[35] The presence of a staggered field, which is kept small, may slightly shift the onset of superconductivity but will have no other effect.

This strategy can allow us, in some of the circumstances in which it has a clear answer, to address the question: why does the 2D Hubbard model have a superconducting ground state? In particular, it is possible that the superconducting ground state of the Hubbard model necessarily has some other type of order coexisting with superconducting order, especially in the underdoped regime. These secondary orders play a role in some theories of the cuprate superconductors. If such an order were present in the Hubbard model, evolution from a purely superconducting initial state would necessarily encounter a phase transition (at the point of onset of this additional order). Therefore, we can determine not only if the Hubbard model has a superconducting ground state, but also whether the ground state exhibits a secondary type of long-ranged order over some range of dopings. We will return to the question “why does the Hubbard model superconduct” when we discuss correlation functions.

The virtue of using these mean field Hamiltonians as starting points for adiabatic evolution is that their physics is fully understood. In particular, their ground states are Slater determinants, which we can efficiently prepare, as we explain in the next section.

2. Preparing Slater Determinants

In this section, we explain an efficient, simple quantum algorithm to prepare Slater determinants. This is then used to prepare the Hubbard ground state by adiabatic evolution from \( U = 0 \) to \( U \neq 0 \).

The standard algorithm is in Ref. 5. This algorithm suffers from some drawbacks. Given a projector \( \rho \), a Slater determinant state is a state

\[ \Psi_{SD}(\rho) = \Pi_{j=1}^{N_e} b_j^\dagger |0\rangle, \]

(13)

where \( |0\rangle \) is the no particle vacuum state, and each operator \( b_j^\dagger \) is a linear combination of \( a_l^\dagger \), where the \( a_l^\dagger \) are fermion creation operators on a given site \( i \). We write \( b_j^\dagger = \sum_{l=1}^{N_e} a_l^\dagger P_{lj} \), where \( P_{lj} \) are complex scalars and are the entries of \( P \), \( n \) is the number of orbitals, and \( N_e \) is the number of electrons. Given any matrix \( \rho \), we can find a matrix \( P \) such that

\[ \rho = PP^\dagger. \]

(14)

Then, using this matrix \( P \) in the definition of the state gives the state \( \Psi_{SD}(\rho) \); note that this definition fixes \( \Psi_{SD}(\rho) \) up to a complex phase. The algorithm in Ref. 5 proceeds by noting that \( U_m = \exp(i \pi (b_m + b_m^\dagger)) \), acting on the vacuum, produces the state \( b_m^\dagger |0\rangle \), up to a phase. Therefore, if the \( b_j^\dagger \) are suitably orthogonalized, successively acting with the unitary \( U_m \) produces the desired state. However, no explicit method for implementing this unitary \( U_m \) is given; if the unitary is implemented by a Trotter method, this will be extremely inefficient. We explain a simpler technique.

Let us first assume that \( \rho \) is real; the extension to the complex Hermitian case will be straightforward. The idea is first to prepare the Slater determinant state

\[ \Psi_{SD}(\rho_0) = \Pi_{j=1}^{N_e} a_j^\dagger |0\rangle. \]

(15)

This corresponds to the case that \( \rho \) is an \( n \)-by-\( n \) diagonal matrix, with ones in the first \( N_e \) entries and zeroes elsewhere. This state has one electron on each of the first \( N_e \) states. It is a product state that can be prepared in linear time (indeed, simply initialize each of the first \( N_e \) spins to up and the other spins to down). Then, we act on this state \( \Psi_{SD}(\rho) \) with a sequence of Givens rotations. A Givens rotation, to borrow the language of matrix diagonalization, is a rotation matrix that acts on two rows or columns at a time. For any given pair, \( i, j \), define the operator \( R_{i,j}(\theta) \) by

\[ R_{i,j}(\theta) = \exp(\theta c_j^\dagger c_i - h.c.). \]

(16)

This is a simple two-qubit gate (up to Jordan-Wigner strings) and is, therefore, more straightforwardly implemented than the unitary \( U_m \) in the standard method described above. Define the \( n \)-by-\( n \) orthogonal matrix \( r_{i,j}(\theta) \) by its matrix elements as follows. Let \( \langle r_{i,j}(\theta) \rangle_{k,k} = 1 \) for \( k \neq i, j \) and let \( \langle r_{i,j}(\theta) \rangle_{k,l} = 0 \) if \( k \neq l \) and \( l \neq i, j \) or if \( k \neq l \) and \( l \neq i, j \). Then, let \( \langle r_{i,j}(\theta) \rangle_{i,i} = \langle r_{i,j}(\theta) \rangle_{j,j} = \cos(\theta) \) while \( \langle r_{i,j}(\theta) \rangle_{i,j} = -\langle r_{i,j}(\theta) \rangle_{j,i} = \sin(\theta) \). Then,

\[ R_{i,j}(\theta) \Psi_{SD}(\rho) = \Psi_{SD}(r_{i,j}(\theta) \rho r_{i,j}(\theta)^{-1}), \]

(17)

where again the equality holds up to the phase ambiguity.
Now, given some desired target $\Psi_{SD}(\rho)$, we claim that one can always find a sequence of at most $nN_e$ Givens rotations that will turn the matrix $\rho$ into the matrix $\rho_0$. Thus, by inverting this sequence of Givens rotations and applying that sequence to the state $\Psi_{SD}(\rho_0)$, we succeed in constructing the desired state. The Givens rotations will be ordered such that the Jordan-Wigner strings can be cancelled as in Ref. 12, so the algorithm takes time $O(N_e n)$. To show our claim that the sequence of Givens rotations exists, note that the effect of a given rotation which transforms $\rho$ to $r_{i,j}(\theta)\rho_1 r_{i,j}^{-1}$ can be achieved by transforming $P$ to $r_{i,j}(\theta)P$. Consider a left singular vector of $P$ with singular value equal to 1; we can find a sequence of at most $n$ Givens rotations that rotate that vector to be equal to the vector $(1,0,...)$. As a result, all the other singular vectors now have amplitude only on sites 2,...,$n$. Find another sequence to transform some other left singular vector with singular value equal to 1 into the vector $(0,1,0,...)$. Continue until all singular vectors with nonzero singular values have amplitudes only on the first $N_e$ sites.

The extension of this procedure to the complex case is simple. Rather than having the Givens rotation $R_{i,j}(\theta)$ depends only a single angle $\theta$, we need to perform a rotation $\exp(2\pi e^{-1})$ for some complex number $z$. All the time estimates remain unchanged up to constant factors.

However, for many systems a much shorter sequence of Givens rotations can be found. Using the same idea as in the fast Fourier transform, if $n$ is a power of 2, there exists a sequence of $n \log_2(n)$ Givens rotations to transform an $n$-by-$n$ matrix to the momentum basis. Thus, we can produce ground states of a system of free fermions on a periodic lattice in time $O(n^2 \log(n))$. This can be easily extended to handle the case of a periodic system with a unit cell larger than a single site. Assume there are states $|i,\vec{x}\rangle$ where $i$ labels an atom in a unit cell and $\vec{x}$ labels a lattice vector. We transform to a momentum basis $|i,\vec{k}\rangle$, and then for each $\vec{k}$ we initialize some Slater determinant in the unit cell in momentum space. If the unit cells have size $O(1)$, the time to initialize each cell is $O(1)$ so the total time is still $O(n^2 \log(n))$.

A further possible optimization is that if $N_e > n/2$, we can instead use Givens rotations to bring the holes to the desired states, rather than the electrons, taking a time $O((n - N_e)n)$ rather than $O(N_e n)$. As an interesting application of this, consider creating the ground state of $N_e = 3$ electrons in $n = 4$ sites, using a hopping Hamiltonian with the sites arranged on a ring. This can be done by initializing a state with electrons on the first 3 sites and then apply a total of three different $R_{i,j}(\theta)$ operators.

3. Numerical Results

We applied the Slater determinant preparation procedure to the Hubbard model on 8 sites arranged in two rows of 4 sites. We started at $U = 0$ and $\epsilon_i = 0$. All horizontal couplings were set equal to 1, while to avoid having an exact ground state degeneracy the vertical couplings were doubled in strength to 2 (note that this requirement depends on the filling factor; for example, at half-filling we would instead prefer to remain at vertical coupling equal to 1 to avoid degeneracies). The ground state of this model is a Slater determinant of free fermion single-particle wavefunctions that we could prepare with a total of 14 Givens rotations for both spin up and spin down. This was done by initializing particles in three out of four sites in the top row of sites, and holes in the remaining sites. Then, 3 Givens rotations were used to create a uniform superposition of the hole in various states in the top row. Finally, 4 Givens rotations were used to create a uniform superposition between top and bottom rows.

We then annealed to $U \neq 0$ while reducing the vertical coupling to unit strength. The results are shown in Fig. 1. As seen, increasing the annealing time increases the success probability, until it converges to 1. The annealing times required to achieve substantial (say, 90%) overlap with the ground state are not significantly different from the preparation approach discussed next based on “joining” plaquettes. While the 8 site results cannot be directly extrapolated, we have a variety of methods to infer the behavior of larger systems, including free fermions, studied later in the paper, and scaling theory.

Note that for this particular model, we know the ground state energy with very high accuracy from running an exact diagonalization routine so we know whether or not we succeed in projecting onto it after phase estimation. The system size is small enough that calculating ground state energies is easy to accomplish on a classical computer using Lanczos methods. Of course, on larger systems requiring a quantum computer to simulate, we will not have access to the exact energy. In this case, it will be necessary to do multiple simulations with different mean-field starting states as described at the start of this section; then, if the annealing is sufficiently slow and sufficiently many samples are taken, the minimum over these simulations will give a good estimate of the ground state energy. Once we have determined the optimum mean-field starting state and annealing path, we can then use this path to re-create the ground state to determine correlation functions.

B. Preparing of $d$-wave resonating valence bond states from plaquettes

1. $d$-wave resonating valence bond states

An alternative approach to adiabatic state preparation has previously been proposed in the context of ultracold quantum gases in optical lattices. In this approach, $d$-wave resonating valence bond (RVB) states are prepared. [18] RVB states are spin liquid states of a Mott insulator, in which the spins on nearby sites are
paired into short-range singlets. RVB states were first introduced as proposed wave functions for spin liquid ground states of Mott insulators, and then conjectured to describe the insulating parent compounds of the cuprate superconductors. In more modern language, they are described as $\mathbb{Z}_2$ topologically-ordered spin liquid states. The idea was that, upon doping, the singlet pairs would begin to move and form a condensate of Cooper pairs. Although the insulating parent compounds of the cuprates are antiferromagnetically-ordered, it is nevertheless possible that the cuprate superconductors are close to an RVB spin liquid ground state and are best understood as doped spin liquids. The RVB scenario has been confirmed for $t$-$J$ and Hubbard models of coupled plaquettes and ladders (consisting of two coupled chains) and remains a promising candidates for the ground state of the Hubbard model and high-temperature superconductors.

Instead of starting from mean-field Hamiltonians, we build up the ground state wave function from small spatial motifs, in this case four-site plaquettes, which are the smallest lattice units on which electrons pair in the Hubbard model. Following the same approach as originally proposed for ultracold quantum gases, we prepare four-site plaquettes in their ground state, each filled with either two or four electrons. These plaquettes then get coupled, either straight to a two-dimensional square lattice, or first to quasi-one dimensional ladders, which are subsequently coupled to form a square lattice.

FIG. 1. Success probability as a function of total annealing time for an annealing path in the Hubbard model that begins at $U = 0$ and vertical hopping equal to 2 and ends at $U = 2.5t$ (blue) or $U = 4t$ (red) and all hoppings equal to 1. The initial state is a Slater determinant state (of free fermion single-particle states). Annealing is linear along the path. The x-axis represents the total time for the anneal. The y-axis represents the probability that the phase estimation routine returns an energy that is within $10^{-3}$ of the ground state energy. This probability is obtained by sampling runs of the phase estimation circuit, giving rise to some noise in the curves.

