


Heisenberg-Limited Ground-State Energy Estimation for Early Fault-Tolerant Quantum Computers

Lin Lin^{1,2,3,*} and Yu Tong^{1,†}

¹*Department of Mathematics, University of California, Berkeley, California 94720, USA*

²*Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

³*Challenge Institute of Quantum Computation, University of California, Berkeley, California 94720, USA*

 (Received 27 February 2021; revised 6 October 2021; accepted 6 January 2022; published 2 February 2022)

Under suitable assumptions, the quantum-phase-estimation (QPE) algorithm is able to achieve Heisenberg-limited precision scaling in estimating the ground-state energy. However, QPE requires a large number of ancilla qubits and a large circuit depth, as well as the ability to perform inverse quantum Fourier transform, making it expensive to implement on an early fault-tolerant quantum computer. We propose an alternative method to estimate the ground-state energy of a Hamiltonian with Heisenberg-limited precision scaling, which employs a simple quantum circuit with one ancilla qubit, and a classical postprocessing procedure. Besides the ground-state energy, our algorithm also produces an approximate cumulative distribution function of the spectral measure, which can be used to compute other spectral properties of the Hamiltonian.

DOI: [10.1103/PRXQuantum.3.010318](https://doi.org/10.1103/PRXQuantum.3.010318)

I. INTRODUCTION

Estimation of the ground-state energy of a quantum Hamiltonian is of immense importance in condensed-matter physics, quantum chemistry, and quantum information. The problem can be described as follows: we have a Hamiltonian H , acting on n qubits, with the eigendecomposition

$$H = \sum_{k=0}^{K-1} \lambda_k \Pi_k,$$

where Π_k is the projection operator into the λ_k eigensubspace and the λ_k 's are increasingly ordered. Each eigenvalue may be degenerate, i.e., the rank of Π_k can be more than 1. We assume that we can access the Hamiltonian H through the time-evolution operator $e^{-i\tau H}$ for some fixed τ . Our goal is to estimate the ground-state energy λ_0 to within additive error ϵ .

Some assumptions are needed, as otherwise this problem is quantum Merlin Arthur (QMA) hard [1–4]. We assume that we are given a state described by its density matrix ρ . Let $p_k = \text{Tr}[\rho \Pi_k]$. Then, if p_0 (i.e., the overlap between

the initial state and the ground state) is reasonably large, we can solve the ground-state energy-estimation problem efficiently. This assumption is reasonable in many practical settings. For example, in quantum chemistry, the Hartree-Fock method usually yields an approximate ground state that is easy to prepare on a quantum computer. At least for relatively small molecular systems, the Hartree-Fock state can often have a large overlap with the exact ground state [5]. Therefore, we may use the Hartree-Fock solution as ρ in this setting. Other candidates for ρ that can be relatively easily prepared on quantum computers have been discussed in Refs. [5–7] and an overview of methods to choose ρ can be found in Ref. [8, Sec. V.A.2].

The computational complexity of this task depends on the desired precision ϵ . Even in the ideal case where the exact ground state is given, this dependence cannot be better than linear in ϵ^{-1} for generic Hamiltonians [9]. This limit is called the Heisenberg limit [10–13] in quantum metrology. This notion is closely related to the time-energy uncertainty principle [9, 14–16]. This optimal scaling can be achieved using the quantum-phase-estimation (QPE) algorithm [17], which we discuss in detail later.

Much work has been done to develop the algorithms for ground-state energy estimation both for near-term quantum devices [18–21], and fully fault-tolerant quantum computers [22–25]. Relatively little work has been done for early fault-tolerant quantum computers [26–29], which we expect to be able to accomplish much more complicated tasks than current and near-term devices but these devices still place significant limitations on the suitable

*linlin@math.berkeley.edu

†yu_tong@berkeley.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

algorithms. Careful resource-cost estimation of performing QPE for the Hubbard model using a surface code to perform quantum error correction has been carried out in Refs. [28,30].

To be specific, we expect such early fault-tolerant quantum computers to have the following characteristics, (1) The number of logical qubits is limited. (2) It is undesirable to have a large number of controlled operations. (3) It is a priority to reduce the circuit depth, e.g., it is better to run a circuit of depth $\mathcal{O}(D)$ for $\mathcal{O}(M)$ times than to run a circuit of depth $\mathcal{O}(DM)$ for a constant number of times, even if use of the shorter circuit entails some additional polylogarithmic factors in the total run time.

In this context, the textbook version of QPE (see, e.g., Refs. [31,32]), which uses multiple ancilla qubits to store the phase and relies on the inverse quantum Fourier transform (QFT), has features that are not desirable on early fault-tolerant quantum computers. Some variants of QPE have been developed to achieve a high confidence level [33–35], which can be important in many applications. However, such modifications require even more ancilla qubits to store multiple estimates of the phase and an additional coherent circuit to perform logical operations. Another possible way to achieve a high confidence level is to utilize a resource state ([36, Sec. II B]) to implement a Kaiser-window filter [37]. This approach requires the same number of ancilla qubits as the textbook version of QPE.

Due to the above considerations, we focus on the variants of QPE that use only very few ancilla qubits (in fact, all algorithms below use only one ancilla qubit). Kitaev’s algorithm (see, e.g., Ref. [3]) uses a simple quantum circuit with one control qubit to determine each bit of the phase individually. However, this method, together with many other algorithms based on it [38,39], are designed for phase estimation with an eigenstate given exactly, which is different from our goal. The semiclassical Fourier transform [40] can simulate QFT+measurement (meaning that all qubits are measured in the end) with only one-qubit gates, classical control, and postprocessing, thus trading the expensive quantum resource for inexpensive classical operations. One can replace the inverse QFT with the semiclassical Fourier transform and this results in a phase-estimation algorithm that uses only one ancilla qubit [41,42]. This approach can be seen as a simulation of the multiple-ancilla qubit version of QPE and is therefore applicable to the case when ρ is not exactly the ground state. Because of these attractive features, this is the version of QPE used in Refs. [28,30]. However, as we explain below in Sec. I A, this type of QPE requires running coherent time evolution for time $\mathcal{O}(p_0^{-1}\epsilon^{-1})$. This leads to a large circuit depth when p_0 is small. Moreover, this approach cannot be used together with the resource state discussed earlier, because the resource state is not a product state.

In this work, the complexity is measured by the time for which we need to perform time evolution with the target Hamiltonian H . We use two metrics: (1) the *maximal evolution time*, which is the maximum length of time for which we need to perform (controlled) coherent time evolution, and (2) the *total evolution time*, which is the sum of all the lengths of time we need to perform (controlled) coherent time evolution. They describe, respectively, the circuit depth and the total run time. Moreover, we are primarily concerned with how they depend on the initial overlap p_0 and the precision ϵ . The dependence on the system size n mainly comes indirectly through p_0 and the conversion between the total evolution time and the run time, which we discuss in more detail later. We present an algorithm that achieves the following goals:

- (1) Achieves Heisenberg-limited precision scaling; i.e., the total time for which we run time evolution is $\tilde{\mathcal{O}}(\epsilon^{-1}\text{poly}(p_0^{-1}))$.
- (2) Uses at most one ancilla qubit.
- (3) The maximal evolution time is at most $\mathcal{O}(\epsilon^{-1}\text{polylog}(\epsilon^{-1}p_0^{-1}))$.

To the best of our knowledge, our algorithm is the first to satisfy all three requirements. In our algorithm, we sample from a simple quantum circuit and use the samples to approximately reconstruct the cumulative distribution function (CDF) of the spectral measure associated with the Hamiltonian. We then use classical postprocessing to estimate the ground-state energy with high confidence. Besides the ground-state energy, our algorithm also produces the approximate CDF, which may be of independent interest. In the above discussion, we assume that the controlled time evolution can be done efficiently. If controlled time evolution is costly to implement, then based on ideas in Refs. [18,43–45], in Appendix E we offer an alternative circuit that uses two ancilla qubits, with some additional assumptions.

The problem of ground-state energy estimation is closely related to that of ground-state preparation but there are important differences. First, having access to a good initial state ρ (with large overlap with the ground state) does not make the energy estimation a trivial task, as even if we have access to the exact ground state, the quantum resources required to perform phase estimation can still be significant. Second, ground-state energy-estimation algorithms do not necessarily involve ground-state preparation. This is true for the algorithm in this work as well as in Refs. [23,24]. Consequently, even though the ground-state preparation algorithms generally have a run time that depends on the spectral gap between the two lowest eigenvalues of the Hamiltonian, the cost of ground-state energy-estimation algorithms may not necessarily depend on the spectral gap.

We remark that although we characterize the scaling as depending on the overlap p_0 , in practice we need to know a lower bound of p_0 , which we denote by η . The dependence on p_0 should more accurately be replaced by a dependence on η . To the best of our knowledge, in order to obtain a rigorous guarantee of the performance, knowledge of η (and the fact that η is not too small) is needed in all previous algorithms related to QPE. This is because in QPE we need knowledge of η to obtain a stopping criterion. We briefly explain this using a simple example. Suppose that we have a Hamiltonian H on n qubits with eigenvalues λ_k (arranged in ascending order) and eigenstates $|\psi_k\rangle$ and that $|\phi_0\rangle$ is an initial guess for the ground state. Furthermore, we assume that $p_0 = |\langle\phi_0|\psi_0\rangle|^2 = 0.01$, $p_1 = |\langle\phi_0|\psi_1\rangle|^2 = 0.5$. We may idealize QPE as an exact energy measurement to simplify the discussion. If we have no *a priori* knowledge of p_0 , then performing QPE on the state $|\phi_0\rangle$ will give us λ_1 with probability $1/2$. If we repeat this $\lesssim 100$ times, most likely all the energies we obtain will be $\geq \lambda_1$. Only when we measure $\gtrsim 100$ times can we reach the correct ground-state energy λ_0 . Hence if we do not know about a lower bound of p_0 , we can never know whether we have stopped the algorithm prematurely.

The main idea of our algorithm is to use a binary-search procedure to gradually narrow down the interval in which the ground-state energy is located. The key component is a subroutine CERTIFY (Algorithm 2) that distinguishes whether the ground-state energy is approximately to the left or right of some given value. This, however, can only be performed up to a certain precision and can fail with nonzero probability. Therefore, our search algorithm needs to account for this fuzzy outcome to produce a final result that is correct with probability arbitrarily close to 1. In the CERTIFY procedure, we use a stochastic method to evaluate the cumulative distribution function associated with the spectral density and this is the key to achieving the Heisenberg scaling. This stochastic method is described in detail in Sec. III.

A. Related works

We first briefly analyze the cost of the textbook version of QPE using multiple ancilla qubits. Although this method has features that are not desirable on early fault-tolerant quantum computers, this analysis is nevertheless helpful for understanding the cost of other variants of QPE. For simplicity, we assume that $\rho = |\phi\rangle\langle\phi|$ is a pure state and that the ground state $|\psi_0\rangle$ is nondegenerate. Approximately, the QPE performs a projective measurement in the eigenbasis of H . With probability p_0 , $|\phi\rangle$ will collapse to the ground state $|\psi_0\rangle$. If this happens, the energy register will then give the ground-state energy λ_0 to precision ϵ . Therefore, we run phase estimation for a total of $\mathcal{O}(p_0^{-1})$ times and take the instance with the minimum value in the energy register. With high probability, this value will be

close to λ_0 . Each single run takes time $\mathcal{O}(\epsilon^{-1})$. The total run-time cost is therefore $\mathcal{O}(p_0^{-1}\epsilon^{-1})$. For simplicity here, we do not consider the run time needed to prepare $|\phi\rangle$.

The above analysis, however, is overly optimistic. Since we need to repeat the phase-estimation procedure for a total of $\mathcal{O}(p_0^{-1})$ times, for an event that only has $\mathcal{O}(p_0)$ probability of happening in a single run, the probability of this event occurring at least once in the total $\mathcal{O}(p_0^{-1})$ repetitions is now $\mathcal{O}(1)$ (which means that we cannot ensure that the error happens with sufficient low probability). In our setting, suppose that the maximal evolution time is T ; then each time we measure the energy register, there is a $\mathcal{O}(T^{-1}\epsilon'^{-1})$ probability that the output will be smaller than $\lambda_0 - \epsilon'$. If we choose $T = \mathcal{O}(\epsilon^{-1})$ as discussed above and we let $\epsilon' = \epsilon/p_0$, then the probability of the minimum of the $\mathcal{O}(p_0^{-1})$ energy-register measurement outputs being smaller than $\lambda_0 - \epsilon/p_0$ is only upper bounded by $\mathcal{O}(1)$ and we no longer have control over the probability of the error being larger than ϵ . This means that there might be a high probability that the error of the ground-state energy will, in the end, be of order ϵ/p_0 instead of ϵ . For a more formal analysis, see Ref. [23, Appendix A]. We numerically demonstrate that this is indeed the case in Fig. 1, in which we show that the error increases as p_0 decreases and that there is a larger probability of the estimate deviating beyond a prescribed tolerance if the maximal evolution time—or, equivalently, the circuit depth—for QPE is fixed.

To avoid this, one can instead choose the maximal evolution time to be $T = \mathcal{O}(p_0^{-1}\epsilon^{-1})$. After repeating $\mathcal{O}(p_0^{-1})$ times, the total run time then becomes $\mathcal{O}(p_0^{-2}\epsilon^{-1})$. The increase in the maximal evolution time can prevent the increase of the error (see Fig. 1). However, the extra p_0^{-1} factor increases the circuit depth and is undesirable.

There are several other algorithms based on phase estimation using a single ancilla qubit [38,39,46] that are designed for different settings from ours: they assume the availability of an exact eigenstate or are designed for obtaining the entire spectrum and thus only work for small systems. In Ref. [47], a method is proposed for estimating the eigenvalues by first estimating $\text{Tr}[\rho e^{-itH}]$ and then performing a classical Fourier transform but no run-time scaling is provided. The semiclassical Fourier transform [40] simulates the QFT in a classical manner and QPE using one ancilla qubit and the semiclassical Fourier transform has the same scaling in terms of the maximal evolution time and the total evolution time.