The first step is preparing the ground state of the Hubbard model on 4-sites plaquettes filled with either two or four electrons. We start from very simple product states and adiabatically evolve them into the ground states of the plaquettes.

To prepare the ground states of plaquettes with four electrons we start from a simple product state $|\psi_0\rangle = \prod_{i,j} c_{i,j}^\dagger \prod_{i,j} c_{i,j} |0\rangle$, using the labeling shown in Fig. 2. Our initial Hamiltonian, of which this state is the ground state has no hopping ($t_{ij} = 0$), but already includes the Hubbard repulsion on all sites. To make the state with doubly occupied state the ground state we add large on-site potentials $\epsilon_2 = \epsilon_3 = 2U + 3t$ on the empty sites. To prepare the ground state of the plaquette we then

1. start with $t_{ij} = 0$, $U_i = U$ and a large potential on the empty sites (here $\epsilon_2 = \epsilon_3 = \epsilon$) so that the initial state is the ground state,
2. ramp up the hoppings $t_{ij}$ from 0 to $t$ during a time $T_1$,
3. ramp the potentials $\epsilon_i$ down to 0 during a time $T_2$.

Time scales $T_1 = T_2 \approx 10t^{-1}$ and $\epsilon = U + 4t$ are sufficient to achieve high fidelity. As we discuss below, it can be better to prepare the state quickly and then project into the ground state through a quantum phase estimation than to aim for an extremely high fidelity in the adiabatic state preparation. To prepare a plaquette with two electrons we start from the product state $|\psi_0\rangle = \prod_{i,j} c_{i,j}^\dagger \prod_{i,j} c_{i,j} |0\rangle$ and use the same schedule, except that we choose the non-zero potentials on three sites of the plaquette ($\epsilon_1 = \epsilon_2 = \epsilon_3 = \epsilon$).

3. Coupling of plaquettes

After preparing an array of decoupled plaquettes, some of them with four electrons and some with two electrons depending on the desired doping, we adiabatically turn on the coupling between the plaquettes. As a test case we use that of Ref. 18 and couple two plaquettes (see Fig. 3), one of them prepared with two electrons and one with four electrons.

Naively coupling plaquettes by adiabatically turning on the intra-plaquette couplings $t_{24}$ and $t_{35}$ will not give
FIG. 3. Coupling of two plaquettes

FIG. 4. Probability of preparing ground state as a function of times $T_1$, $T_2$, $T_3$ in units of $t^{-1}$, following the annealing schedule described above. During time $T_3$, a uniform potential $\epsilon$ was ramped from $t$ to 0 on the plaquette with two electrons to break the symmetry between plaquettes.

the desired result, since the Hamiltonian is reflection symmetric but an initial state with two electrons on one plaquette and four on the other breaks this symmetry. No matter how slowly the anneal is done, the probability of preparing the ground state will not converge to 1. We thus need to explicitly break the reflection symmetry by either first ramping up a small chemical potential on a subset of the sites or – more easily – not completely ramping down the non-zero $\epsilon_i$ values when preparing plaquette ground states. Consistent with Ref. 18 we find that a time $T_3 \approx 50t^{-1}$ is sufficient to prepare the ground state with high fidelity.

FIG. 5. Energy levels during annealing. $y$-axis is energy. $x$-axis is time. From time 0 to time 16 (in units of $t^{-1}$) the hopping is ramping up, time 16 to time 36 the non-uniform $\epsilon$ in each plaquette is ramping down (leaving a uniform $\epsilon = 1$ on the plaquette with two electrons), time 36 to time 76 is joining plaquettes. All levels remain non-degenerate except for a degeneracy that appears at the end of the anneal. After time 76 plaquettes are separated, as described below.

4. Decoupling plaquettes

As proposed in the context of analog quantum simulation, the pairing of holes on plaquettes can be tested by adiabatically decoupling plaquettes.[18] We start by preparing one plaquette with four electrons and a second plaquette with two electrons. We then couple them as described above to end up with a system with two holes relative to the half filled Mott insulator. After adiabatically decoupling the two plaquettes again we measure the number of electrons on each plaquette. If holes bind in the ground state, they will end up on the same plaquette, and we measure an even electron number per plaquette. If they don’t bind each hole will prefer to be on a different plaquette to maximize its kinetic energy and we will end up with three electrons per plaquette.

As shown in Fig. 6, the probability of finding three electrons on a plaquette goes to zero for $U = 2t$ and $U = 4t$ when increasing the time $T_s$ over which we ramp down the inter-plaquette hopping. This indicates pairing and is consistent with the known critical value of $U \approx 4.5t$ for pairing on four-site plaquettes. On the other hand for $U = 6t$ and $U = 8t$ we see an increase in the probability of finding three electrons per plaquette, indicating the
absence of pairing.

C. Trotter Errors and Accelerating adiabatic preparation by using larger time steps

In practice, it suffices for our annealing preparation to obtain a state with sufficiently large overlap onto the ground state that phase estimation will then project onto the ground state with a reasonably large probability. We can thus tolerate some errors in our annealing path. For this reason, we annealed using a larger time step, before doing the phase estimation with a smaller time step.

In fact, this approach of annealing at a larger time step can be turned into an approach which uses a larger time step but still produces a state with very large overlap with the ground state. Note that the larger Trotter-Suzuki step means that we in fact evolve under a modified Hamiltonian, proportional to the logarithm of the unitary describing a single time step. As the time step decreases, this Hamiltonian becomes closer to the desired Hamiltonian. It is then possible to evolve with a larger time step, following the adiabatic path in parameter space, and then at the end of the path one can “anneal in the time step”, gradually reducing the time step to zero.

In our simulation of the anneal using joining two plaquettes, the system was rather insensitive to Trotter errors. We used a relatively large time step, equal to $0.25t^{-1}$ for the annealing, and used a much smaller time step for the phase estimation that made the errors negligible there.

A time step of $0.05t^{-1}$ for phase estimation yielded relative errors below $10^{-3}$ and $0.003t^{-1}$ yielded relative errors below $10^{-5}$. The larger time step used for the adiabatic preparation prevents us from obtaining $100\%$ overlap with the ground state even in the limits of very long annealing times; however, as seen we still obtain very high overlaps in the range of $80\% - 96\%$.

### IV. IMPLEMENTING UNITARY TIME EVOLUTION OF THE HUBBARD MODEL

A. Circuits for the individual terms

For the Hubbard model, we need to implement unitary evolution under two types of terms, hopping terms $c_{p,\sigma}c_{q,\sigma'}^\dagger$ and repulsion terms $c_{p,\sigma}^\dagger c_{p,\sigma'}^\dagger c_{q,\sigma}c_{q,\sigma'}^\dagger$. These terms are a subset of those needed for earlier electronic structure calculations[6]. In addition we will need to implement unitary evolution under a pairing term $c_{p,\sigma}^\dagger c_{q,\sigma'}^\dagger + h.c.$.

To establish notation we summarize in Table I the gates used in the quantum circuits shown in this paper.

| Name                        | Symbol | Matrix Representation |
|-----------------------------|--------|-----------------------|
| Hadamard                    | $\{H\}$ | $\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$ |
| $Y$ basis change            | $Y$    | $\begin{bmatrix} 1/\sqrt{2} & i/\sqrt{2} \\ i/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ |
| Phase gate                  | $T(\theta)$ | $\begin{bmatrix} 1 & 0 \\ 0 & e^{-i\theta} \end{bmatrix}$ |
| $Z$ rotation                | $R_z(\theta)$ | $\begin{bmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{bmatrix}$ |
| controlled not (CNOT)       |        | $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ |

1. The repulsion and chemical potential terms

The repulsion term is an example of what is called an $H_{ppq}$ term and evolution under this term is given in Fig. 7. A chemical potential term, $c_{p,\sigma}^\dagger c_{p,\sigma}$, which we call an $H_{pp}$ term, is not needed for the final Hamiltonian but is useful for annealing. The circuit for this term is shown in Fig. 8. Since all the repulsion terms in the Hubbard model and all the chemical potential terms commute with each other, they can be applied in parallel, thus needing
FIG. 8. Quantum circuit to implement time evolution for a time step $\theta$ under the term $H_{pqqp} = n_{p,\sigma}n_{q,\sigma'} = c_{p,\sigma}^{\dagger}c_{q,\sigma'}^{\dagger}c_{q,\sigma'}c_{p,\sigma}$. Top and bottom lines represent qubits for $p, \sigma$ and $q, \sigma'$. Only $O(N)$ gates and $O(1)$ parallel circuit depth.

2. The hopping and pairing terms

The unitary evolution under the hopping term, that we call an $H_{pq}$ term, is given by the circuit in Fig. 9, and has also been considered previously. It also will be useful to be able to implement evolution under a superconducting pairing term, $\Delta c_{p,\sigma}^{\dagger}c_{q,\sigma'}^{\dagger} + h.c.$ This term is similar to the hopping term, since both are quadratic in the fermionic operators. The hopping and pairing term circuits will be very similar to each other, with the changes involving changing some of the basis change gates. The pairing term will be needed for the adiabatic evolution from the BCS mean-field Hamiltonian suggested in subsection III A 1. Also, it will be needed to measure superconducting correlations present in the Hubbard state: in this case the measurement will be of the correlation function between two different superconducting pairing terms on two pairs of sites separated from each other. Measurement will be discussed more later, but our general procedure for measurement will require knowledge of the unitary which implements evolution under the term.

For the case $\Delta$ real, we wish to implement unitary evolution by $\exp[-i\theta(c_{p,\sigma}^{\dagger}c_{q,\sigma'}^{\dagger} + h.c.)]$. Let us write $X_{p,\sigma}$ to denote the Pauli $X$ operator on the qubit corresponding to spin-orbital $p, \sigma$, and similarly $Y_{p,\sigma}$ and $Z_{p,\sigma}$ denote Pauli $Y$ and $Z$ operators on that spin-orbital. Using the identity that $c^\dagger = (1/2)(X + iY)$ and $c = (1/2)(X - iY)$, up to phase factor associated with the Jordan-Wigner strings, we wish to implement

$$\exp[-i\theta (X_{p,\sigma}X_{q,\sigma'} - Y_{p,\sigma}Y_{q,\sigma'})Z_{JW}], \quad (18)$$

where $Z_{JW}$ denotes the product of the Pauli $Z$ operators on the spin-orbitals intervening between $p, \sigma$ and $q, \sigma'$. In general, we wish to implement this unitary operation controlled by some ancilla. In fact, this circuit is the same as the circuit used to implement the term $c_{p,\sigma}^{\dagger}c_{q,\sigma'}^{\dagger} + h.c.$, up to a sign change on the second spin rotation, as shown in Fig. 10.

For $\Delta$ imaginary, we wish to implement

$$\exp[-i\theta (X_{p,\sigma}Y_{q,\sigma'} + Y_{p,\sigma}X_{q,\sigma'})Z_{JW}]. \quad (19)$$

This circuit is similar to the other pairing and hopping circuits, except that in the first half of the circuit, we apply a Hadamard to the first qubit and a $Y$ basis change to the second, and in the second half we apply the $Y$ basis change to the first qubit and the Hadamard to the second.

B. Optimizing the cost of the Jordan Wigner transformation

In order to implement the unitary time evolution in the Hubbard model, we use a Jordan-Wigner transformation to turn the model into a spin model. The depth of the circuits described in the previous subsection will depend upon the particular ordering of spin-orbitals chosen for this Jordan-Wigner transform. We choose to order so that all qubits corresponding to spin up appear first in some order, and then all qubits corresponding to spin down. Since the hopping terms preserve spin, this simplifies the Jordan-Wigner strings needed. For a given spin, we order the sites in the “snake” pattern shown in Fig. 11.

With this snake pattern, we group the hopping terms into four different sets, as shown. The terms in any given set all commute with each other. The two horizontal sets of terms have no Jordan-Wigner string required, since they involve hopping between neighboring sites with the given ordering pattern, and so these terms can all be executed in parallel. The vertical terms do not commute with each other. However, they nest as in Ref. (12) so that all vertical terms in a given set can be executed with $O(N)$ gates using a depth $O(\sqrt{N})$, where we assume that the geometry is roughly square so the number of sites in a given row is $\sqrt{N}$.

For periodic boundary conditions, we must also handle the terms going from left to right boundary or from top to bottom boundary. The terms going from left to right boundary can be executed in parallel with each other. Each term naively requires a depth proportional to the length of a given horizontal row to calculate the Jordan-Wigner string. Assuming a square geometry, this length is $O(\sqrt{N})$. We can use a tree to calculate the fermionic parity of the sites in the middle of the string, reducing the depth to $\log(N)$ if desired. For the terms from top to bottom edge, there are Jordan-Wigner strings of length $O(N)$, but these terms nest and so can all be executed in
FIG. 9. Quantum circuit to implement time evolution for a time step \( \theta \) under the hopping term \( c_p^{\dagger}c_{q,\sigma} + c_q^{\dagger}c_{p,\sigma} \). Top and bottom lines represent qubits for \( p,\sigma \) and \( q,\sigma \).

FIG. 10. Quantum circuit to implement time evolution for a time step \( \theta \) under the pairing term \( c_p^{\dagger}c_{q,\sigma'} + c_q^{\dagger}c_{p,\sigma} \). Top and bottom lines represent qubits for \( p,\sigma \) and \( q,\sigma' \).

FIG. 11. Snake pattern for ordering of sites. Arrow shows order of sites. Dashed and solid lines represent hopping terms. The hopping terms are grouped into four different sets, two sets of vertical and two sets of horizontal hoppings as shown by different colorations and dash patterns.

parallel, reducing the depth to \( O(\log(N)) \). The fermionic parity on the sites on the middle rows (those other than the top and bottom row) can be calculated using a tree, again in depth \( O(\log(N)) \). The total depth can then be \( O(\log(N)\sqrt{N}) \). In fact, using a tree to compute parities, the total depth can be reduced to polylogarithmic in \( N \).

Another way to reduce the cost of the Jordan-Wigner transform in this geometry is in Ref. 8. This method does require additional qubits.

V. FINDING THE GROUND STATE BY QUANTUM PHASE ESTIMATION

The annealing algorithm generates a state with good overlap with the ground state if the rate is sufficiently slow. VC discusses optimized paths that increase the overlap. However, in fact it is not necessary to obtain a perfect overlap with the ground state; if the overlap with the ground state is reasonably large, then we can apply quantum phase estimation to the state and have a reasonably large chance of projecting onto the ground state. Since the time for the simplest annealing path scales inversely with the gap squared, while the time to use quantum phase estimation to measure energies more accurately than the gap (and thus to determine whether or not one is in the ground state) scales inversely with the gap, up to logarithmic corrections, using quantum phase estimation to project an approximate state onto the ground state may be preferred. The gap will depend upon the annealing path; since we use a symmetry breaking field until nearly the end, the gap only closes at the end of the anneal. In this section we first give a simple improvement to quantum phase estimation that leads to a reduction in depth. We then give some numerical estimates for annealing times for free fermion systems (the simplest case where one can estimate annealing times for a large system numerically).

A. Reducing the number of rotations in quantum phase estimation

While this general setting of using adiabatic preparation and quantum phase estimation has been considered
many times before, we briefly note one simplification of the quantum phase estimation procedure. If the gate set available consists of one- or two-qubit Clifford operations, then this simplification provides a slightly larger than twofold reduction in depth; more importantly, it provides a fourfold reduction in the number of arbitrary rotations, which are expected to be the most costly operations to implement. This simplification is not in any way specific to the Hubbard model, and could be applied in other settings, such as in quantum chemistry. The usual approach to quantum phase estimation is to build controlled unitaries to implement the unitary evolution (by using the ancilla to control the unitaries described above). In this case, if the ancilla qubit is in the $|0\rangle$ state, then no quantum evolution is implemented, while if the ancilla qubit is in the $|1\rangle$ state, then quantum evolution by $U = \exp(-iHt)$ is implemented (up to Trotter-Suzuki error). Then, quantum phase estimation estimates the eigenvalues of this unitary $U$. The simple modification that we propose is that instead, if the ancilla qubit is in the $|0\rangle$ state, then we implement evolution by $\exp(iHt/2)$, while if the ancilla qubit is in the $|1\rangle$ state, then we implement evolution by $\exp(-iHt/2)$. Since the difference between these two evolutions is the same $\exp(-iHt)$, then the same phase estimation procedure on the ancilla will return the identical result for the energy (again up to Trotter-Suzuki error).

This approach means that we have then only to implement unitary evolution for half the time (i.e., $t/2$ rather than $t$), so assuming the same time step for each Trotter step means that the number of Trotter steps required is halved. However, to analyze whether or not this improves the gate depth, we need to determine how the gate depth might change in a given Trotter step. First, let us review the usual method of making circuits controlled to implement phase estimation, and then we will explain how to do it for the modification considered here. As an example, let us consider a circuit to implement a hopping term, as in Fig. 9. The usual technique (see references before) to make this circuit controlled (so that if the ancilla is $|0\rangle$ then the identity operation is performed while if the ancilla is $|1\rangle$ then evolution under the hopping term is performed) is to modify the two $R_x(\theta)$ gates, making them both controlled by an ancilla. A controlled rotation by angle $\theta$ implements no rotation if the ancilla is in the $|0\rangle$ state and implements a rotation by angle $\theta$ if the ancilla is in the $|1\rangle$ state. Different quantum computing architectures may make different elementary gates available. If we have access to arbitrary $Z$ rotations by angle $\theta$, and to Clifford operations, then the controlled rotation by angle $\theta$ is implemented by rotating the target by angle $\theta/2$, then applying a controlled NOT from the ancilla to target, then rotating the target again by angle $-\theta/2$, then again applying a controlled NOT from the ancilla to target, so that the net rotation is $\theta/2 - \theta/2 = 0$ if the ancilla is in the $|0\rangle$ state and is $\theta/2 + \theta/2 = \theta$ if the ancilla is in the $|1\rangle$ state. Thus, the controlled rotation is implemented by doing two arbitrary angle rotations, and additionally applying two Clifford operations, for a total of four gates given the gate set mentioned above.

A crucial point is that the only gates which need to be made controlled are the $R_x(\theta)$ gates. The other gates in Fig. 9 do not need to be modified in any way. The property still holds when we consider our proposed method of implementing evolution by either $\exp(iHt/2)$ or $\exp(-iHt/2)$, depending on the ancilla. In this method, we need to modify the $R_x(\theta)$ gates to implement a rotation by either a positive or negative angle; we term this a "uniformly controlled rotation": a rotation by angle $+\theta$ if the ancilla is in the $|0\rangle$ state and a rotation by angle $-\theta$ if the ancilla is in the $|1\rangle$ state. This can be done by applying a controlled NOT from the ancilla to target, then rotating the target, then again applying a controlled NOT from the ancilla to target. This requires one arbitrary angle rotation and two Clifford operations.

Thus, we find that the total number of gates required to implement the appropriately controlled version of Fig. 9 has been reduced by 1, as one fewer arbitrary rotation is required, while the same Clifford operations are required. So, the depth of a single Trotter step is slightly reduced. Since the number of Trotter steps has been halved, we find indeed a slightly larger than twofold reduction in depth, as claimed. Further, the number of arbitrary rotations is reduced by four. We have described this contraction for the hopping term; however, for any term, the usual technique is to replace arbitrary single qubits rotations by controlled rotations while we suggest instead replacing them with uniformyl controlled rotations.

**B. Annealing Time for Free Fermions**

To make some estimate of the cost of adiabatically preparing a good approximation to the ground state, it is worth considering the case of free fermions. For any system of free fermions with translation symmetry and $k$ sites per unit cell, the annealing problem decouples into many different problems, one for each Fourier mode, each problem describing annealing of a $k$-dimensional subspace. For $k = 2$, this is a Landau-Zener problem. For a single such Landau-Zener problem, the annealing time is expected to scale as $1/\Delta^2$ where $\Delta$ is the gap of the Hamiltonian. However, we have many Landau-Zener problems in parallel, and we want all Fourier modes to reach the ground state. The problem is least serious if the Fermi surface consists just of isolated points (such as in one dimension or for a semimetal in higher dimensions) as there are only a few Fourier modes with small gap. For systems with a Fermi surface, the density of states at the Fermi surface is larger, meaning that there are more modes with small gap. However, the probability of a diabolic transition for a Landau-Zener system is exponentially small in the ratio of the gap squared to the Landau-Zener velocity. Hence, even if there are a large number of modes with gap $\Delta$, this should require only a
logarithmic slowdown in the adiabatic velocity to have a small probability of a diabatic transition. However, one must also worry about effects due to the endpoints of the evolution. See, however, VC for a detailed discussion of improved annealing strategies that overcome some of the effects of the endpoints.

To put some numbers onto these generalities, we considered a system of spinless free fermions at half-filling in one dimension with periodic boundary conditions (this of course is the most advantageous case with only two modes with minimum gap). For each Fourier mode, we compute the evolution using numerically exact techniques (this is a two-dimensional quantum mechanics problem); to determine the probability of the entire system being in the ground state, we multiply the probabilities for each mode. We follow the annealing path

\[ H_s = \sum_{i \text{ even}} \left( c_i^\dagger c_{i+1} + h.c. \right) + s \sum_{i \text{ odd}} \left( c_i^\dagger c_{i+1} + h.c. \right), \tag{20} \]

starting at a fully dimerized state at \( s = 0 \) and moving to a uniform state at \( s = 1 \), where \( c_i^\dagger (c_i) \) creates (annihilates) a spinless fermion on site \( i \). We followed a uniform annealing schedule and used a time discretization with negligible Trotter error. We chose sizes which were not a multiple of 4 to avoid having an exact zero mode.

For a system 1002 sites, with an annealing time \( t = 1000 \), the probability of ending in the ground state is 0.54.... Thus, the expected time to end in the ground state is roughly 2000 plus an roughly additional 2 phase estimation steps on average. Reducing the annealing time to \( t = 500 \), the probability of ending in the ground state is 0.306..., taking slightly less expected annealing time, but more phase estimation steps on average. Increasing to 2002 sites, the situation is worse. For \( t = 1000 \), the probability of ending in the ground state is 0.13.... so the expected time is over 7500, plus over 7 phase estimation steps on average. Because the system is translationally invariant with period 2 for all \( s \), the probability of ending in the ground state is a product over momentum vectors of the probability of ending in the ground state in each momentum vector. Each of these probabilities can be calculated by evolving a 2-by-2 matrix. In all cases, the dominant contribution to the smallness of the probability to end in the ground state was from the lowest energy mode. Thus, small changes in the low energy density of states can lead to dramatic changes in the success probability.

For 1002 sites, the lowest eigenvalue of the free fermion hopping matrix is 0.0188..., so the gap \( \Delta \) of the Hamiltonian is twice that. Hence, in this case, the phase estimation time could be much smaller than the annealing time.

Since rather large Trotter time steps of order 0.25 are sufficient to obtain reasonable overlap with the ground state, we find that about \( 10^4 \) steps are sufficient to prepare the ground state. Using parallel circuit each step can be performed with a small circuit depth that increases only logarithmically with the system size. Hence a parallel circuit depth of about one million gates is sufficient to prepare the ground state.