In order to improve the dependence on p_0 , we may use the high-confidence versions of the phase-estimation algorithm [33–35]. In this method, the maximal evolution time required can be reduced to $\mathcal{O}(\epsilon^{-1} \log(p_0^{-1}))$, through taking the median of several copies of the energy register in a coherent manner. However, this requires using multiple copies of the energy register, together with an additional quantum circuit to compute the medians coherently, which

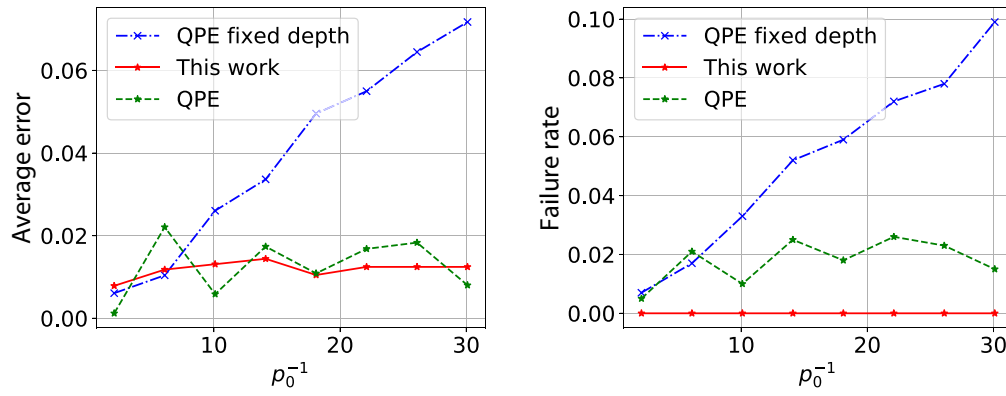


FIG. 1. A comparison of the performance of the textbook version QPE (blue dashed-dotted line) and the method in this work (red solid line) in ground-state energy estimation with a fixed maximal evolution time (300 steps of time evolution with H) and decreasing initial overlap p_0 . The results are benchmarked against QPE with maximal evolution time proportional to p_0^{-1} (green dashed line). To use QPE, either with fixed or $\mathcal{O}(p_0^{-1})$ maximal evolution time, to estimate the ground-state energy, we run QPE for $\mathcal{O}(p_0^{-1})$ times and take the minimum in the energy-measurement outcomes as the ground-state energy estimate. The error is averaged over multiple runs and the failure rate is the percentage of runs that yields an estimate with an error larger than the tolerance 0.04. The Hamiltonian H is the Hubbard Hamiltonian defined in Eq. (F1) with $U = 10$ and the overlap p_0 is artificially tuned.

can be difficult to implement. Note that the semiclassical Fourier transform can only simulate the measurement outcome and does not preserve coherence and therefore, to our knowledge, the high-confidence version of phase estimation cannot be modified to use only a single qubit. In Ref. [23], the authors have used a method called minimum label finding to improve the run time to $\mathcal{O}(p_0^{-3/2} \epsilon^{-1})$ but the implementation of the minimum label finding with limited quantum resources is again difficult.

Besides these algorithms based on phase estimation, several other algorithms have been developed to solve the ground-state energy problem. In Ref. [23], a method has been proposed based on the linear-combination-of-unitaries (LCU) technique that requires running time evolution for duration $\tilde{\mathcal{O}}(p_0^{-1/2} \epsilon^{-3/2})$ and preparing the initial state $\tilde{\mathcal{O}}(p_0^{-1/2} \epsilon^{-1/2})$ times [48]. Assuming that the Hamiltonian H is available in its block encoding [49,50], Ref. [24] uses quantum signal processing [51,52] with a binary-search procedure, which queries the block encoding $\tilde{\mathcal{O}}(p_0^{-1/2} \epsilon^{-1})$ times and prepares the initial state $\tilde{\mathcal{O}}(p_0^{-1/2} \log(\epsilon^{-1}))$ times. To our knowledge, this is the best complexity that has been achieved. However, the block encoding of a quantum Hamiltonian of interest, LCU, and amplitude estimation techniques (used in Ref. [24]) are expensive in terms of the number of ancilla qubits, controlled operations, and logical operations needed.

A very different type of algorithm for ground-state energy estimation is the variational quantum eigensolver (VQE) [19–21], which is a near-term algorithm and has been demonstrated on real quantum computers. The accuracy of the VQE is limited both by the representational power of the variational ansatz and the capabilities of

classical optimization algorithms for the associated non-convex optimization problem. Hence unlike the aforementioned algorithms, there are no provable performance guarantees for VQE-type methods. In fact, some recent results show that solution of the nonconvex optimization problem can be nondeterministic polynomial-time (NP) hard [53]. Furthermore, each evaluation of the energy expectation value to precision ϵ requires $\mathcal{O}(\epsilon^{-2})$ samples due to Monte Carlo sampling. This can, to some extent, be remedied using the methods in Ref. [33,38], at the expense of the larger circuit-depth requirement.

There are also a few options that can be viewed to be in between the VQE and QPE. The quantum-imaginary-time-evolution (QITE) algorithm [54] uses state tomography to turn an imaginary time evolution into a series of real-time Hamiltonian-evolution problems. Inspired by the classical Krylov-subspace method, Refs. [18,55,56] propose to solve the ground-state energy problem by restricting the Hilbert space to a low-dimension space spanned by some eigenstates that are accessible with time evolution. Similar to the VQE, no provable complexity upper bound is known for these algorithms and all algorithms suffer from the ϵ^{-2} scaling due to the Monte Carlo sampling. In fact, the stability of these algorithms remains unclear in the presence of sampling errors.

A more ambitious goal than ground-state energy estimation is to estimate the distribution of all eigenvalues weighted by a given initial state ρ [46,57,58]. Using a quantum circuit similar to that in Kitaev’s algorithm as well as classical postprocessing, Ref. [58] has proposed an algorithm to solve the quantum eigenvalue estimation problem (QEEP). We henceforth refer to this algorithm as the quantum eigenvalue estimation algorithm (QEEA). Suppose that $\|H\| \leq 1/2$ and that the interval

TABLE I. Quantum algorithms for estimating the ground-state energy and whether they satisfy each of the three requirements. We recall that the requirements are (1) achieving the Heisenberg-limited precision scaling, (2) using at most one ancilla qubit, and (3) the maximal evolution time being at most $\mathcal{O}(\epsilon^{-1}\text{polylog}(\epsilon^{-1}p_0^{-1}))$.

| Algorithms | Requirements | | | Other issues |
|------------------------------------|--------------|-----|-----|--------------------------------------|
| | (1) | (2) | (3) | |
| QPE (textbook version) [31,32] | ✓ | ✗ | ✗ | |
| QPE (high-confidence) [33–35] | ✓ | ✗ | ✓ | |
| QPE (semiclassical QFT) [41,42] | ✓ | ✓ | ✗ | |
| QPE (iterative) [3] | ✓ | ✓ | ✓ | Needs exact eigenstate ($p_0 = 1$) |
| LCU approach [23] | ✗ | ✗ | ✗ | |
| Binary-search approach [24] | ✓ | ✗ | ✗ | |
| VQE [19–21] | ✗ | ✓ | ? | No precision guarantee |
| QITE [54] | ✗ | ✓ | ? | Requires state tomography |
| QEEA [58] | ✗ | ✓ | ✓ | |
| Krylov-subspace methods [18,55,56] | ✗ | ✓ | ? | No precision guarantee |
| This work | ✓ | ✓ | ✓ | |

$[-\pi, \pi]$ is divided into M bins of equal size, denoted by $B_j = [-1/2 + j/M, -1/2 + (j+1)/M]$. Then, QEEA estimates the quantities $q_j = \sum_{k:\lambda_k \in B_j} p_k$. Although QEEA was not designed for ground-state energy estimation, one can use this algorithm to find the leftmost bin in which $q_j \geq p_0/2$ and thereby locate the ground-state energy within a bin of size M^{-1} . While the maximal evolution time required scales as $\mathcal{O}(\epsilon^{-1})$, the total evolution time of the original QEEA scales as $\mathcal{O}(\epsilon^{-6})$. We analyze the cost of QEEA in Appendix C and show that the total run time can be reduced to $\mathcal{O}(\epsilon^{-4})$ for the ground-state energy estimation in a straightforward way; yet this is still costly if high precision is required.

To the extent of our knowledge, none of the existing algorithms achieves all three goals. Some can have better maximal evolution time or total evolution-time requirement but the advantage always comes at the expense of some other aspects. In Table I, we list the quantum algorithms discussed in this work and whether they satisfy each of the requirements.

In Table II, we compare the maximal evolution time, the number of repetitions (the number of times we need to run the quantum circuit), and the total evolution time

needed, using the three qubit-efficient methods that require only one ancilla qubit.

Finally, in a gate-based setting, the exact relations between the maximal evolution time and the circuit depth, and between the total evolution time and the total run time, can be affected by the method we use to perform time evolution. Suppose that we have access to a unitary circuit that performs $e^{-i\tau H}$ exactly for some fixed τ . Then, in order to run coherent time evolution for time T , we only need to use a circuit of depth $\mathcal{O}(T)$. Therefore, the circuit depth scales linearly with respect to the maximal evolution time. Similarly, the total run time scales linearly with respect to the total evolution time.

However, if we can only perform time evolution through Hamiltonian simulation, then these relations become more complicated. If advanced Hamiltonian simulation methods [50,52,59] can be used, the additional cost will be asymptotically negligible, since to ensure an ϵ' error for time evolution for time T , the cost is $\mathcal{O}(T\text{polylog}(T\epsilon'^{-1}))$. Hence the cost is only worse than that in the ideal case by a polylogarithmic factor. However, for early fault-tolerant quantum computers, as discussed in Refs. [28,30], Trotter formulas [60] are generally favored. Running time evolution for time T with error at most ϵ' would entail

TABLE II. Comparison of the maximal evolution time, the number of repetitions (the number of times we need to run the quantum circuit), and the total evolution time needed for estimating the ground-state energy to within error ϵ , using the three methods that require only one ancilla qubit: the method in this work; QPE with semiclassical Fourier transform, which uses only one ancilla qubit; and the QEEA in Ref. [58]. The overlap between the initial state and the ground state is assumed to be p_0 . The number of repetitions is also the number of times we need to prepare the initial state. An analysis of the QEEA in Ref. [58] can be found in Appendix C.

| | Maximum evolution time | Repetitions | Total evolution time |
|--|--|--|--|
| This work (Corollary 3) | $\tilde{\mathcal{O}}(\epsilon^{-1}\text{polylog}(p_0^{-1}))$ | $\tilde{\mathcal{O}}(p_0^{-2}\text{polylog}(\epsilon^{-1}))$ | $\tilde{\mathcal{O}}(\epsilon^{-1}p_0^{-2})$ |
| QPE with semiclassical Fourier transform | $\tilde{\mathcal{O}}(\epsilon^{-1}p_0^{-1})$ | $\tilde{\mathcal{O}}(p_0^{-1}\text{polylog}(\epsilon^{-1}))$ | $\tilde{\mathcal{O}}(\epsilon^{-1}p_0^{-2})$ |
| QEEA [58] | $\tilde{\mathcal{O}}(\epsilon^{-1}\text{polylog}(p_0^{-1}))$ | $\tilde{\mathcal{O}}(\epsilon^{-3}p_0^{-2})$ | $\tilde{\mathcal{O}}(\epsilon^{-4}p_0^{-2})$ |

a run time of $\mathcal{O}(T^{1+1/p} \epsilon'^{-1/p})$. The additional cost will therefore prevent us from reaching the Heisenberg limit, though high-order Trotter formulas (i.e., with a large p) can allow us to get arbitrarily close to the Heisenberg limit. If one does not insist on having a Heisenberg-limited scaling, then randomized algorithms [61–63] may lead to a lower gate count when only low precision is required.

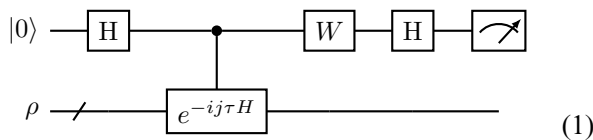
In Appendix D, we analyze the circuit depth and the total run time of our algorithm with time evolution performed using Trotter formulas. We also compare with QPE based on Trotter formulas. We find that when using Trotter formulas, our method has some additional advantage over QPE, achieving a polynomially better dependence on p_0 (i.e., η in Appendix D) in the total run time. The total run time scales like $\epsilon^{-1-o(1)}$ using our algorithm with Trotter formulas and this only approximately reaches the Heisenberg limit ϵ^{-1} in terms of the total run time. However, it is worth noting that none of the other methods can strictly reach the Heisenberg limit using Trotter formulas. Otherwise, we can instead perform Hamiltonian simulation with the exponentially accurate methods to go below the Heisenberg limit, which is an impossible task. Despite the suboptimal asymptotic scaling, with tight error analysis [64–67], Trotter formulas may outperform the advanced Hamiltonian simulation techniques discussed above in terms of the gate complexity, especially when only moderate accuracy is needed.

B. Organization

The rest of the paper is organized as follows. In Sec. II, we introduce the quantum circuit we are going to use and we introduce the CDF, which is going to play an important role in our algorithm, and give an overview of the ground-state energy-estimation algorithm. In Sec. III, we discuss how to approximate the CDF. In Sec. IV, we show that the ground-state energy can be estimated by inverting the CDF and we present the complexity of our algorithm (Corollary 3). In Sec. V, we present the details of our algorithm for postprocessing the measurement data and analyze the complexity.

II. OVERVIEW OF THE METHOD

We want to keep the quantum circuit we use as simple as possible. In this work, we use the following circuit:



where H is the Hadamard gate. We choose $W = I$ or $W = S^\dagger$, where S is the phase gate, depending on the quantity we want to estimate. The quantum circuit is simple and uses

only one ancilla qubit as required. The quantum circuit itself has been used in previous methods [3,58]. However, our algorithm uses a different strategy for querying the circuit and for classical postprocessing and results in lower total evolution time and/or maximal evolution time, achieving the goals (1) and (3).

This circuit requires controlled time evolution, which can be nontrivial to implement. The idea of removing controlled operation in phase estimation has also been considered in Ref. [68]. Here, we can use ideas from Refs. [18,43–45] to remove the need to perform controlled time evolution. But this type of approach requires an eigenstate of H with known eigenvalue that is easy to prepare. In a second-quantized setting, we can simply use the vacuum state. We discuss this in detail in Appendix E.

Using the circuit in Eq. (1), in order to estimate $\text{Re Tr}[\rho e^{-ij\tau H}]$, where j is an arbitrary integer and τ is a real number, we set $W = I$. We introduce a random variable X_j and set it to be 1 when the measurement outcome is 0 and -1 when the measurement outcome is 1. Then,

$$\mathbb{E}[X_j] = \text{Re Tr}[\rho e^{-ij\tau H}]. \tag{2}$$

Similarly, for $\text{Im Tr}[\rho e^{-ij\tau H}]$, we set $W = S^\dagger$ and introduce a random variable Y_j that depends on the measurement outcome in the same way. We have

$$\mathbb{E}[Y_j] = \text{Im Tr}[\rho e^{-ij\tau H}]. \tag{3}$$

The parameter τ is chosen to normalize the Hamiltonian. Specifically, we choose τ so that $\tau \|H\| < \pi/3$. We remark that τ should be chosen to be $\mathcal{O}(\|H\|^{-1})$ and to avoid unnecessary overheads we want its scaling to be as close to $\Theta(\|H\|^{-1})$ as possible.