C. Improved Annealing Paths

There are broadly two strategies for reducing the error in the annealing path. The typical objective in such optimizations is to find a function, \( f(s) \), that satisfies \( f : [0, 1] \rightarrow [0, 1] \) and \( f(0) = 0 \) and \( f(1) = 1 \) such that the diabatic leakage out of the ground state caused by \( U_{\text{adiab}} := e^{\int_0^T H_0 E_0(s) ds} T e^{-i\int_0^T H(f(s)) ds} \) is minimized for a fixed value of \( T \). Here we use the convention that \( T \) is the time ordering operator, \( T \) is the total evolution time allotted and \( s = t/T \) is the dimensionless time, and the phase factor is included to ensure that \( \lim_{T \to \infty} U_{\text{adiab}} = U_{\text{adiab}} \) is well defined.

The first strategy is known as local adiabatic interpolation.[44] The idea of local adiabatic interpolation is to choose \( f(s) \) to increase rapidly when the eigenvalue gap of \( H(f(s)) \) is large and increase slowly when the gap is small. This strategy can substantially reduce the time required to achieve a fixed error tolerance, but does not improve the scaling of the diabatic errors in the state preparation, which scale as \( \mathcal{O}(1/T) \) for Hamiltonians that are sufficiently smooth.

The second strategy, known as boundary cancellation is often diametrically opposed to local adiabatic optimization.[45] The idea behind it is instead of moving at a speed designed to minimize diabatic transitions, you choose a path designed to maximize cancellations between them. There are several strategies for exploiting this, but the simplest strategy to choose \( f(s) \) such that \( \partial_s^2 f(s) \big|_{s=0.1} = 0 \) for \( k = 1, \ldots, m \). Then if \( f(s) \) is at least \( m + 2 \) times differentiable and \( H(f(s)) \) does not explicitly depend on \( T \) then the error in the adiabatic approximation is at most [46]

\[ 2 \max_s \frac{\| \partial^m_{s} H(f(s)) \|}{\Delta(s)^{m+2Tm+1}} + O \left( 1/T^{m+2} \right), \tag{21} \]

where \( \Delta(s) \) is the gap at the given value of \( s \). Equation (21) therefore implies that if

\[ f(s) = \frac{\int_0^s y^m(1-y)^m dy}{\int_0^1 y^m(1-y)^m dy}, \tag{22} \]

then the upper bound on the asymptotic scaling is improved from \( \mathcal{O}(1/T) \) to \( \mathcal{O}(1/T^{m+1}) \). Furthermore, if a diabatic error of \( \delta \) is desired then it suffices to take

\[ T = O \left( \frac{\max_s \| \partial^m_{s} H(f(s)) \|}{\Delta(s)^{1+1/m+1}} \right). \tag{23} \]

This implies that even in the limit of small \( \delta \) the cost of adiabatic state preparation is not necessarily prohibitive because \( m \) can be increased as \( \delta \) shrinks to achieve improved scaling with the error tolerance. Furthermore, the
scaling of the error with the minimum gap becomes near–quadratically smaller than would be otherwise expected (for $\delta$ sufficiently small).

The above argument only discusses the scaling of the evolution time under the assumption that $m$ is fixed. If $m$ grows as $\delta$ shrinks then there may be $m$–dependent prefactors that are ignored in the above analysis. If the Hamiltonian is analytic in a strip about $[0, 1]$ and it remains gapped throughout the evolution then Lidar et al. show that, for fixed $T$, the error grows at most polynomially with $m$. Calculus then shows that the optimal value of $m$ scales logarithmically with the error tolerance and hence such contributions are subpolynomial.

These two strategies are often mutually exclusive when the avoided crossing does not occur near the boundary because if $H(s)$ is analytic then $H(s)$ will have to be larger away from the boundary to compensate for the fact that it is nearly constant in the vicinity of the boundary. This can cause the Hamiltonian to move rapidly through the avoided crossing, which increases the annealing time required. However, the minimum gap often occurs at the end of the evolution when transforming from the initial Hamiltonian (such as the free Fermion model, plaquette preparation, or other tractable approximation) to the interacting theory. This means that the two strategies are often compatible here and that boundary cancellation methods will often be well suited for high–accuracy state preparation. Further adiabatic optimizations could also be achieved by altering (22) to approximate the local–adiabatic path in the region away from from $s = 0, 1$.

We are of course concerned about the evolution time required to perform $U_{\text{adiab}} \approx U_{\text{diab}}$ since it dictates the resources needed to prepare the ground state $\psi_0$ within a fixed error by adiabatic evolution starting from the free fermion ground state $\psi_f$. A related problem is the problem of implementing the projector onto the ground state, $P_0 = |\psi_0\rangle \langle \psi_0|$ within a fixed error; we will wish to be able to do this for a method described in VII B. Given the projector $P_{0f} = |\psi_0\rangle \langle \psi_f|$ we have $P_0 = U_{\text{adiab}} P_{0f} U_{\text{adiab}}^\dagger$; we describe later how to measure $P_{0f}$ and so by conjugating this measurement by $U_{\text{adiab}}$ we can measure $P_0$.

At first glance, the use of a Trotter decomposition may appear to be problematic for adiabaticity because the time–dependent Hamiltonian that describes the decomposition is discontinuous. Such discontinuities are not actually problematic, as can be seen by using the triangle inequality

$$\begin{align*}
&\langle (U_{\text{adiab}} - U_{\text{adiab}}^{TS}) | \psi_0^f \rangle f \rangle
\leq \|U_{\text{adiab}} - U_{\text{adiab}}^{TS}\|
\leq \|U_{\text{adiab}} - U_{\text{diab}}\| + \|U_{\text{diab}} - U_{\text{diab}}^{TS}\|,
\end{align*}$$

where $U_{\text{adiab}}^{TS}$ is a quantum circuit approximation to the finite time evolution $U_{\text{adiab}}$. This shows that the error in the state preparation is at most the sum of the adiabatic error and the error in simulating the adiabatic evolution.

We have a similar bound for error in measuring $P_0$:

$$\begin{align*}
\|U_{\text{adiab}} P_{0f} U_{\text{adiab}}^\dagger - U_{\text{adiab}}^{TS} P_{0f} U_{\text{adiab}}^{TS}\|
&\leq 2 \left( \|U_{\text{adiab}} - U_{\text{diab}}\| + \|U_{\text{diab}} - U_{\text{diab}}^{TS}\| \right).
\end{align*}$$

In order to make the entire simulation error at most $\delta$, it suffices to set both sources of error to $\delta/2$. This creates an issue because the cost of simulation scales at least linearly with $T$, which in turn scales with $1/\delta$. There are a wide array of different Trotter–based formulas that can be used in the simulation but in all such cases the Trotter number (and in turn the simulation cost) scales exponentially with $k$ for integer $k > 0$ as $T^{1+1/2k}/\delta^{1/2k}$, which using (21) is $O \left( \delta^{-2k(m+1)} \right)$. For any fixed $m$, the value of $k$ that leads to the best scaling with $\delta$ can be found by numerically optimization. However, if we approximate this optimal value by taking $2k = m$ for even $m$ then the cost of the Trotterized simulation is $O(\delta^{m-2k}/m)$, which leads to subpolynomial (but not polylogarithmic) scaling of the total cost of simulation in $1/\delta$ if $2k = m$ and the Hamiltonian is sufficiently smooth because the cost of a Trotterized simulation grows exponentially with $k$ unlike the error in the adiabatic approximation.

This shows that the cost of implementing $P_0$ or preparing $\psi_0$ within error $\delta$, $T_0$, is sub–polynomial in $1/\delta$. Since $O(\epsilon^{-1}\log(1/\epsilon))$ repetitions of this circuit are needed and since errors are subadditive it follows that taking $\delta \propto \epsilon^2 / \log(1/\epsilon)$ suffices to make the total error at most $\epsilon$. Thus the cost of implementing $P_0$ scales as $e^{-1+o(1)}$ where the $o(1)$ term is the amalgamated costs of the adiabatic preparation and the logarithmic term from phase estimation.

It also should be noted that high–order Trotter–Suzuki formulas may not be needed in practice as numerical evidence suggests that small Trotter numbers typically suffice for the small Hubbard models that we have considered. This means that low–order methods may often be preferable to their higher–order brethren. Tight error bounds for the second order Trotter–Suzuki formula are given in the appendix.

VI. MEASURING OBSERVABLES

In this section we will discuss the measurement of interesting physical observables. The total energy, which could be measured by phase estimation, is the least interesting quantity. We will instead focus on densities and density correlations, kinetic energies, Green’s functions and pair correlation functions.

Such computations will allow us to understand the physics of the ground state. For instance, if the ground state is superconducting, then we can compute the correlation function of the pair field by the methods described in this section. Its long–distance behavior determines the condensate fraction. We can also compute the expectation value of the kinetic energy; its variation with doping
can help determine whether the system becomes superconducting as a result of kinetic energy gain compared to the non-superconducting state. Finally, more complex correlation functions, which would be very difficult – if not impossible to determine in experiments – could determine how these superconducting properties vary in response to perturbations that enhance or suppress effects such as spin or charge fluctuations.

A. Local observables and equal-time correlations

1. Measuring the density, double occupancy and spin and density correlations

The densities \( n_{i,\sigma} \) are trivially measured by measuring the value of the qubit corresponding to the spin-orbital \((i, \sigma)\). Similarly double occupancies \( n_{i,\uparrow} n_{i,\downarrow} \) can be determined by measuring two qubits, while density correlation functions

\[
n_{i} n_{j} = (n_{i,\uparrow} + n_{i,\downarrow})(n_{j,\uparrow} + n_{j,\downarrow})
\]

and spin correlation functions

\[
S_{i}^{z} S_{j}^{z} = \frac{1}{4}(n_{i,\uparrow} - n_{i,\downarrow})(n_{j,\uparrow} - n_{j,\downarrow})
\]

can be determined directly by measuring the values of four qubits.

By simultaneously measuring all qubits in the computational basis (eigenvectors of the Pauli-Z operator) we can determine density, double occupancy and all spin and density correlations. As an example, suppose we have recorded this sequence of measurement outcomes over several runs (for each run, for every qubit we measure the Pauli-Z operator and we record the measurement outcomes). Given this data, suppose we now wish to estimate a correlation function such as

\[
\langle \psi | n_{i,\uparrow} n_{j,\uparrow} | \psi \rangle
\]

for a pair of sites \( i \neq j \). This is equal to the expectation value of

\[
\frac{Z_{i,\uparrow} + 1}{2} \frac{Z_{j,\uparrow} + 1}{2}.
\]

Given the outcomes of the measurements of \( Z_{i,\uparrow}, Z_{j,\uparrow} \), we simply add 1 to each measurement, multiply the results, and divide by four. We then average this over runs. For example, if we consider a two-site Hubbard model at half-filling and \( U = 0 \), we will find that the two measurement outcomes \( Z_{i,\uparrow} \) and \( Z_{j,\uparrow} \) are perfectly anti-correlated (since there is only one electron with spin up which has equal probability to be on either of the two sites) and so the average of the product will not equal the product of the averages.

Next, we note that there is a general strategy used to measure the expectation value of any unitary operator \( U \), assuming that we can build a circuit that implements a controlled version of this unitary, controlled by some ancilla. Namely, we apply a one-bit phase estimation using the phase estimation circuit of Fig. 12. This is a standard trick; see for example Fig. 9 in Ref. 16. Since we have circuits that implement unitary evolution under various Hamiltonian terms, this enables us to measure these terms. For example, to measure a term \( c_{p,\sigma}^{+} c_{q,\sigma} \), we measure the expectation value of the unitary \( \exp[-i\theta(c_{p,\sigma}^{+} c_{q,\sigma} + h.c.)] \).

However, for the observables we consider, there in fact are much simpler ways of measuring the correlation functions. We give two different strategies. The first strategy involves replacing rotations in some of our unitary gates with measurements and we call it the “stabilizer strategy”; the second introduces a new gate called an “FSWAP”.

2. General strategy for other observables

The stabilizer strategy is a method for measuring observables of the form \( \exp[-i\theta(O_{1} + O_{2} + ...)] \) or of the form \( O_{1} + O_{2} + ... \) where each \( O_{i} \) is a product of some number of Pauli operators, and \( [O_{i}, O_{j}] = 0 \). This form includes many operators of interest to us, including terms in the Hamiltonian such as the kinetic energy as well as other terms such as pair correlation. We call this the stabilizer strategy because of our use of products of Paulis which commute with each other; no assumption is made that any state is a stabilizer state.

3. The Stabilizer Strategy

The stabilizer strategy is a method for measuring observables of the form \( \exp[-i\theta(O_{1} + O_{2} + ...)] \) or of the form \( O_{1} + O_{2} + ... \), where each \( O_{i} \) is a product of some number of Pauli operators, and \( [O_{i}, O_{j}] = 0 \). This form includes many operators of interest to us, including terms in the Hamiltonian such as the kinetic energy as well as other terms such as pair correlation. We call this the stabilizer strategy because of our use of products of Paulis which commute with each other; no assumption is made that any state is a stabilizer state.

If we can measure each \( O_{i} \), then we succeed in measuring the desired operator. Since they commute, we may measure them in any order without needing to recreate the state after measurement. As each \( O_{i} \) is a product of
Paulis, there is some unitary in the Clifford group which maps it onto a Pauli $Z$ operator on some given qubit. Hence, we can measure that $O_i$ by applying that Clifford unitary, then doing a $Z$-basis measurement, and finally undoing the Clifford unitary. Applying this procedure to terms in the Hamiltonian, for which we have previously given circuits to implement $\exp[-i\theta(O_1 + O_2 + \ldots)]$ using sequences of controlled $Z$-basis rotations conjugated by Clifford gates, the measurement circuit amounts to replacing each controlled $Z$-basis rotation in the evolution circuit with a $Z$-basis measurement. For example, if we apply this procedure to the $H_{pq}$ unitary evolution shown in Fig. 9, we arrive at the circuit shown in Fig. 13. Summing the value of the two $Z$-measurements (treating the answers as $\pm 1$ for each measurement) and then dividing by two (this division occurs because the controlled unitaries rotate by an angle equal to half the coupling strength) gives a measurement of the expectation operator of the hopping operator. In fact, we find that the same measurements are needed to measure the pairing operator for $\Delta$ real, since as we have noted, the pairing operator, up to a sign change in the second controlled rotation. Thus, the two measurements we have given simultaneously measure the two commuting operators $c_{p,\sigma}^\dagger c_{q,\sigma'} + h.c.$ and $c_{p,\sigma'}^\dagger c_{q,\sigma} + h.c.$ (to measure the pairing operator, one must instead consider the difference of the two $Z$ measurements).

4. The FSWAP Strategy for the kinetic energy and Green’s functions

For the kinetic energies we will have to measure one-body Green’s functions $c_{i,\sigma}^\dagger c_{j,\sigma} + c_{i,\sigma'}^\dagger c_{j,\sigma'}$. To do this we first swap qubit $j$ with its neighboring qubits until it is neighboring qubit $i$ in the normal ordering and is at one of the positions $k = i \pm 1$. Taking into account the fermionic nature of the electrons we need to use a new fermionic swap gate FSWAP with matrix

$$
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & -1
\end{pmatrix}
$$

(28)

The circuit expressed in standard gates is:

$$
\begin{array}{c}
\otimes \\
\biguplus \\
\downarrow
\end{array}
$$

(29)

We next change the basis to be able to better measure the kinetic energy on the bond which is then $-t_{pq} (c_{p,\sigma}^\dagger c_{q,\sigma} + c_{q,\sigma}^\dagger c_{p,\sigma})$. This basis change is a simple circuit:

$$
\begin{array}{c}
\otimes \\
\biguplus \\
\downarrow
\end{array}
$$

(30)

giving the unitary:

$$
\begin{pmatrix}
\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\
-\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\
0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\
0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0
\end{pmatrix}
$$

(31)

An example circuit containing several measurements in parallel is shown in Fig. 14.

We may then measure the $Z$ eigenvalues of the two qubits and call them $s_1$ and $s_2$. If we measure $s_1 = -1$, there is one particle on the two sites. We then have as estimators for the Green’s function $c_{p,\sigma}^\dagger c_{q,\sigma'} + c_{q,\sigma'}^\dagger c_{p,\sigma}$:

$$
s_2 \delta_{s_1,-1}
$$

(32)

and for the kinetic energy $-t_{pq} (c_{p,\sigma}^\dagger c_{q,\sigma} + c_{q,\sigma}^\dagger c_{p,\sigma})$ the estimator

$$
-t_{pq} s_2 \delta_{s_1,-1}
$$

(33)

Note that if we have $N$ qubits we can do $N/2$ measurements simultaneously, as long as no qubit is involved in two measurements.

Conversely, if the first qubit is +1, there are either 0 or 2 particles on the bond. In the presence of a pairing term, which breaks particle number conservation we can measure non-trivial values for the pair field operator $c_{p,\sigma}^\dagger c_{q,\sigma'} + h.c.$ as $s_2 \delta_{s_1,1}$.

5. The FSWAP Strategy for the pair correlation functions

Pair correlation functions in their most general form are built from the general 2-body terms $c_{i,\sigma}^\dagger c_{j,\sigma'}^\dagger c_{k,-\sigma} c_{l,\sigma}$. This expectation value can be measured after fermion-swapping the qubits to be adjacent and then performing two Green’s function measurements. We first use FSWAP gates to move the qubits $i$, $k$, $j$, and $l$ to four adjacent positions which we will label 1 to 4. We then apply the same circuit as for the Green’s function measurements to both pairs (1, 2) and (3, 4) and measure the $Z$ component of each qubit $m$ (=1, 2, 3, or 4) as $s_m$. The expectation value is then expressed as

$$
\langle c_{i,\sigma}^\dagger c_{k,\sigma'}^\dagger c_{j,\sigma} c_{l,\sigma'} \rangle = \langle s_2 s_4 \delta_{s_1,-1} \delta_{s_3,-1} \rangle
$$

Care must be taken if two indices are the same, e.g. if $i = k$ in above term. In that case things simplify since

$$
\langle c_{i,\tau}^\dagger c_{i,\tau}^\dagger c_{j,\tau} c_{j,\tau} \rangle = \langle n_i \delta_{\tau,\delta_{\tau}} \rangle
$$

(34)
and the measurement just becomes

$$\langle n_{i,\uparrow} c_{j,\downarrow} \dagger n_{k,\downarrow} c_{l,\uparrow} \rangle = \langle n_{i,\uparrow} s_4 \delta_{n_i,-1} \rangle \quad (35)$$

Similarly if \(i = k\) and \(j = l\) we get

$$\langle c_{i,\uparrow} c_{k,\downarrow} c_{j,\downarrow} c_{l,\uparrow} \rangle = \langle n_{i,\uparrow} n_{j,\downarrow} \rangle \quad (36)$$

and the measurement is straightforward.

For the Hubbard model we are interested in singlet pairing and the specific terms we want are:

$$\Delta_{ij,kl} = \frac{1}{2} \left( c_{i,\uparrow}^\dagger c_{j,\downarrow}^\dagger c_{j,\downarrow} - c_{i,\uparrow}^\dagger c_{j,\downarrow} c_{j,\downarrow} \right) \left( c_{k,\uparrow} c_{l,\downarrow} - c_{k,\downarrow} c_{l,\uparrow} \right). \quad (37)$$

Multiplying this out and sorting by spin we get

$$\Delta_{ij,kl} = -\frac{1}{2} \left( c_{i,\uparrow}^\dagger c_{k,\uparrow} + c_{i,\downarrow}^\dagger c_{k,\downarrow} \right) \left( c_{l,\uparrow} c_{j,\downarrow} + c_{l,\downarrow} c_{j,\uparrow} \right) + c_{i,\uparrow}^\dagger c_{k,\downarrow} c_{j,\downarrow} c_{j,\uparrow} c_{k,\downarrow} \quad (38)$$

which can be measured by individually measuring all terms.

Note that even though there may be \(N^4\) different choices for the four indices, we do not need to measure all of these four-point functions. As pairs are expected to be tightly bound in the Hubbard model, choosing \(i\) close to \(j\) and \(k\) close to \(l\), but choosing the pairs \((i,j)\) and \((k,l)\) at as large a distance as possible on a given lattice is sufficient to check for long-range pair correlations. In fact, we can measure \(c_{i,\uparrow}^\dagger c_{j,\downarrow}^\dagger + h.c\) (or, \(c_{i,\uparrow}^\dagger c_{j,\downarrow} - h.c\)) for \(N/2\) pairs \((i,j)\) simultaneously because the operators commute; then, averaging the result over pairs extracts the long wavelength pair correlation. To measure these with minimum depth, one can use nesting strategies similar to those discussed for measuring hopping terms in Sec. IV B.

B. Dynamic correlation functions and gaps

1. Dynamic response in the time domain

The measurement of dynamic correlation functions

$$C_{A,B}(t) = \langle \psi_0 | A(t) B(0) | \psi_0 \rangle \equiv \langle \psi_0 | e^{it\hat{H}} A^\dagger e^{-it\hat{H}} B | \psi_0 \rangle \quad (39)$$

for time-independent Hamiltonians such as the Hubbard model are typically presented not in the time domain but after a Fourier transform in the frequency domain:

$$S_{A,B}(\omega) = \int_{-\infty}^{\infty} dt C_{A,B}(t) e^{i\omega t} \quad (40)$$

The standard way of measuring them has been explained in several references, most explicitly in Ref. 16. For unitary operators \(A, B\), the product \(A^\dagger e^{it\hat{H}} B e^{-it\hat{H}}\) is a unitary, and can be measured using the circuit of Fig. 12. One simple improvement is to note that it is not necessary to control the evolutions \(e^{it\hat{H}} e^{-it\hat{H}}\) and it suffices to control \(A^\dagger, B\); then, if the ancilla bit is \(|0\rangle\) so that the operators \(A^\dagger, B\) are not done, the product \(e^{it\hat{H}} e^{-it\hat{H}}\) is equal to the identity as desired. The uncontrolled time evolution will require only half as many arbitrary angle rotations as the controlled time evolution for many gate sets (see Sec. V A for discussion of implementing controlled time evolution).

A further improvement is to note that the final evolution \(e^{-it\hat{H}}\) just gives an overall phase acting on the ground state, and so can be replaced by \(Z\) rotation of the ancilla without being implemented, giving another factor of two reduction in depth. A final improvement
combines the idea in Ref. 16 of controlling one operation and anti-controlling another with one of the ideas in Sec. V A of implementing either forward or backward evolution in time. We prepare the ancilla in the state $|+\rangle$. If the ancilla is $|0\rangle$, we implement $e^{itH/2}B$ and if the ancilla is $|1\rangle$ we implement $e^{-itH/2}A$. The measurement of the ancilla in the $X$ basis gives the desired expectation value. The controlled time evolution, by either $e^{itH/2}$ or $e^{-itH/2}$ depending on whether the ancilla is $|0\rangle$ or $|1\rangle$, can be done as described in Sec. V A. This requires half the number of time steps as that required implement the uncontrolled evolution by $e^{itH}$, assuming that the same timestep is used for both evolutions; the only additional cost is an additional CNOT gate for each arbitrary angle $Z$ rotation in the evolution and since these are Clifford gates the cost is much smaller than the cost of an arbitrary angle rotation for many gate sets.