We can define a spectral measure of τH associated with ρ . The spectral measure is

$$p(x) = \sum_{k=0}^{K-1} p_k \delta(x - \tau \lambda_k), \quad x \in [-\pi, \pi]. \tag{4}$$

Here, K is the number of different eigenvalues, the λ_k 's are the distinct eigenvalues arranged in ascending order, and each p_k is the corresponding overlap, as defined in Sec. I. We extend it to a 2π -periodic function by $p(x + 2\pi) = p(x)$, so that the Fourier transform can be performed on the interval $[0, 2\pi]$ instead of the whole real line, which leads to a discrete Fourier spectrum. Note that because of the assumption $\tau \|H\| < \pi/3$, within the interval $[-\pi, \pi]$, $p(x)$ is supported in $(-\pi/3, \pi/3)$. Next, we consider the CDF associated with this measure.

We define the 2π -periodic Heaviside function by

$$H(x) = \begin{cases} 1, & x \in [2k\pi, (2k+1)\pi), \\ 0, & x \in [(2k-1)\pi, 2k\pi), \end{cases} \quad (5)$$

where $k \in \mathbb{Z}$. The CDF is usually defined by $C(x) = \sum_{k:\lambda_k \leq x} p_k$. This is, however, not a 2π -periodic function and thus will create technical difficulties in later discussions. Therefore, instead of the usual definition, we define

$$C(x) = (H * p)(x), \quad (6)$$

where the asterisk (*) denotes convolution. There is ambiguity at the jump discontinuities and we define the values of $C(x)$ at these points by requiring $C(x)$ to be right continuous. We check that this definition agrees with the usual definition when $x \in (-\pi/3, \pi/3)$, which is the interval that contains all the eigenvalues of τH :

$$\begin{aligned} C(x) &= \int_{-\pi}^{\pi} H(y)p(x-y)dy = \int_0^{\pi} p(x-y)dy \\ &= \int_{x-\pi}^x p(y)dy = \int_{-\pi}^x p(y)dy = \sum_{k:\lambda_k \leq x} p_k. \end{aligned}$$

Consequently $C(x)$ is a right-continuous nondecreasing function in $(-\pi/3, \pi/3)$.

If we could evaluate the CDF, then we would be able to locate the ground-state energy. This is because the CDF is a piecewise constant function. Each of its jumps in the interval $(-\pi/3, \pi/3)$ corresponds to an eigenvalue of τH . In order to find the ground-state energy, we only need to find where $C(x)$ jumps from zero to a nonzero value. However, in practice we cannot evaluate the CDF exactly. We will see that we are able to approximate—in a certain sense, as will be made clear later—the CDF using a function that we call the approximate CDF (ACDF). To this end, we first define an approximate Heaviside function $F(x) = \sum_{|j| \leq d} \hat{F}_j e^{ijx}$ such that

$$|F(x) - H(x)| \leq \epsilon, \quad x \in [-\pi + \delta, -\delta] \cup [\delta, \pi - \delta]. \quad (7)$$

The construction of this function is provided in Lemma 6, where \hat{F}_j is written as $\hat{F}_{d,\delta,j}$. Here, the parameters d and δ need to be chosen to control the accuracy of this approximation and their choices are discussed later. We also omit the d and δ dependence in the subscripts for simplicity. With this $F(x)$, we define the ACDF by

$$\tilde{C}(x) = (F * p)(x). \quad (8)$$

In Sec. III, we discuss how to evaluate this ACDF using the circuit in Eq. (1). The ACDF and CDF are related through

the following inequalities:

$$C(x - \delta) - \epsilon \leq \tilde{C}(x) \leq C(x + \delta) + \epsilon, \quad (9)$$

for any $|x| \leq \pi/3$, $0 < \delta < \pi/6$ and $\epsilon > 0$. We prove these inequalities in Appendix B. Given the statistical estimation of the ACDF $\tilde{C}(x)$, these inequalities enable us to estimate where the jumps of the CDF occur, which leads to an estimate of the ground-state energy.

By approximately evaluating the ACDF $\tilde{C}(x)$ for certain chosen x , and through Eq. (9), we can perform a binary search to locate the ground-state energy in smaller and smaller intervals. The algorithm to do this and the total computational cost required to estimate the ground-state energy to precision ϵ at a confidence level $1 - \vartheta$ are discussed in Secs. IV and V. A schematic illustration of our algorithm is shown in Fig. 2.

III. EVALUATING THE ACDF

In this section, we discuss how to evaluate the ACDF $\tilde{C}(x)$. We first expand it in the following way:

$$\begin{aligned} \tilde{C}(x) &= \sum_{|j| \leq d} \hat{F}_j \int_{-\pi}^{\pi} p(y) e^{ij(x-y)} dy \\ &= \sum_{|j| \leq d} \hat{F}_j e^{ijx} \text{Tr}[\rho e^{-ij\tau H}], \end{aligned} \quad (10)$$

where the spectral measure $p(x)$ is defined in Eq. (4). In going from the first line to the second line in the above equation, we use the fact that

$$\int_{-\pi}^{\pi} p(y) e^{-ijy} dy = \sum_{k=0}^{K-1} \text{Tr}[\rho \Pi_k] e^{-ij\tau \lambda_k} = \text{Tr}[\rho e^{-ij\tau H}].$$

One might want to evaluate each $\text{Tr}[\rho e^{-ij\tau H}]$ using Monte Carlo sampling, since this quantity is equal to $\mathbb{E}[X_j + iY_j]$. If we want to evaluate all $\text{Tr}[\rho e^{-ij\tau H}]$ to any accuracy at all, we need to sample each X_j and Y_j at least once. Then, the total evolution time is at least $\tau \sum_{|j| \leq d} |j| = \Omega(\tau d^2)$. Later, we will see that we need to choose $d = \mathcal{O}(\epsilon^{-1} \text{polylog}(\epsilon^{-1} p_0^{-1}))$ to ensure that the ground-state energy estimate has an additive error smaller than ϵ . Hence this total evolution time would give rise to a ϵ^{-2} dependence in the run time.

In order to avoid this ϵ^{-2} dependence, instead of evaluating all the terms we stochastically evaluate Eq. (10) as a whole. The idea we are going to describe is inspired by the unbiased version of the multilevel Monte Carlo method [69,70]. We define a random variable J that is drawn from

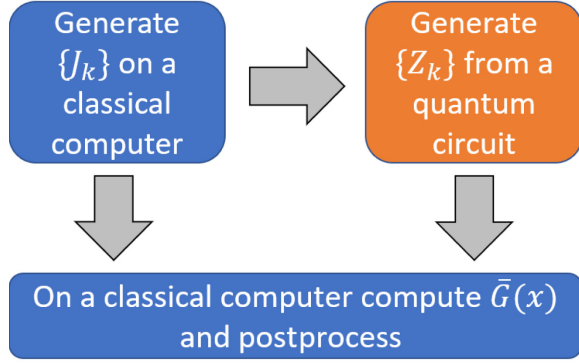


FIG. 2. An illustration of the classical and quantum components of our algorithm: (1) generate samples $\{J_k\}$ from Eq. (11); (2) use $\{J_k\}$ to generate $\{Z_k\}$ according to Eq. (13); (3) compute $\bar{G}(x)$ through Eq. (16). The ground-state energy estimate can be obtained through postprocessing as discussed in Sec. IV. Only Step (2) needs to be performed on a quantum computer.

$\{-d, -d+1, \dots, d\}$, with probability

$$\Pr[J = j] = \frac{|\hat{F}_j|}{\mathcal{F}}, \quad (11)$$

where the normalization factor $\mathcal{F} = \sum_{|j| \leq d} |\hat{F}_j|$. We let θ_j be the argument of \hat{F}_j , i.e., $\hat{F}_j = |\hat{F}_j|e^{i\theta_j}$. Then,

$$\begin{aligned} \mathbb{E}[(X_J + iY_J)e^{i(\theta_J + Jx)}] &= \sum_{|j| \leq d} \mathbb{E}[X_j + iY_j]e^{i(\theta_j + jx)} \Pr[J = j] \\ &= \frac{1}{\mathcal{F}} \sum_{|j| \leq d} \text{Tr}[\rho e^{-ij\tau H}] e^{ijx} \hat{F}_j \\ &= \frac{\tilde{C}(x)}{\mathcal{F}}, \end{aligned} \quad (12)$$

where we use Eqs. (2) and (3). For simplicity, we write X_J and Y_J into a complex random variable

$$Z = X_J + iY_J \in \{\pm 1 \pm i\}. \quad (13)$$

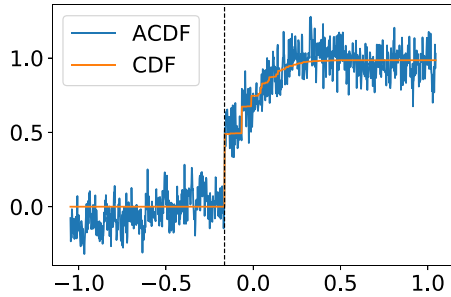


FIG. 3. $\bar{G}(x)$ and the CDF $C(x)$, for $x \in [-\pi/3, \pi/3]$ (left) and the enlarged view around $\tau\lambda_0$ (right), the ground-state energy for τH , where H is the Hamiltonian for the eight-site Hubbard model with $U/t = 4$ at half filling. The dashed vertical line is $x = \tau\lambda_0$. The parameters are $\delta = 2 \times 10^{-4}$, $d = 2 \times 10^4$, and $\tau = \pi/(4\|H\|)$. In total, 3000 samples are used.

Therefore, we can use

$$G(x; J, Z) = \mathcal{F} Z e^{i(\theta_J + Jx)} \quad (14)$$

as an unbiased estimate of $\tilde{C}(x)$. The variance can be bounded by

$$\text{var}[G(x)] \leq \mathcal{F}^2 \mathbb{E}[|X_J|^2 + |Y_J|^2] \leq 2\mathcal{F}^2. \quad (15)$$

Here, we use the fact that $|X_j|, |Y_j| \leq 1$.

From the above analysis, we can generate N_s independent samples of (J, Z) , denoted by (J_k, Z_k) , $k = 1, 2, \dots, N_s$, and then take the average

$$\bar{G}(x) = \frac{1}{N_s} \sum_{k=1}^{N_s} G(x; J_k, Z_k), \quad (16)$$

which can be used to estimate $\tilde{C}(x)$ in an unbiased manner. The variance is upper bounded by $2\mathcal{F}^2/N_s$. In order to make the variance upper bounded by a given σ^2 , we need $N_s = \mathcal{O}(\mathcal{F}^2/\sigma^2)$. The expected total evolution time is

$$N_s \tau \mathbb{E}[|J|] = \frac{\mathcal{F} \tau}{\sigma^2} \sum_{|j| \leq d} |\hat{F}_j| |j|.$$

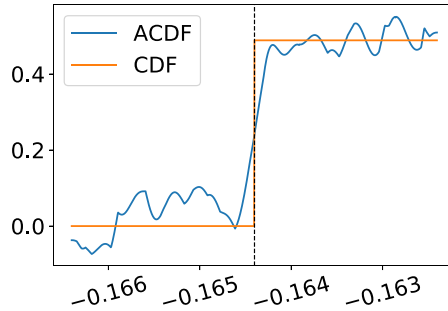
Furthermore, by Lemma 6 (iii), we have $|\hat{F}_j| \leq C|j|^{-1}$ for some constant C . Therefore,

$$\mathcal{F} = \mathcal{O}(\log(d)), \quad \sum_{|j| \leq d} |\hat{F}_j| |j| = \mathcal{O}(d).$$

The number of samples and the expected total evolution time are therefore

$$N_s = \mathcal{O}\left(\frac{\log^2(d)}{\sigma^2}\right), \quad N_s \tau \mathbb{E}[|J|] = \mathcal{O}\left(\frac{\tau d \log(d)}{\sigma^2}\right), \quad (17)$$

respectively. We can see that in this way we avoid the d^2 dependence, which shows up in a term-by-term evaluation.



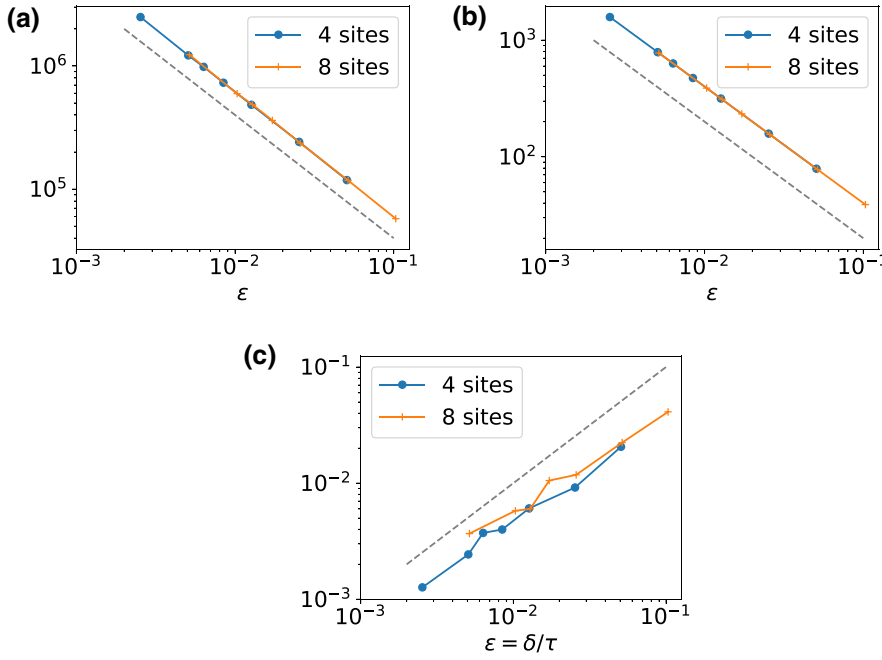


FIG. 4. The total evolution time (a), the maximal evolution time (b), and the average ground-state energy-estimation error (c) for the four-site and eight-site Hubbard model with $U/t = 4$ at half filling. The horizontal axis is the error threshold $\epsilon = \delta/\tau$. In (a) and (b), the gray dashed lines have slope -1 and in (c), the gray dashed line (with slope 1) shows the value of ϵ . For each δ , d is chosen to be $d = 4/\delta$, with 1800 samples, and $\tau = \pi/(4\|H\|)$. The maximal evolution time is $\tau d = 4\tau/\delta$.