2. Dynamic response in the frequency domain by phase estimation

A dynamic correlation function of an operator with its adjoint, such as

$$C_A(t) = \langle \psi_0 | A^\dagger(t) A(0) | \psi_0 \rangle$$

$$\equiv \langle \psi_0 | e^{+itH} A^\dagger e^{-itH} A | \psi_0 \rangle,$$

can be simplified after a Fourier transform into the frequency domain:

$$S_A(\omega) = \int_{-\infty}^{\infty} dt \ e^{i\omega t} C_A(t)$$

$$= \int_{-\infty}^{\infty} dt \ e^{i\omega t} \langle \psi_0 | e^{+itH} A^\dagger e^{-itH} A | \psi_0 \rangle$$

$$= \int_{-\infty}^{\infty} dt \ e^{i\omega t} \sum_n \langle \psi_0 | e^{+itH} A^\dagger e^{-itH} | \psi_n \rangle \langle \psi_n | A | \psi_0 \rangle$$

$$= \int_{-\infty}^{\infty} dt \ e^{i(\omega - (E_n - E_0)) t} \langle \psi_0 | A^\dagger | \psi_n \rangle \langle \psi_n | A | \psi_0 \rangle$$

$$= \sum_n \langle \psi_n | A | \psi_0 \rangle^2 \delta(\omega - (E_n - E_0)).$$

Instead of performing a “simple sampling” and measuring $C_A(t)$ for all times $t$ and then Fourier-transforming it and hoping to get these delta-functions resolved a new and better approach is to do “importance sampling” and measure the energies of the eigenstates $|n\rangle$ with eigenenergies $E_n$ directly with the weight $\langle \psi_n | A | \psi_0 \rangle^2$ with which they appear in above sum.

This can be achieved by phase estimation. If we apply phase estimation to the state $A|\psi_0\rangle$ then eigenstate $|\psi_n\rangle$ will be picked just with the weight $A \langle \psi_n | A | \psi_0 \rangle^2$ and a histogram of the measured energies thus directly measures the dynamic structure factor $S_A(\omega)$.

The retarded correlation function:

$$\chi_A(t) = \theta(t) \langle \psi_0 | [A^\dagger(t), A(0)] | \psi_0 \rangle$$

can then be obtained as follows. First, we write the retarded correlation function in terms of its real and imaginary parts, $\chi_A(\omega) = \chi''_A(\omega) + i\chi'_A(\omega)$. The imaginary part of the retarded correlation function, $\chi'_A(\omega)$ is then obtained from the fluctuation-dissipation theorem: $\chi''_A(\omega) = (1 - e^{-i\beta\omega}) S_A(\omega)$; at zero-temperature, $\beta = \infty$, and the second term in parentheses vanishes. The real part $\chi''_A(\omega)$ is then obtained by the Kramers-Kronig relation:

$$\chi_A(\omega) = \mathcal{P} \int_{-\infty}^{\infty} \frac{d\omega'}{\pi} \frac{\chi''_A(\omega)}{\omega' - \omega}$$

By choosing $A = Z_{p,\sigma}$ we can measure the local dynamical spin density correlations. Explicitly we find that

$$\langle Z_{p,\sigma}(0) Z_{p',\sigma'}(t) \rangle = 4\langle n_{p,\sigma}(0) n_{p',\sigma'}(t) \rangle$$

$$-2n_{p,\sigma}(0) - 2n_{p',\sigma'}(t) + 1.$$  

Since $n_{p',\sigma'}(t) = n_{p',\sigma'}(0)$ the last three terms do not depend on $t$ and only contribute a constant, which does not give any contribution to the interesting finite-$\omega$ structure factor. Local dynamical charge and spin correlations are straightforwardly calculated from the spin density correlations.

Similarly, local single-particle excitations can be measured by choosing

$$A = c_{p,\sigma}^\dagger + c_{p,\sigma} = \prod_{q<p} Z_{q,\sigma} X_{q,\sigma}.$$

Whether a particle or a hole excitation is realized in the excited state that is picked out by phase estimation can easily be determined by measuring the total particle number of that state.

Using circuits similar to the preparation of Slater determinants, we can measure dynamical correlation functions not only locally but for arbitrary single-particle wave functions. In particular, using momentum eigenstates one can obtain the momentum-resolved electron and hole spectral functions as measured in angle resolved photo emission (ARPES).

VII. NON-DESTRUCTIVE MEASUREMENTS

For the small systems that we have explicitly simulated, the time needed to adiabatically prepare a good estimate of the ground state is small compared to the phase estimation time. However, this will change as one increases the system size, and the time to adiabatically prepare the ground state will ultimately dominate, at least if one uses an annealing path that linearly interpolates between initial and final Hamiltonian. The reason is the time to adiabatically prepare the ground state is expected to scale as $\Delta^{-2}$ for a system with spectral gap $\Delta$, while resolving energies to accuracy $\Delta$ with phase estimation takes time $1/\Delta$ and thus for small $\Delta$ the phase estimation is faster. We note that the improved higher-order annealing paths in V C do alleviate this problem.
This problem is slightly alleviated by the fact that we are content to prepare by annealing a state with significant overlap (say, 1/2 overlap) with the ground state, as then phase estimation will give us a significant chance of projecting onto the ground state. However, once we have prepared a ground state \( \psi_0 \) by a combination of annealing and phase estimation, it will be desirable to measure properties of the state without destroying it. Here we present two approaches for such non-destructive measurements. Section VII A discusses an approach based on the Hellman-Feynman theorem, which not only is non-destructive but also scales better in the error than the simple way of repeatedly preparing the ground state and measuring. An alternate approach based on "recovering" the state after measurements may be more useful in some circumstances and is discussed in Sec. VII B.

A. Hellman-Feynman based approach

According to the Hellman-Feynman theorem the derivative of the ground state energy \( E_0(\lambda) \) with respect to a perturbation \( \lambda \) is just the expectation value of the operator \( O \) in the ground state \( |\Psi_0(\lambda)\rangle \):

\[
\frac{d}{d\lambda} E_0(\lambda) = \frac{d}{d\lambda} \langle \Psi_0(\lambda) | H + \lambda O | \Psi_0(\lambda) \rangle = \langle \Psi_0(\lambda) | O | \Psi_0(\lambda) \rangle. \tag{47}
\]

An alternative and superior way of measuring expectation values of arbitrary observables \( O \) is to adiabatically add a small perturbation \( \lambda O \) to the Hamiltonian \( H \).

This opens a way to measuring observables without destroying the ground state wave function \( |\Psi_0(0)\rangle \) of \( H \). We adiabatically evolve the ground state wave function \( |\Psi_0(0)\rangle \) to \( |\Psi_0(\lambda)\rangle \) by slowly increasing \( \lambda \) to its final (small) value, and then perform another quantum phase estimation to determine the energy \( E_0(\lambda) \). The ground state expectation value can be estimated through

\[
\langle \Psi_0(0) | O | \Psi_0(0) \rangle = \frac{d}{d\lambda} E_0(\lambda) |_{\lambda=0} = \frac{E_0(\lambda) - E_0(0)}{\lambda} + O(\lambda^2). \tag{48}
\]

For example, to measure a Green’s function we add a “hopping” term \( \lambda(c_{p,\sigma}^\dagger c_{q,\sigma} + c_{q,\sigma}^\dagger c_{p,\sigma}) \) to the Hamiltonian.

In order to obtain an estimate with an error \( \epsilon \) we need to choose \( \lambda = O(\epsilon) \), and then measure the energy to an accuracy \( \epsilon^2 \). Phase estimation then requires time \( O(\epsilon^{-2}) \), which scales the same as repeated preparation with destructive measurements and thus only gets a constant improvement.

The scaling can be improved by using a symmetric estimator for the derivative

\[
\langle \Psi_0(0) | O | \Psi_0(0) \rangle = \frac{E_0(\lambda) - E_0(-\lambda)}{2\lambda} + O(\lambda^3), \tag{49}
\]

which reduces the approximation error. To get the energy and thus allows a larger value of \( \lambda = O(\epsilon^{1/2}) \), which requires phase estimation only to an accuracy of \( O(\epsilon^{3/2}) \), and a time scaling as \( O(\epsilon^{-3/2}) \). In general, using a \( k \)-th order estimator for the derivative requires \( O(k) \) energy measurements, but with an error scaling as \( O(\epsilon^k) \), we can choose \( \lambda = O(\epsilon^{1/(k-1)}) \), and require time \( O(k\epsilon^{-1/(k-1)}) \). Asymptotically for small \( \epsilon \) we can thus estimate the expectation value non-destructively with an effort scaling as \( O(-\log(\epsilon)) \), obtaining a near-quadratic speedup.

In the Hellman-Feynman approach, there may be no need to do an adiabatic evolution. In many cases we could simply start in the state \( \psi_0 \), then apply a phase estimation to expectation value of \( H + \lambda O \), and then apply another phase estimation to estimate the ground state energy of \( H \). The absolute value squared of the overlap between the ground state of \( H \) and the ground state of \( H + \lambda O \) is equal to \( 1 - O(\lambda^2/\Delta^2) \), and so for \( \lambda << \Delta \) the overlap is close to unity. So, with probability close to 1, the energy given by phase estimation of \( H + \lambda O \) will in fact be the desired ground state energy, and with again probability close to 1 the second phase estimation will return to the ground state of \( H \).

B. Non-Destructive Measurements With Quadratic Speedup

This annealing procedure becomes less useful for operators \( O \) which require a complicated quantum circuit. Suppose we wish to measure the expectation value of a time dependent correlation such as \( O = S^z(t)S^z(0) \), where \( t > 0 \) and where \( S^z(t) = \exp(iHt)S^z\exp(-iHt) \). In this case, the quantum circuit to measure \( O \) (given below) is much more complicated; thus the Hamiltonian \( H + \epsilon O \) becomes significantly more costly to do phase estimation on.

In this section, we explain a technique to measure arbitrary projection operators in a non-destructive fashion. We then further explain a quadratic speedup of this approach. Our initial approach will give many independent binary measurements with outcome probability determined by the expectation value of the operator. This will require a number of samples proportional to the square of the inverse error. On the other hand, the improved approach will exploit phase estimation to speed this up quadratically (up to log factors).

As we have explained previously, we have quantum circuits that implement projective measurements of all the operators we are interested in, such as number operator, hopping operator, and so on. Further, we have a general technique for implementing a projective measurement of any unitary that can be controlled.

1. Recovery Map

Suppose we wish to implement a projective measurement with \( k \) possible outcomes, for some given \( k \). We describe these outcomes by projectors \( Q_1, \ldots, Q_k \), with
\[ \sum_i Q_i = I, \] and we refer to this measurement as “measuring \( Q \)”. The ground state \( \psi_0 \) can be written as a linear combination of states in the range of each of these projectors:

\[ \psi_0 = \sum_i a_i \phi_i, \quad (50) \]

where

\[ Q_i \phi_i = \phi_i, \quad (51) \]

and \( |\phi_i| = 1 \). Assume for the moment that we can also implement a measurement \( P_0 \) which projects onto the ground state \( \psi_0 \) (we describe how to do this to sufficient accuracy below). Then, the algorithm for nondestructive measurement is: first, begin in the ground state. Then, measure \( Q \). Then, measure \( P_0 \). If the result is that one indeed is in the ground state, then we have restored the ground state; at this point we can either re-measure the projector \( Q \) to obtain better statistics, or make some other measurement of a different observable. If instead the measurement of \( P_0 \) reveals that we are not in the ground state, we re-measure \( P_0 \) and then re-measure \( Q \). We repeat this process of re-measuring \( Q \) and re-measuring \( P_0 \) until we are in the ground state. The basic idea is similar to that in Ref. 48.