In Fig. 3, we show the plot of the ACDF obtained through our method for the Fermi-Hubbard model. The details on this numerical experiment can be found in Appendix F. We can estimate the ground-state energy from the ACDF in a heuristic manner: we let

$$x^* = \inf\{x : \bar{G}(x) \geq \eta/2\}$$

and x^*/τ is an estimate for the ground-state energy λ_0 . Here, η is chosen so that $p_0 \geq \eta$. In Sec. V, we describe a more elaborate method to achieve the prescribed accuracy and confidence level. However, this heuristic method seems to work reasonably well in practice. In Fig. 4, we show the scaling of the ground-state energy-estimation error, the total evolution time, and the maximal evolution time, with respect to $\delta = \tau\epsilon$ (δ here is the parameter needed to construct $\{\hat{F}_j\}$ using Lemma 6), where ϵ is the allowed error. Both the total evolution time and the maximal evolution time are proportional to ϵ^{-1} . The details on this numerical experiment can also be found in Appendix F.

IV. ESTIMATING THE GROUND-STATE ENERGY

In this section, we discuss how to estimate the ground-state energy with a guaranteed error bound and confidence level from the samples generated on classical and quantum circuits discussed in Secs. II and III. First, we note that the CDF $C(x) = 0$ for all $-\pi/3 < x < \tau\lambda_0$, and $C(x) > 0$ for all $\tau\lambda_0 \leq x < \pi/3$. Therefore, getting the ground-state energy out of the CDF can be seen as inverting the CDF: we only need to find the smallest x such that $C(x) > 0$. One might consider performing a binary search to find such

a point but we run into a problem immediately: we only have access to estimates of $C(x)$ with statistical noise and we cannot tell if the estimate is greater than zero is due to $C(x) > 0$ or is merely due to statistical noise. We therefore need to make the search criterion more robust to noise.

Note that the CDF cannot take values between 0 and p_0 : $C(x) \geq p_0$ for $\tau\lambda_0 \leq x < \pi/3$ and $C(x) = 0$ for $-\pi/3 < x < \tau\lambda_0$. Now suppose that we know $p_0 \geq \eta$; then for any x , rather than distinguishing between $C(x) = 0$ and $C(x) > 0$, we instead distinguish between $C(x) = 0$ and $C(x) \geq \eta/2$ (here, $\eta/4$ is chosen to be consistent with later discussion and it can be any number between 0 and 1 times η). In this setting, if the estimate of $C(x)$ is larger than $\eta/4$, then we tend to believe that $C(x) \geq \eta/2$, and if the estimate is smaller than $\eta/4$, then we tend to believe that $C(x) = 0$. Thus we can tolerate an error that is smaller than $\eta/4$.

It may appear that we can find the ground-state energy by performing a binary search for the point at which $C(x)$ first becomes larger than $\eta/2$. However, we can only estimate the continuous function $\tilde{C}(x)$, which cannot uniformly approximate $C(x)$. This is because $C(x)$ has many jump discontinuities (each of which corresponds to an eigenvalue). As a result, we cannot perform this binary-search procedure directly.

From the above discussion, we need a search criterion that can be checked via $\tilde{C}(x)$. We consider the following criterion:

Problem 1 (Inverting the CDF): *For $0 < \delta < \pi/6$, $0 < \eta < 1$, find $x^* \in (-\pi/3, \pi/3)$ such that*

$$C(x^* + \delta) > \eta/2, \quad C(x^* - \delta) < \eta. \quad (18)$$

First, we verify that this can be checked via $\tilde{C}(x)$. In Eq. (9), if we choose $x = x^*$, $\epsilon = \eta/6$, then $\tilde{C}(x^*) > (2/3)\eta$ implies $C(x^*) > \eta/2$ and $\tilde{C}(x^*) < (5/6)\eta$ implies $C(x^*) < \eta$. Therefore, we only need to find x^* satisfying $(2/3)\eta < C(x^*) < (5/6)\eta$ to satisfy this criterion. Second, we show that an x^* satisfying this criterion gives us an estimate of the ground-state energy to within additive error δ/τ . Suppose that we choose $\eta > 0$ so that $p_0 \geq \eta$. Then, if we solve Problem 1, we will find an x^* such that $C(x^* + \delta) > \eta/2 > 0$ and $C(x^* - \delta) < \eta \leq p_0$. $C(x^* + \delta) > 0$ indicates that $x^* + \delta \geq \tau\lambda_0$. Since $C(x)$ cannot take a value between 0 and p_0 , $C(x^* - \delta) < p_0$ indicates $C(x^* - \delta) = 0$, and thus $x^* - \delta < \tau\lambda_0$. Hence we know that $|x^* - \tau\lambda_0| \leq \delta$. If we choose $\delta = \tau\epsilon$ and $\tilde{\lambda}_0 = x^*/\tau$, then

$$|\tilde{\lambda}_0 - \lambda_0| \leq \epsilon.$$

Then, $\tilde{\lambda}_0$ is our desired estimate.

Note that Eq. (18) is a weaker requirement than $\eta/2 < C(x^*) < \eta$, for which due to the discontinuity of $C(x)$ the required x^* may not exist. However, an x^* satisfying Eq. (18) must exist. In fact, let $a = \sup\{x \in (-\pi/3, \pi/3) : C(x) \leq \eta/2\}$ and $b = \inf\{x \in (-\pi/3, \pi/3) : C(x) \geq \eta\}$. Then, because $C(x)$ is monotonously increasing, $a \leq b$, and any $x^* \in [a - \delta, b + \delta]$ satisfies Eq. (18).

Using the samples $\{J_k\}$ and $\{Z_k\}$ generated on classical and quantum circuits, respectively, we are able to solve Problem 1.

Theorem 2 (Inverting the CDF): *With samples $\{J_k\}_{k=1}^M$ satisfying $|J_k| \leq d$ and $\{Z_k\}_{k=1}^M$, generated according to Eqs. (11) and (13), respectively, we can solve Problem 1 on a classical computer with probability at least $1 - \vartheta$, for $d = \mathcal{O}(\delta^{-1} \log(\delta^{-1} \eta^{-1}))$ and $M = \mathcal{O}(\eta^{-2} \log^2(d) (\log \log(\delta^{-1}) + \log(\vartheta^{-1})))$. The classical postprocessing cost is*

$$\tilde{\mathcal{O}}(\eta^{-2} \log^3(\delta^{-1}) \log(\vartheta^{-1})). \quad (19)$$

To generate the samples $\{Z_k\}_{k=1}^M$ on a quantum circuit, the expected total evolution time and the maximal evolution time are

$$\tau M \mathbb{E}[|J|] = \tilde{\mathcal{O}}(\tau \delta^{-1} \eta^{-2} \log(\vartheta^{-1})), \quad (20)$$

and

$$\tau d = \mathcal{O}(\tau \delta^{-1} \log(\delta^{-1} \eta^{-1})), \quad (21)$$

respectively.

The meanings of the symbols used in this theorem can be found in Table III in Appendix G. We prove this theorem by constructing the algorithm for classical postprocessing in Sec. V. Since solving Problem 1 enables us to estimate

the ground-state energy as discussed above, from Theorem 2 we have the following corollary:

Corollary 3 (Ground-state energy): *With samples $\{J_k\}_{k=1}^M$ satisfying $|J_k| \leq d$ and $\{Z_k\}_{k=1}^M$, generated according to Eqs. (11) and (13), respectively, we can estimate the ground-state energy λ_0 to within additive error ϵ on a classical computer with probability at least $1 - \vartheta$, if $p_0 \geq \eta$ for some known η , $d = \mathcal{O}(\epsilon^{-1} \tau^{-1} \log(\epsilon^{-1} \tau^{-1} \eta^{-1}))$, and $M = \mathcal{O}(\eta^{-2} \log^2(d) (\log \log(\epsilon^{-1} \tau^{-1}) + \log(\vartheta^{-1})))$. The classical postprocessing cost is $\mathcal{O}(\eta^{-2} \text{polylog}(\epsilon^{-1} \tau^{-1} \eta^{-1}))$. The expected total evolution time and the maximal evolution time are $\mathcal{O}(\epsilon^{-1} \eta^{-2} \text{polylog}(\epsilon^{-1} \tau^{-1} \eta^{-1}))$ and $\mathcal{O}(\epsilon^{-1} \text{polylog}(\epsilon^{-1} \tau^{-1} \eta^{-1}))$, respectively.*

Usually, the Heisenberg limit is defined in terms of the root-mean-square error (RMSE) of the estimate. In this paper, we focus on ensuring the error of the ground-state energy to be below a threshold ϵ with probability at least $1 - \vartheta$. From Corollary 3, our algorithm only has a logarithmic dependence on ϑ^{-1} and the error can be at most $2\|H\|$. We can easily ensure that the RMSE is $\mathcal{O}(\epsilon)$ using the result by choosing $\vartheta = \mathcal{O}(\epsilon^2 \|H\|^{-2})$. We can see that the total evolution-time scaling with respect to ϵ is still $\tilde{\mathcal{O}}(\epsilon^{-1})$.

Remark 4 (System-size dependence): *One might note the absence of an explicit system-size dependence in the evolution-time scaling in Theorem 2 and Corollary 3. This is because, as mentioned before in Sec. I, the total evolution time depends on the system size indirectly through two parameters, τ and η . Moreover, if we consider the dependence of the total run time on the system size, we also need to account for the overhead that comes from performing Hamiltonian simulation. This overhead and the scaling of η with respect to the system size can be highly problem-specific and are independent from the tasks we are considering in this paper. Because the Hamiltonian norm can generally be upper bounded by a polynomial of the system size and the total evolution-time dependence on τ^{-1} is polylogarithmic, τ contributes a polylogarithmic overhead in the system-size dependence.*

V. INVERTING THE CDF

In this section, we prove Theorem 2 by constructing the classical postprocessing algorithm to solve Problem 1 using samples from a quantum circuit. Since we want to search for an x^* satisfying the requirement in Eq. (18), a natural idea is to use binary search. Our setting is somewhat different from the usual binary-search setting but we show that a similar approach still works. The current setting differs from the setting of binary search mainly in two ways: first, any $x^* \in [\tau\lambda_0 - \delta, \tau\lambda_0 + \delta]$ satisfies the requirement Eq. (18) and can therefore be a target. When

performing binary search, we want to be able to tell if the target is to the left or right of a given x but here the targets may be on both sides of x . When this happens, there is some uncertainty as to how the algorithm will proceed next. However, in our algorithm we show that this does not present a problem. Also, because this algorithm is based on random samples, there is some failure probability in each search step. We use a majority-voting procedure to suppress the failure probability, so that in the end the algorithm will produce a correct answer with probability arbitrarily close to 1.

We suppose that we are given independent samples of (J, Z) defined in Eqs. (11) and (13) generated from a quantum circuit. We denote these samples by $\{(J_k, Z_k)\}_{k=1}^M$. We divide them into N_b batches of size N_s , where $N_s N_b = M$. This division is for the majority-voting procedure we mentioned above. The maximal evolution time needed to generate these samples is proportional to $\max_k |J_k| \leq d$. The expected total evolution time we will need is proportional to $M\mathbb{E}[|J|]$.

We first reduce Problem 1 into a decision problem. For any $x \in (-\pi/3, \pi/3)$, one of the following must be true:

$$C(x + \delta) > \eta/2, \quad \text{or} \quad C(x - \delta) < \eta. \quad (22)$$

If there is a subroutine that tells us which one of the two is correct, or randomly picks one when both are correct, then we can use it to find x^* . We assume that such a subroutine, which uses $\{(J_k, Z_k)\}_{k=1}^M$, exists and denote it by the name $\text{CERTIFY}(x, \delta, \eta, \{(J_k, Z_k)\})$. The subroutine returns either 0 or 1: 0 for $C(x + \delta) > \eta/2$ being true and 1 for $C(x - \delta) < \eta$ being true.

In Algorithm 1, with $\text{CERTIFY}(x, \delta, \eta, \{(J_k, Z_k)\})$, we describe the algorithm to solve Problem 1. We denote this algorithm by $\text{INVERT_CDF}(\delta, \eta, \{(J_k, Z_k)\})$. It runs as follows. We start with $x_{0,0} = -\pi/3$ and $x_{1,0} = \pi/3$. They are chosen so that $C(x_{1,0}) > \eta/2$ and $C(x_{0,0}) < \eta$. Let ℓ be the number of iterations that we perform and let $\ell = 0$ at the beginning. At each iteration, we let $x_\ell = (x_{0,\ell} +$

$x_{1,\ell})/2$ and run $\text{CERTIFY}(x_\ell, (2/3)\delta, \eta, \{(J_k, Z_k)\})$. This tells us either $C(x_\ell + (2/3)\delta) > \eta/2$ or $C(x_\ell - (2/3)\delta) < \eta$. If the former, then we let $x_{0,\ell+1} = x_{0,\ell}$ and $x_{1,\ell+1} = x_\ell + (2/3)\delta$, and if the latter, we let $x_{0,\ell+1} = x_\ell + (2/3)\delta$ and $x_{1,\ell+1} = x_{1,\ell}$. This is done so that for each ℓ , we have

$$C(x_{0,\ell}) < \eta, \quad C(x_{1,\ell}) > \eta/2. \quad (23)$$

We then let $\ell \leftarrow \ell + 1$ and go to the next iteration. The algorithm stops once $x_{1,\ell} - x_{0,\ell} \leq 2\delta$. We denote the total number of iterations by L . The output is $x_L = (x_{0,L} + x_{1,L})/2$. Because Eq. (23) holds for each iteration, we have

$$C(x_L - \delta) \leq C(x_{0,L}) < \eta, \quad C(x_L + \delta) \geq C(x_{1,L}) > \eta/2.$$

Thus we can see that x_L satisfies the requirements for x^* in Problem 1. The next question is, how many iterations does it take to satisfy the stopping criterion? Regardless of the outcome of the CERTIFY subroutine, we always have

$$x_{1,\ell+1} - x_{0,\ell+1} = \frac{1}{2}(x_{1,\ell} - x_{0,\ell}) + \frac{2}{3}\delta.$$

From this, we can see that

$$x_{1,\ell} - x_{0,\ell} = \frac{2\pi/3 - (4/3)\delta}{2^\ell} + \frac{4}{3}\delta.$$

Therefore, it takes $L = \mathcal{O}(\log(\delta^{-1}))$ iterations for the algorithm to stop.

Next, we discuss how to construct the subroutine $\text{CERTIFY}(x, \delta, \eta, \{(J_k, Z_k)\})$. While we cannot directly evaluate the CDF $C(x)$ for any x , we can estimate the ACDF $\tilde{C}(x)$ using the data $\{J_k\}$ and $\{Z_k\}$. We can let $\epsilon = \eta/8$ in Eq. (7) and choose $d = \mathcal{O}(\delta^{-1} \log(\delta^{-1} \eta^{-1}))$ according to Lemma 6. Then, by Eq. (9), we have $C(x - \delta) \leq \tilde{C}(x) + \eta/8$ and $C(x + \delta) \geq \tilde{C}(x) - \eta/8$. One of the following must be true:

$$\tilde{C}(x) > (5/8)\eta, \quad \text{or} \quad \tilde{C}(x) < (7/8)\eta, \quad (24)$$

then the former implies $C(x + \delta) > \eta/2$ and the latter $C(x - \delta) < \eta$. Therefore, the CERTIFY subroutine only needs to decide which one of the two is correct or to output a random choice when both are correct.

As discussed in Sec. III, $\bar{G}(x)$ is an unbiased estimate of $\tilde{C}(x)$. We use $\{J_k\}$ and $\{Z_k\}$ to obtain N_b samples for $\bar{G}(x)$, denoted by $\bar{G}_r(x)$, via

$$\bar{G}_r(x) = \frac{1}{N_s} \sum_{k=1}^{N_s} G(x; J_{(r-1)N_s+k}, Z_{(r-1)N_s+k}),$$

for $r = 1, 2, \dots, N_b$. Here, $G(x; J, Z)$ is defined in Eq. (14). For each r , we compare $\bar{G}_r(x)$ with $(3/4)\eta$. If $\bar{G}_r(x) > (3/4)\eta$ for a majority of batches, then we tend to

Input: $\delta, \eta, \{(J_k, Z_k)\}$
 $x_0 \leftarrow -\pi/3, x_1 \leftarrow \pi/3;$
while $x_1 - x_0 > 2\delta$ **do**
 $x \leftarrow (x_0 + x_1)/2;$
 $u \leftarrow \text{CERTIFY}(x, (2/3)\delta, \eta, \{(J_k, Z_k)\});$
if $u = 0$ **then**
 $x_1 \leftarrow x + (2/3)\delta;$
else
 $x_0 \leftarrow x - (2/3)\delta;$
end if
end while
Output: $(x_0 + x_1)/2$

Algorithm 1. INVERT_CDF .

Input: $x, \delta, \eta, \{(J_k, Z_k)\}$
 $b \leftarrow 0, c \leftarrow 0;$
for $r = 1, 2, \dots, N_b$ **do**
 $\bar{G}_r(x) \leftarrow (1/N_s) \sum_{k=1}^{N_s} G(x; J_{(r-1)N_s+k}, Z_{(r-1)N_s+k});$
 $\{G(x; J, Z) \text{ defined in (14)}\}$
if $\bar{G}_r(x) > (3/4)\eta$ **then**
 $c \leftarrow c + 1;$
end if
end for
if $c \leq B/2$ **then**
 $b \leftarrow 1;$
end if
Output: b

Algorithm 2. CERTIFY.

believe $\tilde{C}(x) > (5/8)\eta$ and output 0 for $C(x + \delta) > \eta/2$. Otherwise, we tend to believe that $\tilde{C}(x) < (7/8)\eta$ and output 1 for $C(x - \delta) < \eta$. This is the majority-voting procedure that we mentioned earlier. For the pseudocode for the subroutine, see Algorithm 2.

In the CERTIFY subroutine, an error occurs when $\tilde{C}(x) > (5/8)\eta$ yet a majority of estimates $\bar{G}_r(x)$ are smaller than $(3/4)\eta$, or when $\tilde{C}(x) < (7/8)\eta$ yet a majority of estimates $\bar{G}_r(x)$ are larger than $(3/4)\eta$. We need to make the probability of this kind of error occurring upper bounded by ν . First, we assume that $\tilde{C}(x) > (5/8)\eta$. Then, for each r , by Markov's inequality, we have

$$\Pr[\bar{G}_r(x) < (3/4)\eta] \leq \frac{64 \text{var}[\bar{G}_r(x)]}{\eta^2}.$$

We want to make this probability at most $1/4$. Therefore, we need $\text{var}[\bar{G}_r(x)] \leq \eta^2/256$. To ensure this, by Eq. (17), in which we let $\sigma^2 = \eta^2/256$, we can choose

$$N_s = \mathcal{O}\left(\frac{\log^2(d)}{\eta^2}\right). \quad (25)$$

Then, by the Chernoff bound, the probability of the majority of estimates $\bar{G}_r(x)$ being smaller than $(3/4)\eta$ is at most $e^{-C'N_b}$ for some constant C' . In order to make this probability bounded by ν , we only need to let $N_b = \mathcal{O}(\log(\nu^{-1}))$.

In the algorithm INVERT_CDF, the subroutine CERTIFY is used $L = \mathcal{O}(\log(\delta^{-1}))$ times. If an error occurs in a single run of CERTIFY with probability at most ν , then in the total L times we use this subroutine, the probability of an error occurring is at most $L\nu$. Therefore, in order to ensure that an error occurs with probability at most ϑ in INVERT_CDF, we need to set $\nu = \vartheta/L$. Therefore, $N_b = \mathcal{O}(\log(L\vartheta^{-1})) = \mathcal{O}(\log \log(\delta^{-1}) + \log(\vartheta^{-1}))$.

The above analysis shows that in order to solve Problem 1, the total evolution time is $M\mathbb{E}[|J|] = N_b N_s \mathbb{E}[|J|]$.

We evaluate $N_s \mathbb{E}[|J|]$ by Eq. (17), in which we let $\sigma^2 = \eta^2/256$ as discussed before when we estimate how large N_s needs to be in Eq. (25). Multiplying this by N_b , we have Eq. (20). Note here that we do not need to multiply by L because in each CERTIFY subroutine, we can reuse the same $\{J_k\}, \{Z_k\}$. The maximal evolution time required is τd and this leads to Eq. (21). The main cost in classical postprocessing comes from evaluating $\bar{G}_r(x)$. This needs to be done LN_b times. Each evaluation involves $\mathcal{O}(N_s) = \mathcal{O}(\eta^{-2} \log^2(d))$ arithmetic operations. The total run time for classical postprocessing is therefore $LN_b N_s = LM$, which leads to Eq. (19). Thus we obtain all the cost estimates in Theorem 2 and prove the theorem.

VI. DISCUSSION AND CONCLUSIONS

In this paper, we present an algorithm to estimate the ground-state energy with Heisenberg-limited precision scaling. The quantum circuit that we use requires only one ancilla qubit and the maximal evolution time needed per run has a polylogarithmic dependence on the overlap p_0 . Such dependence on p_0 is exponentially better than that required by QPE using a similarly structured circuit using the semiclassical Fourier transform, as discussed in Sec. I A. Both rigorous analysis and numerical experiments are done to validate the correctness and efficiency of our algorithm.

Although our algorithm has a near-optimal dependence on the precision, the dependence on p_0 (more precisely, on its lower bound η), which scales as p_0^{-2} in Corollary 3, is far from optimal compared to the $p_0^{-1/2}$ scaling in Refs. [23,24]. Whether one can achieve this $p_0^{-1/2}$ scaling without using a quantum circuit with substantially larger maximal evolution time, and without using such techniques as LCU or block encoding, remains an open question.

The probabilistic choice of the simulation time according to Eq. (11) plays an important role in reducing the total evolution time. However, we may partially derandomize the algorithm following the spirit of the multilevel Monte Carlo (MLMC) method [71] in the classical setting. The method that we develop for computing the approximate CDF in Sec. III is, in fact, a quite general approach for evaluating expectation values from matrix functions. This method can act as a substitute of the LCU method in many cases, especially in a near-term setting. The use of this method to compute other properties of the spectrum, such as the spectral density, is a direction for future work.

ACKNOWLEDGMENTS

This work was partially supported by the Air Force Office of Scientific Research under Award No. FA9550-18-1-0095 (L.L. and Y.T.) and by the Department of Energy under Grant No. DE-SC0017867 and under the

Quantum Systems Accelerator program (L.L.). We thank Andrew Baczewski and Barbara Terhal for helpful discussions. L.L. is a Simons Investigator.

APPENDIX A: CONSTRUCTING THE APPROXIMATE HEAVISIDE FUNCTION

In this appendix, we construct the approximate Heaviside function satisfying the requirement in Eq. (7). We first need to construct a smeared Dirac function, which we use as a mollifier in constructing the approximate Heaviside function. To the best of our knowledge, this particular version of the smeared Dirac function has not been proposed in previous works.

Lemma 5: We define $M_{d,\delta}(x)$ by

$$M_{d,\delta}(x) = \frac{1}{\mathcal{N}_{d,\delta}} T_d \left[1 + 2 \frac{\cos(x) - \cos(\delta)}{1 + \cos(\delta)} \right],$$

where $T_d(x)$ is the d th Chebyshev polynomial of the first kind, and

$$\mathcal{N}_{d,\delta} = \int_{-\pi}^{\pi} T_d \left[1 + 2 \frac{\cos(x) - \cos(\delta)}{1 + \cos(\delta)} \right] dx.$$

Then:

- (i) $|M_{d,\delta}(x)| \leq (1/\mathcal{N}_{d,\delta})$ for all $x \in [-\pi, -\delta] \cup [\delta, \pi]$, and $M_{d,\delta}(x) \geq -(1/\mathcal{N}_{d,\delta})$ for all $x \in \mathbb{R}$.
- (ii) $\int_{-\pi}^{\pi} M_{d,\delta}(x) dx = 1$, $1 \leq \int_{-\pi}^{\pi} |M_{d,\delta}(x)| dx \leq 1 + (4\pi/\mathcal{N}_{d,\delta})$.
- (iii) When $\tan(\delta/2) \leq 1 - 1/\sqrt{2}$, we have

$$\mathcal{N}_{d,\delta} \geq C_1 e^{d\delta/\sqrt{2}} \sqrt{\frac{\delta}{d}} \operatorname{erf}(C_2 \sqrt{d\delta}),$$

for some constants C_1 and C_2 that do not depend on d or δ .

Proof. We first note that, by the property of Chebyshev polynomials, when $x \in [-\pi, -\delta] \cup [\delta, \pi]$, i.e., $\cos(x) \leq \cos(\delta)$, we have $|T_d(1 + 2\{[\cos(x) - \cos(\delta)]/[1 + \cos(\delta)]\})| \leq 1$. This proves the first inequality in (i). Note that when $x \in [-\delta, \delta]$, $T_d(1 + 2\{[\cos(x) - \cos(\delta)]/[1 + \cos(\delta)]\}) \geq -1$. Combine this inequality, the first inequality in (i), and the fact that $M_{d,\delta}(x)$ is 2π periodic we prove the second inequality in (i).

The first part of (ii) is obvious because of the definition of $\mathcal{N}_{d,\delta}$. For the second part, we have $\int_{-\pi}^{\pi} |M_{d,\delta}(x)| dx \geq \int_{-\pi}^{\pi} M_{d,\delta}(x) dx = 1$. Also,

$$\begin{aligned} \int_{-\pi}^{\pi} |M_{d,\delta}(x)| dx &= \left(\int_{-\pi}^{-\delta} + \int_{\delta}^{\pi} \right) |M_{d,\delta}(x)| dx \\ &+ \int_{-\delta}^{\delta} M_{d,\delta}(x) dx \leq \frac{4\pi}{\mathcal{N}_{d,\delta}} \end{aligned}$$

$$\begin{aligned} &+ \left(\int_{-\pi}^{-\delta} + \int_{\delta}^{\pi} \right) M_{d,\delta}(x) dx \\ &+ \int_{-\delta}^{\delta} M_{d,\delta}(x) dx = 1 + \frac{4\pi}{\mathcal{N}_{d,\delta}}. \end{aligned} \quad (\text{A1})$$

We now prove (iii). This requires lower bounding $T_d(1 + 2\{[\cos(x) - \cos(\delta)]/[1 + \cos(\delta)]\})$ when $x \in [-\delta, \delta]$. For δ small enough so that

$$\max_x 2 \frac{\cos(x) - \cos(\delta)}{1 + \cos(\delta)} = 2 \tan^2(\delta/2) \leq 3 - \sqrt{2},$$

which is equivalent to $\tan(\delta/2) \leq 1 - 1/\sqrt{2}$, we can use Ref. [72, Lemma 13] to provide a lower bound for the $x \in [-\delta, \delta]$ case:

$$\begin{aligned} &T_d \left[1 + 2 \frac{\cos(x) - \cos(\delta)}{1 + \cos(\delta)} \right] \\ &\geq \frac{1}{2} \exp \left[\sqrt{2}d \sqrt{\frac{\cos(x) - \cos(\delta)}{1 + \cos(\delta)}} \right]. \end{aligned} \quad (\text{A2})$$

By the elementary inequality $|\sin(x)| \leq |x|$, we have

$$\begin{aligned} \sqrt{\frac{\cos(x) - \cos(\delta)}{1 + \cos(\delta)}} &= \sqrt{\tan^2\left(\frac{\delta}{2}\right) - \frac{\sin^2(x/2)}{\cos^2(\delta/2)}} \\ &= \tan\left(\frac{\delta}{2}\right) \sqrt{1 - \frac{\sin^2(x/2)}{\sin^2(\delta/2)}} \\ &\geq \tan\left(\frac{\delta}{2}\right) \left(1 - \frac{\sin^2(x/2)}{\sin^2(\delta/2)}\right) \\ &\geq \tan\left(\frac{\delta}{2}\right) \left(1 - \frac{x^2}{4 \sin^2(\delta/2)}\right). \end{aligned}$$

Substituting this into Eq. (A2), we have

$$\begin{aligned} &T_d \left[1 + 2 \frac{\cos(x) - \cos(\delta)}{1 + \cos(\delta)} \right] \\ &\geq \frac{1}{2} e^{\sqrt{2}d \tan(\delta/2)} \exp \left[-\frac{dx^2}{\sqrt{2} \sin(\delta)} \right]. \end{aligned} \quad (\text{A3})$$

Then,

$$\begin{aligned} \mathcal{N}_{d,\delta} &\geq \int_{-\delta}^{\delta} T_d \left[1 + 2 \frac{\cos(x) - \cos(\delta)}{1 + \cos(\delta)} \right] dx - 2\pi \\ &\geq \frac{1}{2} e^{\sqrt{2}d \tan(\delta/2)} \sqrt{\frac{\sqrt{2}\pi \sin(\delta)}{d}} \operatorname{erf} \left(\sqrt{\frac{d}{\sqrt{2} \sin(\delta)}} \delta \right) \\ &\quad - 2\pi \geq C_1 e^{d\delta/\sqrt{2}} \sqrt{\frac{\delta}{d}} \operatorname{erf}(C_2 \sqrt{d\delta}), \end{aligned}$$

for $\delta \in (0, \pi/2)$ and some constants $C_1, C_2 > 0$. This proves (iii). \blacksquare

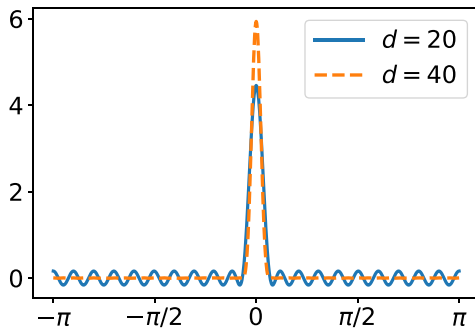


FIG. 5. Illustration of $M_{d,\delta}(x)$ for $\delta = 0.2, d = 20, 40$.