The rest of this subsubsection is focused on calculating the probability of returning to the ground state; most readers may wish to skip to the next subsubsection, which gives the more important quadratic speedup. The main reason for introducing the recovery map in this subsubsection is that it can be used after the quadratic speedup (see below). Let us compute the probability of returning to the ground state in this process. If the initial measurement of \( Q \) gives outcome \( i \) then the resulting state is \( \phi_i \); since, \( |\langle \psi_0 | \phi_i \rangle|^2 = |a_i|^2 \), we have a probability \( |a_i|^2 \) of returning to the ground state after measuring \( P_0 \). We now analyze the case that this measurement of \( P_0 \) reveals that we are not in the ground state; then the resulting state is equal to

\[ \frac{1}{\sqrt{1 - |a_i|^2}} (1 - P_0) \phi_i - \frac{a_i}{\sqrt{1 - |a_i|^2}} \psi_0, \quad (52) \]

up to a phase. Then, the probability that the subsequent measurement of \( Q \) gives outcome \( j \) is equal to

\[ |\langle \phi_j | \frac{1}{\sqrt{1 - |a_i|^2}} \phi_i - \frac{a_i}{\sqrt{1 - |a_i|^2}} \psi_0 \rangle|^2 = \frac{|\delta_{ij} - a_j a_i \pi_i|^2}{1 - |a_i|^2}, (53) \]

and the resulting state is exactly equal to \( \phi_j \). Repeating this, we find that after every measurement of \( Q \) with outcome \( i \), the probability that the measurement of \( P_0 \) will return to the ground state is \( |a_i|^2 \), while if we do not return to the ground state the probability that the next measurement of \( Q \) will give outcome \( j \) is given by Eq. (53).

The transitions are then governed by a Markov chain with transition probabilities

\[ T_{i \leftarrow i} = (1 - |a_i|^2)^2 = 1 - 2|a_i|^2 + |a_i|^4, \quad (54) \]

\[ T_{j \leftarrow i, j \neq i} = |a_i|^2 |a_j|^2 \quad (55) \]

\[ T_{0 \leftarrow i} = |a_i|^2, \quad (56) \]

where \( T_{j \leftarrow i} \) is that, starting in state \( \phi_i \), a measurement of \( P_0 \) reveals that one is not in the ground state, and then a measurement of \( Q \) gives outcome \( j \), while \( T_{0 \leftarrow i} \) is the probability that, starting in state \( \phi_i \), does leave on in the ground state. The initial measurement of \( Q \) at the start of the Markov chain gives state \( \phi_i \) with probability \( |a_i|^2 \). We would like to work out the expected number of measurements before the algorithm terminates back in the ground state.

To do this calculation, we re-interpret the probabilities, saying that with probability \( 2|a_i|^2 \) the system makes some transition; half the time this transition leads to state 0, while the other half the time the system makes a transition from state \( i \) to some state \( j \), and state \( j \) is chosen with probability proportional to \( |a_j|^2 \). See the factor \(-2|a_i|^2 \) in Eq. (54) which we interpret as the negative of the probability of there being a transition; note also that \( j \) may equal \( i \), so some of the “transitions” do not actually lead to a change in state: this is the factor of \( |a_i|^4 \) in Eq. (54). This re-interpretation only makes sense if \( |a_i|^2 \leq 1/2 \), of course but it will allow us to compute the expected time exactly and as the expected time is analytic, this calculation will then work for all \( |a_i|^2 \).

The advantage of this re-interpretation is that if a transition occurs, and if the transition leads to a state other than 0, then the probability that that state is equal to \( j \) is proportional to \( |a_j|^2 \), which is the same probability distribution as the chain began in.

Starting in state \( i \), the expected number of steps before a transition is \( 1/2|a_i|^2 \). Thus, averaging over states \( i \) with initial probability distribution \( |a_i|^2 \), the expected number of steps before a transition is \( k/2 \), where \( k \) is the number of possible outcomes of measurement \( Q \). Half of these transitions return to state 0, while the other half do not, leading to some state \( j \neq 0 \) with probability distribution proportional to \( |a_j|^2 \). Thus, the average number of steps before returning to the ground state is

\[ N_{\text{steps}} = k. \quad (57) \]

Note that both the number of phase estimations required and the number of measurements of \( Q \) required are equal to \( k \).
2. Quadratic Speedup

Suppose we wish to measure a projector $Q$. Define unitaries $U = 2Q - 1$ and $V = 2P_0 - 1$. Assume that $P_0$ has rank one. Applying any sequence of projections $P_0$ or $Q$ to the ground state $\psi_0$ gives some state in the space spanned by the non-orthogonal basis vectors $\psi_0, Q\psi_0$. Hence, we will work in this two-dimensional space for our analysis (the degenerate cases that $Q\psi_0 = \psi_0$ or $Q\psi_0 = 0$ can be handled easily too and one can verify that the final result will be correct in these cases also). In this subspace, in the non-degenerate case, we can write

$$V = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$U = \begin{pmatrix} 2\cos^2(\theta) - 1 & -2\cos(\theta)\sin(\theta) \\ -2\cos(\theta)\sin(\theta) & 2\sin^2(\theta) - 1 \end{pmatrix},$$

where the angle $\theta$ is such that $\langle \psi_0|Q|\psi_0\rangle = \cos^2(\theta)$. We emphasize that the above two equations are written in the **orthogonal** basis $\psi_0, Z(Q\psi_0 - \cos^2(\theta)\psi_0)$, where $Z$ is some normalization factor. Thus, the unitary $UV$ restricted to the subspace has eigenvalues

$$\exp(\pm 2i\theta).$$

We now build a quantum circuit that adds an additional ancilla that controls both $U$ and $V$, hence controls whether or not $UV$ is applied. We then implement phase estimation using this ancilla to determine the eigenvalues of $UV$. We use the ground state $\psi_0$ as input to the phase estimation circuit. Let $T_0$ be the time required to implement $UV$. Using the ancilla to control $(UV)^n$, for various powers of $n$, one is able to measure the eigenvalue of $UV$ to accuracy $\epsilon$ in a time proportional to $T_0\epsilon^{-1}\log(\epsilon^{-1})$; in this case, we take $n$ of order $\epsilon^{-1}$.

In fact, $UV$ has two distinct eigenvalues, and the phase estimation will return one of the two answers randomly. However, since the eigenvalues differ only in sign, and the desired expectation value $\langle \psi_0|Q|\psi_0\rangle$ is equal to $\cos^2(\theta)$, both eigenvalues give the same answer for this expectation value.

After implementing this phase estimation to the desired accuracy, we then are left with some state in the given subspace. If we then wish to recover the ground state (to recover some other observable), we can apply the recovery map above by alternating measurements of $Q$ and $P_0$ until the ground state is restored.

3. Measuring $P_0$

The above procedure relies on the ability to measure $P_0$. The key is to define an operation that will only identify whether or not one is in the ground state, without revealing additional information. If instead the measurement gives several bits of information on the energy of the state, the procedure does not work; intuitively, each $\phi_i$ is a coherent superposition of different energy eigenstates and making this more detailed measurement destroys this information.

We now describe how to do this; we call this procedure a “coherent phase estimation”. We have the ability to perform a phase estimation on the Hamiltonian $H$ to determine the energy of a state. This phase estimation is given by a quantum circuit, in which several control bits are initialized in a state $|0\rangle$. Then, a Hadamard is applied to each control bit. Then, the control bits are used to control the application of the Hamiltonian for a set period of time. Finally, the Hadamard is re-applied, and the control bits are measured in the computational basis. Then, classical post-processing is applied to extract the energy of the state. This process can be performed in serial (using only one control bit) or in parallel; the parallel approach uses more control bits but reduces the depth of the circuit. The measurement of the energy is not deterministic; however, one can employ a larger number of control bits to increase the accuracy.

To measure $P_0$, we use this phase estimation in its parallel form, but do not measure the outcomes of the control bits. We initialize an additional “outcome” bit to $|0\rangle$. We then use a unitary quantum circuit to implement the classical post-processing done to determine the energy, storing that energy value coherently. Then a unitary quantum circuit determines if that energy is near the ground state value (which we have determined in advance by a phase estimation before doing any measurements), and if so, it flips the value of the outcome bit. Finally, we uncompute, reversing the classical post-processing and phase estimation steps, and then at the end we measure the outcome bit. This procedure is essentially that in Ref. 48. While this procedure does not implement a projective measurement, but rather implements a POVM, if sufficient numbers of control bits are used, the measurement can be arbitrarily close to the desired projective measurement.

Performing this measurement of $P_0$ does require additional qubits. The number of qubits required scales proportional to the number of bits of accuracy desired; however, since this accuracy is of order $\Delta$, if $\Delta$ is only polynomially small in system size then this requires only logarithmically many extra bits. Additionally, the more accurately we implement a projective measurement rather than a POVM, the more control bits required, but this overhead also only scales logarithmically.

An alternate method of implementing $P_0$ is to note that we can easily implement the projector onto the ground state of the Hamiltonian at the start of the annealing process if our initial Hamiltonian is of a sufficiently simple form, such as a free fermion Hamiltonian or a Hamiltonian with a product ground state. For a general free fermion Hamiltonian, we can use Givens rotations to rotate to the case that the ground state simply has some
spin-orbitals occupied and some empty; we then measure whether we have the desired pattern or not. Call this projector $P^f_0$. Then, if $U_{\text{adiab}}$ is the unitary which describes the approximately adiabatic evolution from the free Fermion Hamiltonian onto the desired final Hamiltonian, we can approximate $P^f_0 \approx U_{\text{adiab}} P^f_0 U_{\text{adiab}}^\dagger$. The accuracy of this approximation depends upon the annealing path; see the discussion in VC, which shows that we can use this strategy to achieve a (near) quadratic speedup. The improved annealing paths discussed there are most useful for the particular approach here where we need very high accuracy in the annealing; in all other applications in the paper, we need only moderate accuracy in the annealing, sufficient to get a large overlap with the ground state so that phase estimation can then project onto the ground state with reasonably large probability. We therefore

VIII. CONCLUSIONS

In this paper we have presented a comprehensive strategy for using quantum computers to solve models of strongly correlated electrons, using the Hubbard model as a prototypical example. Similar strategies can be used for generalizations of the Hubbard model and other discrete quantum lattice models, including, but not limited to, the $t$-$J$ model, frustrated magnets, or bosonic models coupled to (static) gauge fields, which all suffer from negative sign problems in quantum Monte Carlo simulations on classical computers. We go beyond previous papers [5, 14–16] that discussed simulations of the Hubbard model on quantum computers by giving complete details of all steps needed to learn about the properties of the Hubbard model.

In particular we presented a strategy to prepare trial ground states starting from various mean-field states or RVB states on decoupled plaquettes. While finding the ground state of quantum lattice models is QMA-hard[3, 49, 50] this is a statement about the worst case. We expect that, as experience has shown, many experimentally relevant models have ground states with either no long-range order (so-called quantum liquids) or with broken symmetries that are qualitatively well described by mean-field theories. In fact, if a model has such peculiar properties that it is hard to find its ground state on a quantum computer, then also a material described by that model will have a hard time reaching its ground state. Quick low-temperature thermalization in experiments is thus an indication that the ground state of the model describing its properties should also be easy to prepare on a quantum computer. The opposite is the case, e.g. for (classical) spin glasses, where finding the ground state is NP-hard but also spin glass materials never thermalize nor reach the ground state.

We have furthermore presented an efficient and deterministic quantum algorithm to prepare arbitrary Slater determinants as initial state, which scales better than the algorithms of Refs. 5 and 16. This can be used to prepare ground states of various candidate mean-field Hamiltonians, from which an adiabatic evolution to the ground state of the Hubbard model can be attempted.

To implement time evolution under the Hubbard and mean-field Hamiltonians we have given explicit quantum circuits for all terms, and discussed the size of the Trotter time step required to achieve sufficiently small errors. We discuss how time evolution under the individual terms in the Hubbard model can be efficiently parallelized, ultimately requiring $O(N)$ qubits and gates with only $O(\log N)$ parallel circuit depth for one time step, which allows efficient simulations of very large systems. We gain additional constant factors over previous approaches, by optimizing the phase estimation algorithm to reduce the required number of (expensive) rotation gates by a factor of four, assuming that our gate set consists of one- or two-qubit Clifford operations and arbitrary single qubit rotations. We furthermore propose to use a larger Trotter time step for the adiabatic state preparation (whose time scale is controlled by the inverse gap squared), to prepare a very good (but not perfect) guess for the ground state, and then refine this to the exact ground state by doing only the final quantum phase estimation (whose time scale is shorter since it is controlled by the inverse gap) with a small time step.