A plot of $M_{d,\delta}$ is shown in Fig. 5. As we can see, it roughly takes the shape of a Dirac function. We then use it as a mollifier to approximate the Heaviside function using the convolution of $M_{d,\delta}$ and the Heaviside function.

Lemma 6: *Let $H(x)$ be the periodic Heaviside function defined in Eq. (5). For any $\delta \in (0, \pi/2)$ such that $\tan(\delta/2) \leq 1 - 1/\sqrt{2}$ and $\epsilon > 0$, there exists $d = \mathcal{O}(\delta^{-1} \log(\delta^{-1}\epsilon^{-1}))$ and a 2π -periodic function $F_{d,\delta}(x)$ of the form*

$$F_{d,\delta}(x) = \frac{1}{\sqrt{2\pi}} \sum_{k=-d}^d \hat{F}_{d,\delta,k} e^{ikx},$$

satisfying

- (i) $-\epsilon/2 \leq F_{d,\delta}(x) \leq 1 + \epsilon$ for all $x \in \mathbb{R}$
- (ii) $|F_{d,\delta}(x) - H(x)| \leq \epsilon$ for all $x \in [-\pi + \delta, -\delta] \cup [\delta, \pi - \delta]$
- (iii) $|\hat{F}_{d,\delta,k}| \leq 2(1 + \epsilon)/(\sqrt{2\pi}|k|)$ for $k \neq 0$.

Proof. We first construct the function $F_{d,\delta}(x)$. Let $M_{d,\delta}(x)$ be the mollifier in Lemma 5. Because of Lemma 5 (i) and (ii), $M_{d,\delta}(x)$ can be used as to mollify nonsmooth functions. Also, because $T_d(x)$ is a polynomial of degree d , the Fourier coefficients

$$\hat{M}_{d,\delta,k} = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} M_{d,\delta}(x) e^{-ikx} dx$$

are nonzero only for $-d \leq k \leq d$. Also,

$$|\hat{M}_{d,\delta,k}| \leq \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} |M_{d,\delta}(x)| dx = \frac{1 + \epsilon}{\sqrt{2\pi}}. \quad (\text{A4})$$

We construct $F_{d,\delta}$ by mollifying the Heaviside function with $M_{d,\delta}(x)$:

$$F_{d,\delta}(x) = (M_{d,\delta} * H)(x) = \int_{-\pi}^{\pi} M_{d,\delta}(x') H(x - x') dx'. \quad (\text{A5})$$

We then show we can choose $d = \mathcal{O}(\delta^{-1} \log(\delta^{-1}\epsilon^{-1}))$ to satisfy (ii). We have

$$\begin{aligned} |F_{d,\delta}(x) - H(x)| &= \left| \int_{-\pi}^{\pi} M_{d,\delta}(x') (H(x - x') - H(x)) dx' \right| \\ &\leq \int_{-\pi}^{\pi} M_{d,\delta}(x') |H(x - x') - H(x)| dx'. \end{aligned}$$

For any x such that $|x| \in [\delta, \pi - \delta]$, we first consider the case where $|x'| < \delta$. In this case, $H(x - x') = H(x)$ and therefore the integrand $M_{d,\delta}(x') |H(x - x') - H(x)| = 0$. Then, we consider the case where $|x'| \geq \delta$. By Lemma 5 (i), we have $M_{d,\delta}(x') \leq 2/\mathcal{N}_{d,\delta}$ and as $|H(x - x') - H(x)| \leq 1$, $M_{d,\delta}(x') |H(x - x') - H(x)| \leq 2/\mathcal{N}_{d,\delta}$. Thus for any x such that $|x| \in [\delta, \pi - \delta]$,

$$|F_{d,\delta}(x) - H(x)| \leq \frac{4\pi}{\mathcal{N}_{d,\delta}}. \quad (\text{A6})$$

If we want to keep the approximation error for $x \in [-\pi + \delta, -\delta] \cup [\delta, \pi - \delta]$ to be below ϵ , we will need, by Lemma 5 (i) and Eq. (A5),

$$C_1 e^{d\delta/\sqrt{2}} \sqrt{\frac{\delta}{d}} \operatorname{erfc}(C_2 \sqrt{d\delta}) \geq \frac{4\pi}{\epsilon}.$$

It can be checked that we can choose $d = \mathcal{O}(\delta^{-1} \log(\epsilon^{-1}\delta^{-1}))$ to achieve this.

We then show that this choice of d ensures (i) as well. From Eq. (A1),

$$F_{d,\delta}(x) \leq \int_{-\pi}^{\pi} |M_{d,\delta}(y)| dy \leq 1 + \frac{4\pi}{\mathcal{N}_{d,\delta}} \leq 1 + \epsilon$$

and by the second inequality in Lemma 5 (i),

$$F_{d,\delta}(x) \geq -\frac{1}{\mathcal{N}_{d,\delta}} \int_{-\pi}^{\pi} H(y) dy = -\frac{2\pi}{\mathcal{N}_{d,\delta}} \geq -\frac{\epsilon}{2}.$$

Finally, we prove that our construction satisfies (iii). Because $F_{d,\delta}(x)$ is defined through a convolution, its Fourier coefficients can be obtained through

$$\hat{F}_{d,\delta,k} = \sqrt{2\pi} \hat{M}_{d,\delta,k} \hat{H}_k,$$

where \hat{H}_k 's are the Fourier coefficients of the rectangle function $H(x)$. Therefore, $\hat{F}_{d,\delta,k} \neq 0$ only for $-d \leq k \leq d$.

Because of Eq. (A3), we have

$$|\hat{F}_{d,\delta,k}| \leq (1 + \epsilon)|\hat{H}_k|.$$

Since when $k \neq 0$,

$$\begin{aligned} \hat{H}_k &= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} H(x) e^{-ikx} dx \\ &= \begin{cases} \frac{2}{i\sqrt{2\pi}k} & 2 \nmid k \\ 0 & 2 \mid k, \end{cases} \end{aligned}$$

we have (iii). ■

APPENDIX B: THE RELATION BETWEEN THE CDF AND THE APPROXIMATE CDF

In this appendix, we prove Eq. (9). Let $0 < \delta < \pi/6$. First, we have a 2π -periodic function $F(x)$ from Lemma 6 that satisfies

$$|F(x) - H(x)| \leq \epsilon, \quad x \in [-\pi + \delta, -\delta] \cup [\delta, \pi - \delta],$$

and $F(x) \in [0, 1]$ for all $x \in \mathbb{R}$ (In Appendix A we have $F(x) \in [-\epsilon/2, 1 + \epsilon]$. Here we slightly shift and rescale $F(x)$ so that $F(x) \in [0, 1]$ for the discussion in this appendix.). We further define $F_L(x) = F(x - \delta)$ and $F_R(x) = F(x + \delta)$. They satisfy

$$\begin{aligned} |F_L(x) - H(x)| &\leq \epsilon, \quad x \in [-\pi + 2\delta, 0] \cup [2\delta, \pi], \\ |F_R(x) - H(x)| &\leq \epsilon, \quad x \in [-\pi, -2\delta] \cup [0, \pi - 2\delta]. \end{aligned} \quad (\text{B1})$$

We define some functions related to the ACDF as follows:

$$\tilde{C}_L(x) = (F_L * p)(x), \quad \tilde{C}_R(x) = (F_R * p)(x). \quad (\text{B2})$$

Then, we have

$$\tilde{C}_L(x) = \tilde{C}(x - \delta), \quad \tilde{C}_R(x) = \tilde{C}(x + \delta). \quad (\text{B3})$$

The functions $\tilde{C}_L(x)$ and $\tilde{C}_R(x)$ can be used to bound $C(x)$. Because of Eq. (B1), the fact that $p(x)$ is supported in $(-\pi/3, \pi/3)$ in $[-\pi, \pi]$, $\delta < \pi/6$, and that $H(y)$ and $F_L(y)$ both take value in $[0, 1]$, for $x \in (-\pi/3, \pi/3)$, we have

$$\begin{aligned} |\tilde{C}_L(x) - C(x)| &\leq \int_{-\pi}^{\pi} p(x-y)|H(y) - F_L(y)| dy \\ &\leq \epsilon + \int_0^{2\delta} p(x-y) dy \\ &= \epsilon + C(x) - C(x - 2\delta). \end{aligned}$$

Therefore,

$$\tilde{C}_L(x) \geq C(x) - [\epsilon + C(x) - C(x - 2\delta)] = C(x - 2\delta) - \epsilon.$$

Similarly, we have

$$\tilde{C}_R(x) \leq C(x) + [\epsilon + C(x + 2\delta) - C(x)] = C(x + 2\delta) + \epsilon.$$

Combining these two inequalities with Eq. (B3), we have

$$C(x - 2\delta) \leq \tilde{C}(x - \delta) + \epsilon, \quad C(x + 2\delta) \geq \tilde{C}(x + \delta) - \epsilon.$$

This proves Eq. (9).

APPENDIX C: OBTAINING THE GROUND-STATE ENERGY BY SOLVING THE QEEP

Here, we discuss how to obtain the ground-state energy using algorithm in Ref. [58] to solve the QEEP. The cost of solving the QEEP as analyzed in Ref. [58] scales as ϵ^{-6} . However, the cost can be much reduced for the problem of ground-state energy estimation. For simplicity, we assume that $\|H\| < \pi/3$ and that τ is chosen to be 1.

In order to find the interval of size 2ϵ containing the ground-state energy, we first divide the interval $[-\pi/3, \pi/3]$ into M bins of equal size smaller than 2ϵ . We then define the indicator function associated with an interval $[a, b]$ to be

$$1_{[a,b]}(x) = \begin{cases} 1, & x \in [a, b], \\ 0, & x \notin [a, b]. \end{cases}$$

In QEEP, the goal is to estimate $\text{Tr}[\rho 1_{[a,b]}(H)]$, where $[a, b]$ is one of the M bins, to within precision $\mathcal{O}(\epsilon)$. However, in our setting, if we know $p_0 \geq \eta$, one can estimate $\text{Tr}[\rho 1_{[a,b]}(H)]$ to within error $\mathcal{O}(\eta)$. If we obtain $\text{Tr}[\rho 1_{[a,b]}(H)] < \eta$ with high confidence, then we know that the ground-state energy λ_0 is not in this interval. If we know $\text{Tr}[\rho 1_{[a,b]}(H)] > \eta/2$ with high confidence, then there is an eigenvalue in $[a, b]$. If the above task can be done, then we choose the leftmost bin in which $\text{Tr}[\rho 1_{[a,b]}(H)] > \eta/2$. This will enable us to solve the ground-state energy-estimation problem.

To estimate $\text{Tr}[\rho 1_{[a,b]}(H)]$, in Ref. [58] the indicator function $1_{[a,b]}(x)$ has first been approximated using a truncated Fourier series [58, Appendix A], similar to what we do in Sec. A. The number of terms N_{term} and the maximal evolution time T both scale like ϵ^{-1} . In Ref. [58], the author has proposed estimating each Fourier mode $\text{Tr}[\rho e^{-ijH}]$ to within error $\mathcal{O}(\epsilon/N_{\text{term}})$. Because here the estimation precision is $\mathcal{O}(\eta)$ rather than $\mathcal{O}(\epsilon)$, we should instead estimate $\text{Tr}[\rho e^{-ijH}]$ to within error $\mathcal{O}(\eta/N_{\text{term}}) = \mathcal{O}(\eta\epsilon)$. Because we are using Monte Carlo sampling, this requires $\mathcal{O}(\eta^{-2}\epsilon^{-2})$ samples. We need the same number of samples for each $\text{Tr}[\rho e^{-ijH}]$ and therefore the total time we need to run time evolution is $\mathcal{O}(N_{\text{term}} T \eta^{-2} \epsilon^{-2}) = \mathcal{O}(\eta^{-2} \epsilon^{-4})$. We omit polylogarithmic factors in the complexity.

However, if the analysis is done more carefully, the dependence on ϵ can be improved. First, one should note that the error for each $\text{Tr}[\rho e^{-ijH}]$ is independent and the estimate is unbiased (if we do not consider the Fourier approximation error), as is the case in our algorithm (Sec. III). Therefore, the total error for estimating $\text{Tr}[\rho 1_{[a,b]}(H)]$ accumulates sublinearly. More precisely, let the error for estimating $\text{Tr}[\rho e^{-ijH}]$ be ϵ_j , with variance σ_j^2 , and let the coefficient for $\text{Tr}[\rho e^{-ijH}]$ be A_j . Then, the total error $\sum_j A_j \epsilon_j$ has variance $\sum_j A_j^2 \sigma_j^2$. Therefore, the total error is roughly $\sqrt{\sum_j A_j^2 \sigma_j^2}$ instead of the linearly accumulated error $\sum_j A_j \sigma_j$. These two can have different asymptotic scaling depending on the magnitude of A_j . Because of this, one can in fact choose to estimate $\text{Tr}[\rho e^{-ijH}]$ to within error $\mathcal{O}(\eta/\sqrt{N_{\text{term}}}) = \mathcal{O}(\eta\epsilon^{-1/2})$. This saves an ϵ^{-1} factor in the total run time. Furthermore, one can choose to evaluate the approximate indicator function in a stochastic way, as we do in Sec. III. By taking into account the decay of Fourier coefficients, similar to Lemma 6 (iii), it is possible to further reduce the complexity.

APPENDIX D: COMPLEXITY ANALYSIS FOR USING TROTTER FORMULAS

In this appendix, instead of using the maximal evolution time and the total evolution time to quantify the complexity, we directly analyze the circuit depth and the total run time when the time evolution is simulated using Trotter formulas. We suppose that the Hamiltonian H can be decomposed as $H = \sum_\gamma H_\gamma$, where each of H_γ can be efficiently exponentiated. A p th-order Trotter formula applied to $e^{-i\tau H}$ with r Trotter steps gives us a unitary operator U_{HS} that approximates $e^{-i\tau H}$ with error

$$\|U_{\text{HS}} - e^{-i\tau H}\| \leq C_{\text{Trotter}} \tau^{p+1} r^{-p},$$

where C_{Trotter} is a prefactor, for which the simplest bound is $C_{\text{Trotter}} = \mathcal{O}((\sum_\gamma \|H_\gamma\|)^{p+1})$. Tighter bounds in the form of a sum of commutators are proved in Refs. [65,73].

1. The algorithm in this work

Our algorithm requires approximating Eq. (10) to precision η (as in Theorem 3, η is a lower bound of $p_0/2$) using Trotter formulas. Suppose that we are using a p th-order Trotter formula. Then, we want

$$\left\| \sum_j \hat{F}_j e^{ijx} \text{Tr}[\rho e^{-ij\tau H}] - \sum_j \hat{F}_j e^{ijx} \text{Tr}[\rho U_{\text{HS}}^j] \right\| = \mathcal{O}(\eta).$$

Since the left-hand side can be upper bounded by

$$\sum_j |\hat{F}_j| |j| \|e^{-i\tau H} - U_{\text{HS}}\| = \mathcal{O}(d \|e^{-i\tau H} - U_{\text{HS}}\|)$$

by Lemma 6 (iii), we only need to choose r so that

$$C_{\text{Trotter}} \tau^{p+1} r^{-p} = \mathcal{O}(\eta d^{-1}).$$

Therefore, we can choose

$$r = \max\{1, \tilde{\mathcal{O}}(d^{1/p} \eta^{-1/p} C_{\text{Trotter}}^{1/p} \tau^{1+1/p})\}.$$

The maximal evolution time in Corollary 3 tells us how many times we need to use the operator U_{HS} (multiplied by a factor τ). Multiplying this by r , we have the maximal circuit depth that we need, which is

$$dr = \tilde{\mathcal{O}}(\max\{\tau^{-1} \epsilon^{-1}, \epsilon^{-1-1/p} \eta^{-1/p} C_{\text{Trotter}}^{1/p}\}). \quad (\text{D1})$$

Similarly, we have the total run time

$$\tilde{\mathcal{O}}(\max\{\tau^{-1} \epsilon^{-1} \eta^{-2}, \epsilon^{-1-1/p} \eta^{-2-1/p} C_{\text{Trotter}}^{1/p}\}). \quad (\text{D2})$$

If we fix H and let $\epsilon, \eta \rightarrow 0$, then we can see that this gives us an extra $\epsilon^{-1/p} \eta^{-1/p}$ factor in the circuit depth and total run time, compared to the maximal evolution time and the total evolution time, respectively.

2. Quantum phase estimation

We then analyze the circuit depth and total run-time requirement for estimating the ground-state energy with QPE, where the time evolution is performed using Trotter formulas. We analyze the multi-ancilla qubit version of QPE and the result is equally valid for the single-ancilla qubit version using the semiclassical Fourier transform.

In QPE, when we replace all exact time evolution with U_{HS} , we would like to ensure that the probability of obtaining an energy measurement close to the ground-state energy remains bounded away from 0 by $\Omega(\eta)$. Therefore, the probability distribution of the final measurement outcome should be at most $\mathcal{O}(\eta)$ away from the original distribution in terms of the total variation distance.

Because the only part of QPE that depends on the time-evolution operator is the multiply controlled unitary

$$\sum_{j=0}^{J-1} |j\rangle\langle j| \otimes e^{-ij\tau H},$$

which is replaced by

$$\sum_{j=0}^{J-1} |j\rangle\langle j| \otimes U_{\text{HS}}^j$$

when we use Trotter formulas, we only need to ensure that the difference between the two operators is upper bounded by $\mathcal{O}(\eta)$ in terms of the operator norm. Therefore, we need

$$J \|e^{-ij\tau H} - U_{\text{HS}}\| = \mathcal{O}(\eta).$$

As discussed in Sec. I A, we need to choose $J = \mathcal{O}(\tau^{-1} \epsilon^{-1} \eta^{-1})$ (we need the τ^{-1} factor to account for

rescaling H and p_0 in Sec. I A is replaced by η). Following the same analysis as in the previous section, we need to choose the number of Trotter steps for approximating e^{-itH} to be

$$r = \max\{1, \mathcal{O}(J^{1/p} \eta^{-1/p} C_{\text{Trotter}}^{1/p} \tau^{1+1/p})\}.$$

Therefore, the circuit depth needed is

$$Jr = \mathcal{O}(\max\{\tau^{-1} \epsilon^{-1} \eta^{-1}, \epsilon^{-1-1/p} \eta^{-1-2/p} C_{\text{Trotter}}^{1/p}\}), \quad (\text{D3})$$

and the total run time is

$$\mathcal{O}(\max\{\tau^{-1} \epsilon^{-1} \eta^{-2}, \epsilon^{-1-1/p} \eta^{-2-2/p} C_{\text{Trotter}}^{1/p}\}). \quad (\text{D4})$$

Again, if we fix H and let $\epsilon, \eta \rightarrow 0$, then we can see that this gives us an extra $\epsilon^{-1/p} \eta^{-2/p}$ factor in the circuit depth and total run time, compared to the maximal evolution time and the total evolution time, respectively. This is worse by a factor of $\eta^{-1/p}$ than the cost using our algorithm.

APPENDIX E: THE CONTROL-FREE SETTING

In this appendix we introduce, as an alternative to the quantum circuit in Eq. (1), a circuit that does not require controlled time evolution. This construction is mainly based on the ideas in Refs. [43–45]. We introduce the construction of the circuit and discuss how to use the measurement results from the circuit to construct a random variable \tilde{Z} satisfying

$$\mathbb{E}[\tilde{Z}] = \text{Tr}[\rho e^{-itH}], \quad (\text{E1})$$

for any given t . Then, choosing $t = j\tau$, we are able to replace X_j and Y_j with $\text{Re} \tilde{Z}$ and $\text{Im} \tilde{Z}$, respectively, while satisfying Eqs. (2) and (3). In order to remove the need to perform controlled time evolution of H , we need some additional assumptions:

- (1) The initial state ρ is a pure state $|\phi_0\rangle$, prepared using a unitary circuit U_I .
- (2) We have a reference eigenstate $|\psi_R\rangle$ of H corresponding to a known eigenvalue λ_R . This eigenstate can be efficiently prepared using a unitary circuit U_R .
- (3) $\langle \psi_R | \phi_0 \rangle = 0$.

The last assumption $\langle \psi_R | \phi_0 \rangle = 0$ implies that $\langle \psi_R | e^{-itH} | \phi_0 \rangle = 0$ for all $t \in \mathbb{R}$, because $|\psi_R\rangle$ is an eigenvector of e^{-itH} . All of these are reasonable assumptions for a second-quantized fermionic Hamiltonian: we choose $|\psi_R\rangle$ to be the vacuum state, $\lambda_R = 0$, and $|\phi_0\rangle$ to be the Hartree-Fock state, which can be prepared efficiently [74]. Naturally, $\langle \psi_R | \phi_0 \rangle = 0$ because of particle-number conservation.

With these assumptions, we let

$$\alpha = \langle \phi_0 | e^{-it(H-\lambda_R)} | \phi_0 \rangle.$$

We also define

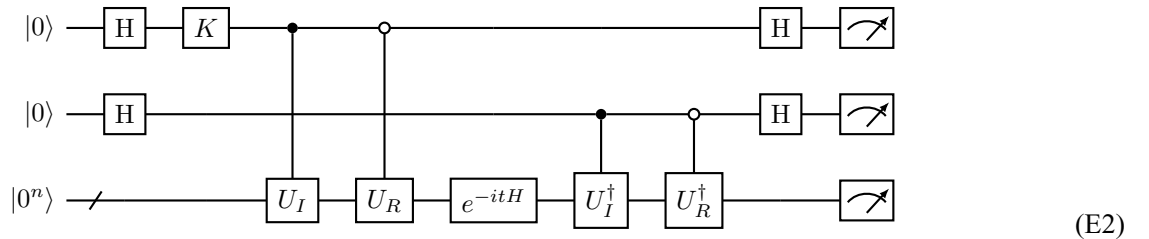
$$|\Psi_{0,\pm}\rangle = \frac{1}{\sqrt{2}}(|\psi_R\rangle \pm |\phi_0\rangle), \quad |\Psi_{1,\pm}\rangle = \frac{1}{\sqrt{2}}(|\psi_R\rangle \pm i|\phi_0\rangle).$$

With these states, we can express α in terms of expectation values:

$$\langle \Psi_{0,+} | e^{-itH} | \Psi_{0,\pm} \rangle = \frac{1}{2} e^{-i\lambda_R t} (1 \pm \alpha),$$

$$\langle \Psi_{0,+} | e^{-itH} | \Psi_{1,\pm} \rangle = \frac{1}{2} e^{-i\lambda_R t} (1 \pm i\alpha).$$

In Refs. [43,45], it is assumed that we have unitary circuits to prepare $|\Psi_{0,\pm}\rangle$ and $|\Psi_{1,\pm}\rangle$. However, it is not immediately clear how these circuits are constructed. Here, we take a slightly different approach. The circuit diagram is as follows:



(E2)

In this circuit, we choose $K = I$ for the real part of α or the phase gate S for the imaginary part of α . This circuit uses three registers, with the first two containing one qubit each and the third one containing n qubits.

We first analyze the probability of different measurement outcomes when $K = I$. When we run the above circuit and measure all the qubits, the probability of the measurement outcomes of the first two qubits being (b_1, b_2) and the rest of the qubits being all 0 is

$$\begin{aligned} p_{0,(b_1,b_2)} &= \begin{cases} |\langle \Psi_{0,+} | e^{-iH} | \Psi_{0,+} \rangle|^2 / 4, & b_1 = b_2 \\ |\langle \Psi_{0,+} | e^{-iH} | \Psi_{0,-} \rangle|^2 / 4, & b_1 \neq b_2 \end{cases} \\ &= \frac{1}{16} (1 + |\alpha|^2 + 2(-1)^{b_1+b_2} \operatorname{Re} \alpha). \end{aligned}$$

Here, we use the fact that $|\langle \Psi_{0,+} | e^{-iH} | \Psi_{0,+} \rangle| = |\langle \Psi_{0,-} | e^{-iH} | \Psi_{0,-} \rangle|$.

Similarly, when $K = S$, the probability of the measurement outcomes of the first two qubits being (b_1, b_2) and the rest of the qubits being all 0 is

$$\begin{aligned} p_{1,(b_1,b_2)} &= \begin{cases} |\langle \Psi_{0,+} | e^{-iH} | \Psi_{1,+} \rangle|^2 / 4, & b_1 = b_2 \\ |\langle \Psi_{0,+} | e^{-iH} | \Psi_{1,-} \rangle|^2 / 4, & b_1 \neq b_2 \end{cases} \\ &= \frac{1}{16} (1 + |\alpha|^2 - 2(-1)^{b_1+b_2} \operatorname{Im} \alpha). \end{aligned}$$

Based on the above analysis, we construct the random variable \tilde{Z} in the following way: we first run the circuit with $K = I$ and denote the measurement outcomes of the first two qubits by (b_1, b_2) . If the third register returns all 0 when measured, then we let $\tilde{X} = (-1)^{b_1+b_2}$. Otherwise, we let $\tilde{X} = 0$. Similarly, we define a random variable \tilde{Y} for $K = S$. We have

$$\mathbb{E}[\tilde{X}] = p_{0,(0,0)} + p_{0,(1,1)} - p_{0,(0,1)} - p_{0,(1,0)} = \frac{1}{2} \operatorname{Re} \alpha,$$

and

$$\mathbb{E}[\tilde{Y}] = p_{1,(0,0)} + p_{1,(1,1)} - p_{1,(0,1)} - p_{1,(1,0)} = -\frac{1}{2} \operatorname{Im} \alpha.$$

Therefore, we can define

$$\tilde{Z} = 2e^{-i\lambda_R t} (\tilde{X} - i\tilde{Y}).$$

Then,

$$\mathbb{E}[\tilde{Z}] = e^{-i\lambda_R t} \alpha = \operatorname{Tr}[\rho e^{-iH}].$$

Thus we can see that this new random variable \tilde{Z} satisfies Eq. (E1). Compared to the Z in the main text, this new random variable has a slightly larger variance:

$$\operatorname{var}[\tilde{Z}] \leq \mathbb{E}[|\tilde{Z}|^2] \leq 8.$$

This, however, does not change the asymptotic complexity.

APPENDIX F: DETAILS ON THE NUMERICAL EXPERIMENTS

In Fig. 3, we apply the procedure described in Sec. III to approximate the CDF of the Fermi-Hubbard model, described by the Hamiltonian

$$H = -t \sum_{\langle j,j' \rangle, \sigma} c_{j,\sigma}^\dagger c_{j',\sigma} + U \sum_j \left(n_{j,\uparrow} - \frac{1}{2} \right) \left(n_{j,\downarrow} - \frac{1}{2} \right), \quad (\text{F1})$$

where $c_{j,\sigma}$ ($c_{j,\sigma}^\dagger$) denotes the fermionic annihilation (creation) operator on the site j with spin $\sigma \in \{\uparrow, \downarrow\}$. $\langle \cdot, \cdot \rangle$ denotes sites that are adjacent to each other. $n_{j,\sigma} = c_{j,\sigma}^\dagger c_{j,\sigma}$ is the number operator. The sites are arranged into a one-dimensional chain, with the open boundary condition.

We first evaluate $\bar{G}(x)$ defined in Eq. (16) and the result is shown in Fig. 3. We use a classical computer to simulate the sampling from the quantum circuit. The initial state ρ is chosen to be the Hartree-Fock solution, which has an overlap of around 0.4 with the exact ground state. We can see that $\bar{G}(x)$ closely follows the CDF and even though there is significant noise from Monte Carlo sampling, the jump corresponding to the ground-state energy is clearly resolved.

Then, we consider estimating the ground-state energy from $\bar{G}(x)$. In this numerical experiment, we use a heuristic approach and the rigorous approach that comes with a provable error bound and confidence level is discussed in Secs. IV and V. We obtain the estimate by

$$x^* = \inf\{x : \bar{G}(x) \geq \eta/2\},$$

and x^*/τ is an estimate for the ground-state energy λ_0 . We expect $x^* \in [\tau\lambda_0 - \delta, \tau\lambda_0 + \delta]$. Here, η is chosen so that $p_0 \geq \eta$.

The error of the estimated ground-state energy, the total evolution time, and the maximal evolution time are shown in Fig. 4, in which we choose $U/t = 4$ for the Hubbard model. In the right-hand panel of Fig. 4, we can see that the line for the total evolution time runs parallel to the line for the maximal evolution time. Because the maximal evolution time scales linearly with respect to δ^{-1} and this plot uses logarithmic scales for both axes, we can see that the total evolution time has a δ^{-1} scaling and is therefore inversely proportional to the allowed error of the ground-state energy estimation.

APPENDIX G: FREQUENTLY USED SYMBOLS

In this appendix we list the symbols that are frequently used in this work.