We have finally discussed approaches to obtain a quadratic speedup in measurements by proposing two non-destructive measurement strategies, one based on the Hellman-Feynman theorem, and another based on recovering the ground state after a (destructive) measurement of a single qubit. We also introduce a new approach of measuring dynamic structure factors and spectral functions directly in frequency space, using ideas similar to angle-resolved photoemission experiments.

Estimating the gate counts required to simulate the Hubbard model, we find that even on lattices with more than $N \approx 1000$ sites – which should be large enough to learn most interesting properties – can be simulated on small scale quantum computers using only a few thousand qubits and a parallel circuit depth of about one million gates. This number is based on our estimate of a few thousand time steps with a circuit depth of not more than a few hundred gates each. The scaling of the number of Trotter steps depends upon the gap. We expect the gap will scale roughly inversely with the linear size or perhaps the square of the linear size. Many interacting systems have this scaling. This means that even with logical gate times of the order of 1µs, the ground state of the Hubbard model can be prepared within a few seconds. We thus believe that scale quantum computers will be a very powerful tool for the investigation of many problems in the field of strongly correlated electron models that are currently out of reach of any classical algorithm.
ACKNOWLEDGMENTS

We acknowledge useful discussions with Bela Bauer. This work was supported by Microsoft Research, by the Swiss National Science Foundation through the National Competence Center in Research NCCR QSIT, and the ERC Advanced Grant SIMCOFE. The authors acknowledge hospitality of the Aspen Center for Physics, supported by NSF grant # 1066293.

[1] R. P. Feynman, International Journal of Theoretical Physics 21, 467 (1982).
[2] A. Y. Kitaev, arXiv:quant-ph/9511026.
[3] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, Classical and quantum computation, 47 (American Mathematical Soc., 2002).
[4] M. A. Nielsen and I. L. Chuang, Quantum computation and quantum information (Cambridge university press, 2010).
[5] G. Ortiz, J. Gubernatis, E. Knill, and R. Laflamme, Physical Review A 64, 022319 (2001).
[6] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, Annals of Physics 298, 210 (2002), quant-ph/0003137.
[7] F. Verstraete and J. I. Cirac, J. Stat. Mech. 0509, P09012 (2005).
[8] H. F. Trotter, Proc. Amer. Math. Soc. 10, 545 (1959).
[9] M. Suzuki, Communications in Mathematical Physics 51, 183 (1976).
[10] D. Wecker, B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer, Phys. Rev. A 90, 022305 (2014).
[11] M. B. Hastings, D. Wecker, B. Bauer, and M. Troyer, Quantum Info. Comput. 15, 1 (2015).
[12] D. Poulin, M. B. Hastings, D. Wecker, N. Wiebe, A. Doherty, and M. Troyer, Preprint (2014), arXiv:1406.4920.
[13] D. Abrams and S. Lloyd, Physical Review Letters 79, 2586 (1997).
[14] D. Abrams and S. Lloyd, Physical Review Letters 83, 5102 (1999).
[15] R. Somma, G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Phys. Rev. A 65, 042323 (2002).
[16] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, 292, 472 (2001).
[17] S. Trebst, U. Schollwöck, M. Troyer, and P. Zoller, Phys. Rev. Lett. 96, 250402 (2006).
[18] A. Hamma and D. A. Lidar, Phys. Rev. Lett. 100, 030502 (2008).
[19] M. C. Gutzwiller, Phys. Rev. Lett. 10, 159 (1963).
[20] J. Hubbard, Proc. R. Soc. London Ser. A 276, 238 (1963).
[21] S. Sachdev, arXiv e-prints (2010), arXiv:1012.0299 [hep-th].
[22] E. Fradkin, S. A. Kivelson, and J. M. Tranquada, ArXiv e-prints (2014), arXiv:1407.4480 [cond-mat.supr-con].
[23] C. M. Varma, Phys. Rev. B 55, 14554 (1997).
[24] A system in which insulating behavior is caused by electron-electron interactions in a half-filled band (which would otherwise be metallic) is called a Mott insulator.
[25] D. Thouless, Proc. Phys. Soc.London 86, 893 (1965).
[26] In two dimensions, long-ranged magnetic order is impossible at temperatures $T > 0$ and must be stabilized by interactions between two-dimensional layers. Hence, while the temperature scale at which antiferromagnetic correlations become significant is $\sim J_s$, the inter-layer exchange.
[27] A. Hamma and D. A. Lidar, Journal of Physics A: Mathematical and General 23, 5777 (1990).
Appendix A: Error bounds for time-dependent Trotter Formulas

Simulating the dynamics of time–dependent Hamiltonians on quantum computers is a more subtle issue than simulating time–independent dynamics and results proven for the time–independent case do not necessarily transfer over. This issue is significant here because of issues surrounding simulating adiabatic state preparation. Bounds for the error in high–order Trotter formulas are known, however, the scaling of these bounds with the number of terms in the Hamiltonian is known to be loose. Since the Hamiltonians needed for our purposes have a large number of terms, tight estimates of the error scaling may be important for determining the cost of performing adiabatic state preparation for Fermionic systems.

It may be tempting to think that we can neglect the errors in the Trotter formula that are incurred by approximating $H(t)$ by a piecewise constant Hamiltonian because the time derivatives of the Hamiltonian are small for a slow adiabatic evolution, but this is not necessarily true because adiabatic theorems only require that the derivatives of the Hamiltonian are small relative to an appropriate power of the minimum eigenvalue gap. This means that in some circumstances adiabatic evolution may actually involve a rapid passage. Furthermore, since we are interested in high accuracy state preparation even though such errors may be small they will not necessarily be negligible compared to our target error tolerance.

In this section we analyze these errors in more detail. In particular, we will examine the error in the second order order Trotter–Suzuki formula

$$\mathcal{T}e^{-i \int_0^1 H(u) du} \approx \prod_{j=1}^m e^{-i H_j(1/2)/2} \prod_{j=m}^1 e^{-i H_j(1/2)/2},$$  \hspace{1cm} (A1)

where we approximate the evolution from time 0 to 1 by a single second-order Trotter-Suzuki step. The bounds that we present are a generalization of those in Ref. 11 to the time–dependent case. Bounds for the Trotter formula are given in Ref. 51.

Our main result is that

$$\left\| \mathcal{T} e^{-i \int_0^1 H(u) du} - \prod_{j=1}^m e^{-i H_j(1/2)/2} \prod_{j=m}^1 e^{-i H_j(1/2)/2} \right\| \leq \frac{1}{24} \max_s \|H''(s)\| + \frac{1}{12} \|[H'(1/2), H(1/2)]\|$$

$$+ \sum_{j=1}^m \left\| \left[ H_j(1/2), \sum_{k>j} H_k(1/2) \right], H_j(1/2) \right\|$$

$$+ \left\| \left[ \sum_{k>j} H_k(1/2), H_j(1/2) \right], \sum_{\ell>j} H_\ell(1/2) \right\|. \hspace{1cm} (A2)$$

This result reduces to the bound of 11 and agrees with the asymptotic error scaling predicted in 52, up to a use of the triangle inequality and a small multiplicative factor, if $H$ is time-independent.

We prove the bound in two steps. First, we will show that

$$\left\| \mathcal{T} e^{-i \int_0^1 H(u) du} - e^{-i H(1/2)} \right\| \leq \frac{1}{24} \max_s \|H''(s)\| + \frac{1}{12} \|[H'(1/2), H(1/2)]\|. \hspace{1cm} (A3)$$

Then, we use the bound of Ref. 11 that

$$\left\| e^{-i H(1/2)} - \prod_{j=1}^m e^{-i H_j(1/2)/2} \prod_{j=m}^1 e^{-i H_j(1/2)/2} \right\| \leq \sum_{j=1}^m \left\| \left[ H_j(1/2), \sum_{k>j} H_k(1/2) \right], H_j(1/2) \right\|$$

$$+ \left\| \left[ \sum_{k>j} H_k(1/2), H_j(1/2) \right], \sum_{\ell>j} H_\ell(1/2) \right\|. \hspace{1cm} (A4)$$

Eq. (A2) then follows from the triangle inequality. We remark that in Ref. 11, a typographical error led to $\|[A, H(x)], H(x)\| \leq \|[A, B], A\| + \|[A, B], B\|$ being replaced by $\|[A, H(x)], H(x)\| \leq \|[A, B], A\| + \|[A, B], B\|$ in the language of that paper in appendix B (see text above Eq. (B3) of that paper). This pair of missing $\|$ symbols propagated to later works although it does not affect any of the scaling results cited in later papers. In the above equation, we have corrected this error, so that the right hand-side is a sum of two distinct norms for each $j$, rather than a norm of a sum.

To show Eq. (A3), we write $H', \ldots$ to denote the first, second, ... derivatives of $H$ with respect to time. By Taylor’s theorem, $H(u) = H(1/2) + (u - 1/2) H'(1/2) + \int_{1/2}^u (u - s) H''(s) ds$. Hence $\mathcal{T} e^{-i \int_0^1 H(u) du}$ is

$$\mathcal{T} e^{-i \int_0^1 \left( H(1/2) + (u - 1/2) H'(1/2) + \int_{1/2}^u H''(s)(u-s) ds \right) du}.$$
and thus
\[
\| T e^{-i \int_0^1 H(u) du} - T e^{-i \int_0^1 (H(1/2) + (u-1/2)H'(1/2)) du} \| \\
\leq \int_0^1 \left\| \int_{1/2}^u H''(s) (u-s) ds \right\| du \\
\leq \frac{1}{24} \max_s \| H''(s) \|. \tag{A5}
\]

We next bound the difference
\[
\left\| T e^{-i \int_0^1 (H(1/2) + (u-1/2)H'(1/2)) du} - e^{-iH(1/2)} \right\|.
\]
Transforming to the interaction representation gives
\[
T e^{-i \int_0^1 (H(1/2) + (u-1/2)H'(1/2)) du} = e^{-i(1/2)H(1/2)} \left( T e^{-i \int_0^1 (u-1/2)H'_u(1/2) du} \right) \times e^{-i(1/2)H(1/2)},
\]
where we define
\[H'_u(1/2) \equiv e^{-i(1/2-u)H(1/2) H'(1/2)} e^{i(1/2-u)H(1/2)}.
\]
So by a unitary rotation
\[
\| T e^{-i \int_0^1 (H(1/2) + (u-1/2)H'(1/2)) du} - e^{-iH(1/2)} \| \\
= \| T e^{-i \int_0^1 (u-1/2)H'_u(1/2) du} - 1 \| \\
= \| T e^{-i \int_0^1 (u-1/2)H'_u(1/2) du} - T e^{-i \int_0^1 (u-1/2)H'(1/2) du} \|.
\]
Note that \( \| H'_u(1/2) - H'(1/2) \| \) can be written as
\[
\left\| \int_0^1 \partial_x \left( e^{-i(1/2-u)(1-x)H(1/2)} H'(1/2) e^{i(1/2-u)(1-x)H(1/2)} \right) dx \right\|
\]
which can be used to show that \( \| H'_u(1/2) - H'(1/2) \| \leq |u - 1/2| \| [H'(1/2), H(1/2)] \| \), and so
\[
\| T e^{-i \int_0^1 (u-1/2)H'_u(1/2) du} - T e^{-i \int_0^1 (u-1/2)H'(1/2) du} \|
\leq \int_0^1 |u - 1/2|^2 \| [H'(1/2), H(1/2)] \| du \\
= \frac{1}{12} \| [H'(1/2), H(1/2)] \|. \tag{A7}
\]
Eq. (A3) then follows from Eqs. (A5, A6, A7) and the triangle inequality.