TABLE III. Frequently used symbols in this work.

| Symbol | Meaning |
|----------------|--|
| H | The Hamiltonian for which we want to estimate the ground-state energy. |
| ρ | The initial state from which we perform time evolution and measurement. |
| p_k | The overlap between ρ and the k th-lowest eigensubspace. |
| τ | A renormalization factor satisfying $\tau \ H\ \leq \pi/4$. |
| $p(x)$ | The spectral density associated with τH and ρ . |
| $C(x)$ | The cumulative distribution function defined in Eq. (6). |
| $\tilde{C}(x)$ | The approximate CDF defined in Eq. (8). |
| $G(x)$ | An unbiased estimate of the ACDF $\tilde{C}(x)$ defined in Eq. (14). |
| $\bar{G}(x)$ | The average of multiple samples of $G(x)$, defined in Eq. (16). |
| J_k | An integer drawn from the distribution Eq. (11), signifying the number of steps in the time evolution. $ J_k \leq d$. |
| Z_k | A sample generated on a quantum circuit from two measurement outcomes. Defined in Eq. (13). Can only take value $\pm 1 \pm i$. |
| d | The maximal possible value of $ J_k $. |
| δ | In the context of Corollary 3, we choose $\delta = \tau\epsilon$ where ϵ is the allowed error of the ground-state energy. |
| ϑ | The allowed failure probability. |

[1] D. Aharonov, D. Gottesman, S. Irani, and J. Kempe, The power of quantum systems on a line, *Commun. Math. Phys.* **287**, 41 (2009).

[2] J. Kempe, A. Kitaev, and O. Regev, The complexity of the local Hamiltonian problem, *SIAM J. Comput.* **35**, 1070 (2006).

[3] A. Y. Kitaev, A. Shen, and M. N. Vyalyi, *Classical and Quantum Computation, Graduate Studies in Mathematics (American Mathematical Soc., 2002)*, Vol. 47.

[4] R. Oliveira and B. M. Terhal, The complexity of quantum spin systems on a two-dimensional square lattice, Preprint [ArXiv:quant-ph/0504050](https://arxiv.org/abs/quant-ph/0504050) (2005).

[5] N. M. Tubman, C. Mejuto-Zaera, J. M. Epstein, D. Hait, D. S. Levine, W. Huggins, Z. Jiang, J. R. McClean, R. Babbush, M. Head-Gordon, *et al.*, Postponing the orthogonality catastrophe: Efficient state preparation for electronic structure simulations on quantum devices, Preprint [ArXiv:1809.05523](https://arxiv.org/abs/1809.05523) (2018).

[6] R. Babbush, J. McClean, D. Wecker, A. Aspuru-Guzik, and N. Wiebe, Chemical basis of Trotter-Suzuki errors in

quantum chemistry simulation, *Phys. Rev. A* **91**, 022311 (2015).

[7] K. Sugisaki, S. Nakazawa, K. Toyota, K. Sato, D. Shiomi, and T. Takui, Quantum chemistry on quantum computers: A method for preparation of multiconfigurational wave functions on quantum computers without performing post-Hartree-Fock calculations, *ACS Cent. Sci.* **5**, 167 (2018).

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

[9] Y. Atia and D. Aharonov, Fast-forwarding of Hamiltonians and exponentially precise measurements, *Nat. Commun.* **8**, 1572 (2017).

[10] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum Metrology, *Phys. Rev. Lett.* **96**, 010401 (2006).

[11] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, *Nat. Photonics* **5**, 222 (2011).

[12] M. Zwiernz, C. A. Pérez-Delgado, and P. Kok, General Optimality of the Heisenberg Limit for Quantum Metrology, *Phys. Rev. Lett.* **105**, 180402 (2010).

[13] M. Zwiernz, C. A. Pérez-Delgado, and P. Kok, Ultimate limits to quantum metrology and the meaning of the Heisenberg limit, *Phys. Rev. A* **85**, 042112 (2012).

[14] Y. Aharonov and D. Bohm, Time in the quantum theory and the uncertainty relation for time and energy, *Phys. Rev.* **122**, 1649 (1961).

[15] Y. Aharonov, S. Massar, and S. Popescu, Measuring energy, estimating Hamiltonians, and the time-energy uncertainty relation, *Phys. Rev. A* **66**, 052107 (2002).

[16] A. M. Childs, J. Preskill, and J. Renes, Quantum information and precision measurement, *J. Mod. Opt.* **47**, 155 (2000).

[17] A. Y. Kitaev, Quantum measurements and the Abelian stabilizer problem, Preprint [ArXiv:quant-ph/9511026](https://arxiv.org/abs/quant-ph/9511026) (1995).

[18] W. J. Huggins, J. Lee, U. Baek, B. O’Gorman, and K. B. Whaley, A non-orthogonal variational quantum eigensolver, *New J. Phys.* **22**, 073009 (2020).

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

[20] P. J. O’Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, *et al.*, Scalable Quantum Simulation of Molecular Energies, *Phys. Rev. X* **6**, 031007 (2016).

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

[22] D. S. Abrams and S. Lloyd, Quantum Algorithm Providing Exponential Speed Increase for Finding Eigenvalues and Eigenvectors, *Phys. Rev. Lett.* **83**, 5162 (1999).

[23] Y. Ge, J. Tura, and J. I. Cirac, Faster ground state preparation and high-precision ground energy estimation with fewer qubits, *J. Math. Phys.* **60**, 022202 (2019).

[24] L. Lin and Y. Tong, Near-optimal ground state preparation, *Quantum* **4**, 372 (2020).

[25] D. Poulin and P. Wocjan, Preparing Ground States of Quantum Many-Body Systems on a Quantum Computer, *Phys. Rev. Lett.* **102**, 130503 (2009).

- [26] R. Babbush, J. R. McClean, M. Newman, C. Gidney, S. Boixo, and H. Neven, Focus beyond quadratic speedups for error-corrected quantum advantage, *PRX Quantum* **2**, 010103 (2021).
- [27] K. E. Booth, B. O’Gorman, J. Marshall, S. Hadfield, and E. Rieffel, Quantum-accelerated constraint programming, Preprint [ArXiv:2103.04502](#) (2021).
- [28] E. T. Campbell, Early fault-tolerant simulations of the Hubbard model, *Quantum Sci. Technol.* **7**, 015007 (2021).
- [29] D. Layden, First-order Trotter error from a second-order perspective, Preprint [ArXiv:2107.08032](#) (2021).
- [30] I. D. Kivlichan, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, W. Sun, Z. Jiang, N. Rubin, A. Fowler, A. Aspuru-Guzik, *et al.*, Improved fault-tolerant quantum simulation of condensed-phase correlated electrons via trotterization, *Quantum* **4**, 296 (2020).
- [31] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Quantum algorithms revisited, *Proc. R. Soc. London, Ser. A* **454**, 339 (1998).
- [32] M. A. Nielsen and I. Chuang, Quantum computation and quantum information, *Am. J. Phys.* **70**, 558 (2002).
- [33] E. Knill, G. Ortiz, and R. D. Somma, Optimal quantum measurements of expectation values of observables, *Phys. Rev. A* **75**, 012328 (2007).
- [34] D. Nagaj, P. Wocjan, and Y. Zhang, Fast amplification of QMA, *Quantum Inf. Comput.* **9**, 1053 (2009).
- [35] D. Poulin and P. Wocjan, Sampling from the Thermal Quantum Gibbs State and Evaluating Partition Functions with a Quantum Computer, *Phys. Rev. Lett.* **103**, 220502 (2009).
- [36] R. Babbush, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, A. Paler, A. Fowler, and H. Neven, Encoding Electronic Spectra in Quantum Circuits with Linear T Complexity, *Phys. Rev. X* **8**, 041015 (2018).
- [37] Y. R. Sanders, D. W. Berry, P. C. Costa, L. W. Tessler, N. Wiebe, C. Gidney, H. Neven, and R. Babbush, Compilation of fault-tolerant quantum heuristics for combinatorial optimization, *PRX Quantum* **1**, 020312 (2020).
- [38] D. Wang, O. Higgott, and S. Brierley, Accelerated Variational Quantum Eigensolver, *Phys. Rev. Lett.* **122**, 140504 (2019).
- [39] N. Wiebe, C. Granade, A. Kapoor, and K. M. Svore, Bayesian inference via rejection filtering, Preprint [ArXiv:1511.06458](#) (2015).
- [40] R. B. Griffiths and C.-S. Niu, Semiclassical Fourier Transform for Quantum Computation, *Phys. Rev. Lett.* **76**, 3228 (1996).
- [41] D. W. Berry, B. L. Higgins, S. D. Bartlett, M. W. Mitchell, G. J. Pryde, and H. M. Wiseman, How to perform the most accurate possible phase measurements, *Phys. Rev. A* **80**, 052114 (2009).
- [42] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Entanglement-free Heisenberg-limited phase estimation, *Nature* **450**, 393 (2007).
- [43] S. Lu, M. C. Bañuls, and J. I. Cirac, Algorithms for quantum simulation at finite energies, Preprint [ArXiv:2006.03032](#) (2020).
- [44] T. E. O’Brien, S. Polla, N. C. Rubin, W. J. Huggins, S. McArdle, S. Boixo, J. R. McClean, and R. Babbush, Error mitigation via verified phase estimation, Preprint [ArXiv:2010.02538](#) (2020).
- [45] A. Russo, K. Rudinger, B. Morrison, and A. Baczewski, Evaluating energy differences on a quantum computer with robust phase estimation, Preprint [ArXiv:2007.08697](#) (2020).
- [46] T. E. O’Brien, B. Tarasinski, and B. M. Terhal, Quantum phase estimation of multiple eigenvalues for small-scale (noisy) experiments, *New J. Phys.* **21**, 023022 (2019).
- [47] R. Somma, G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Simulating physical phenomena by quantum networks, *Phys. Rev. A* **65**, 042323 (2002).
- [48] In this paper, we use the following asymptotic notations besides the usual \mathcal{O} notation: we write $f = \Omega(g)$ if $g = \mathcal{O}(f)$; $f = \Theta(g)$ if $f = \mathcal{O}(g)$ and $g = \mathcal{O}(f)$; $f = \tilde{\mathcal{O}}(g)$ if $f = \mathcal{O}(g \text{ polylog}(g))$.
- [49] S. Chakraborty, A. Gilyén, and S. Jeffery, The power of block-encoded matrix powers: Improved regression techniques via faster hamiltonian simulation, Preprint [ArXiv:1804.01973](#) (2018).
- [50] G. H. Low and I. L. Chuang, Hamiltonian simulation by qubitization, *Quantum* **3**, 163 (2019).
- [51] A. Gilyén, Y. Su, G. H. Low, and N. Wiebe, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* (ACM, Phoenix, AZ, USA, 2019), p. 193.
- [52] G. H. Low and I. L. Chuang, Optimal Hamiltonian Simulation by Quantum Signal Processing, *Phys. Rev. Lett.* **118**, 010501 (2017).
- [53] C. Bittel and M. Kliesch, Training variational quantum algorithms is NP-hard—even for logarithmically many qubits and free fermionic systems, Preprint [ArXiv:2101.07267](#) (2021).
- [54] M. Motta, C. Sun, A. T. K. Tan, M. J. O’Rourke, E. Ye, A. J. Minnich, F. G. S. L. Brandão, and G. K.-L. Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, *Nat. Phys.* **16**, 205 (2019).
- [55] R. M. Parrish and P. L. McMahon, Quantum filter diagonalization: Quantum eigendecomposition without full quantum phase estimation, Preprint [ArXiv:1909.08925](#) (2019).
- [56] N. H. Stair, R. Huang, and F. A. Evangelista, A multireference quantum Krylov algorithm for strongly correlated electrons, *J. Chem. Theory Comput.* **16**, 2236 (2020).
- [57] J. Emerson, S. Lloyd, D. Poulin, and D. Cory, Estimation of the local density of states on a quantum computer, *Phys. Rev. A* **69**, 050305 (2004).
- [58] R. D. Somma, Quantum eigenvalue estimation via time series analysis, *New J. Phys.* **21**, 123025 (2019).
- [59] D. W. Berry, A. M. Childs, and R. Kothari, in *2015 IEEE 56th Annual Symposium on Foundations of Computer Science* (IEEE, 2015), p. 792.
- [60] M. Suzuki, General theory of fractal path integrals with applications to many-body theories and statistical physics, *J. Math. Phys.* **32**, 400 (1991).
- [61] D. W. Berry, A. M. Childs, Y. Su, X. Wang, and N. Wiebe, Time-dependent Hamiltonian simulation with l1-norm scaling, *Quantum* **4**, 254 (2020).
- [62] E. Campbell, Random Compiler for Fast Hamiltonian Simulation, *Phys. Rev. Lett.* **123**, 070503 (2019).

- [63] C.-F. Chen, H.-Y. Huang, R. Kueng, and J. A. Tropp, Quantum simulation via randomized product formulas: Low gate complexity with accuracy guarantees, Preprint [ArXiv:2008.11751](https://arxiv.org/abs/2008.11751) (2020).
- [64] A. M. Childs and Y. Su, Nearly Optimal Lattice Simulation by Product Formulas, *Phys. Rev. Lett.* **123**, 050503 (2019).
- [65] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu, Theory of Trotter Error with Commutator Scaling, *Phys. Rev. X* **11**, 011020 (2021).
- [66] M. C. Tran, S.-K. Chu, Y. Su, A. M. Childs, and A. V. Gorshkov, Destructive Error Interference in Product-Formula Lattice Simulation, *Phys. Rev. Lett.* **124**, 220502 (2020).
- [67] C. Yi and E. Crosson, Spectral analysis of product formulas for quantum simulation, Preprint [ArXiv:2102.12655](https://arxiv.org/abs/2102.12655) (2021).
- [68] S. Boixo and R. D. Somma, Parameter estimation with mixed-state quantum computation, *Phys. Rev. A* **77**, 052320 (2008).
- [69] C.-H. Rhee and P. W. Glynn, in *Proceedings of the 2012 Winter Simulation Conference (WSC)* (IEEE, Berlin, Germany, 2012), p. 1.
- [70] C.-H. Rhee and P. W. Glynn, Unbiased estimation with square root convergence for SDE models, *Oper. Res.* **63**, 1026 (2015).
- [71] M. B. Giles, Multilevel Monte Carlo methods, *Acta Numer.* **24**, 259 (2015).
- [72] L. Lin and Y. Tong, Optimal polynomial based quantum eigenstate filtering with application to solving quantum linear systems, *Quantum* **4**, 361 (2020).
- [73] Y. Su, H.-Y. Huang, and E. T. Campbell, Nearly tight Trotterization of interacting electrons, Preprint [ArXiv:2012.09194](https://arxiv.org/abs/2012.09194) (2020).
- [74] I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K.-L. Chan, and R. Babbush, Quantum Simulation of Electronic Structure with Linear Depth and Connectivity, *Phys. Rev. Lett.* **120**, 110501 (2018).