Computing Ground State Properties with Early Fault-Tolerant Quantum Computers

Ruizhe Zhang * Guoming Wang† Peter Johnson‡

September 28, 2021

Abstract

Significant effort in applied quantum computing has been devoted to the problem of ground state energy estimation for molecules and materials. Yet, for many applications of practical value, additional properties of the ground state must be estimated. These include Green's functions used to compute electron transport in materials and the one-particle reduced density matrices used to compute electric dipoles of molecules. In this paper, we propose a quantum-classical hybrid algorithm to efficiently estimate such ground state properties with high accuracy using low-depth quantum circuits. We provide an analysis of various costs (circuit repetitions, maximal evolution time, and expected total runtime) as a function of target accuracy, spectral gap, and initial ground state overlap. This algorithm suggests a concrete approach to using early fault tolerant quantum computers for carrying out industry-relevant molecular and materials calculations.

*ruizhe@utexas.edu. The University of Texas at Austin. This work was done while R.Z. was a research intern at Zapata Computing Inc.

†guoming.wang@zapatacomputing.com. Zapata Computing Inc.

‡peter@zapatacomputing.com. Zapata Computing Inc.

1 Introduction

One of the primary applications of quantum computing is the simulation of materials and molecules, which are inherently quantum mechanical. It is hoped that future powerful quantum computers will be used in the development of materials and drug discovery [CRAG18]. Although they have yet to realize commercial application, quantum computers have been improving at a rapid rate, increasing the demand for quantum algorithms with high-impact use cases. To date, the main focus of quantum algorithm development for quantum chemistry and materials has been on ground state energy estimation [CRO+19]. This problem is mathematically formulated as estimating the lowest eigenvalue of the Hamiltonian matrix that characterizes the physical system. One of the first quantum chemistry applications of quantum computers was the use quantum phase estimation for estimating the ground state energy of small molecules [AGDLHG05]. More recently, the variational quantum eigensolver algorithm [PMS+14a] was developed to use near-term intermediate-scale quantum (NISQ) computers to solve the ground state energy estimation problem.

However, in characterizing materials or analyzing small molecules for drug discovery, one often needs to estimate properties of the ground state beyond just the energy. These include transport properties [MW92], electric dipole moments [Jen17], and molecular forces [OSS+19]. Such properties depend on expectation values of observables O with respect to the ground state of a Hamiltonian H. The problem of estimating such quantities was studied in [Amb14, GY19, GPY20], showing that it is even harder, in a complexity theoretic sense, than the ground state energy estimation problem in general. The algorithms considered in these works were developed for idealistic quantum computers. In particular, the quantum circuit depths involved in these methods are too deep to even be implemented on early fault-tolerant quantum computers. An approach to estimating ground state properties that is more amenable to near-term quantum computers is to use the variational quantum eigensolver algorithm [MMS+19, OSS+19]. This gives a means to prepare an approximation to the ground state, from which properties can be estimated. However, recent work has suggested that VQE alone is not practical for solving problems of industrial relevance [GRB+20]; methods which invoke some degree of quantum amplification (e.g. [WKJC21]) seem necessary in order for quantum computers to compete with state-of-the-art methods in quantum chemistry and materials. Further issues with the variational quantum eigensolver and its variants are that there are no guarantees on the quality of the output ground state and that heuristic optimization methods struggle to prepare high-fidelity ground states.

This motivates the development of quantum algorithms for ground state property estimation which are both reliable *and* able to be run on near-term quantum computers (e.g. early fault-tolerant quantum devices). The central question that this paper addresses is then:

Is it possible to estimate ground state properties of a Hamiltonian reliably using early fault-tolerant quantum computers?

In this paper, we provide an affirmative answer to this question. Furthermore, we propose an algorithm for the ground state property estimation using low-depth quantum circuits. The main theorem is stated as follows:

Theorem 1.1 (Main theorem, informal)**.** *Given a Hamiltonian H and an observable O. Suppose we have access to a unitary U_I that prepares a state $|\phi_0\rangle$ that has non-trivial overlap with the ground state $|\psi_0\rangle$ of H. Then, there exists an algorithm to estimate $\langle\psi_0| O |\psi_0\rangle$ with high accuracy and low-depth: the maximal Hamiltonian evolution time is $\widetilde{O}(\gamma^{-1})$, where γ is the spectral gap of H.*

We note that the maximal evolution time, which is the maximal length of time we need to perform coherent time evolution, can roughly determine the depth of the quantum circuit. Our

result achieves a nearly-linear dependence on γ^{-1} and only poly-logarithmic on the accuracy ϵ^{-1}, which improves the $\widetilde{O}(\epsilon^{-1})$ maximal evolution time in the ground state energy estimation algorithms [LT21, Som19]. On the other hand, our result does not violate the Heisenberg limit because the total evolution time still depends on $\mathrm{poly}(\epsilon^{-1})$. Hence, our algorithm has a great advantage when the Hamiltonian's spectral gap is much larger than the estimation accuracy, making it easier to be implemented in the early fault-tolerant devices.

Organization. In Section 2 we formally state the problem of ground state property estimation. In Section 3 we review the method developed in [LT21] for estimating ground state energies. In the next three sections we explain our main algorithms and give an analysis for their performances starting from the simplest case and building to the most-involved, general case. Section 4 presents the case of a unitary observable which commutes with the Hamiltonian. Section 5 presents the case of a unitary observable which does not necessarily commute with the Hamiltonian. Section 6 describes the case of a general observable. Then, Section 7 gives two applications of the ground state property estimation algorithm. Section 8 gives a discussion of the results and presents some open questions.

2 Ground State Property Estimation Problem

In this section, we will formally define the ground state property estimation problem. This problem was initially studied by Ambainis [Amb14] as the approximate simulation problem (APX-SIM), and he proved that APX-SIM is $\mathsf{P}^{\mathsf{QMA}[\log]}$-complete.

Problem 2.1 (Approximate simulation (APX-SIM), [Amb14])**.** *Given a k-local Hamiltonian H, an ℓ-local observable O, and real numbers a, b, ϵ such that $b - a \geq 1/\mathrm{poly}(n)$, and $\epsilon \geq 1/\mathrm{poly}(n)$, for n the number of qubits the Hamiltonian H acts on, decide:*

- ***Yes case:*** *H has a ground state $|\psi_0\rangle$ such that $\langle\psi_0| O |\psi_0\rangle \leq a$,*
- ***No case:*** *for any state $|\psi\rangle$ with $\langle\psi| H |\psi\rangle \leq \lambda_0 + \epsilon$ where λ_0 is the ground state energy of H, it holds that $\langle\psi_0| O |\psi_0\rangle \geq b$.*

In the follow-up works, APX-SIM was shown to be $\mathsf{P}^{\mathsf{QMA}[\log]}$-complete even for 5-local Hamiltonian and 1-local observable [GY19], and also for some physics models like 2D Heisenberg model and 1D line model [GPY20, WBG20]. However, these previous studies only considered the worst-case hardness, and thus they defined Problem 2.1 as a decision problem. For the purpose of designing efficient algorithms, we first define the "search version" of APX-SIM as follows:

Problem 2.2 (Search version of APX-SIM)**.** *Given a Hamiltonian H, an observable O, and $\epsilon \in (0, 1)$, with $\Omega(1)$ probability, estimate $\langle\psi_0| O |\psi_0\rangle$ with an additive/multiplicative error at most ϵ.*

In general, Problem 2.2 will not be more tractable than Problem 2.1. Thus, we may need some prior information about the Hamiltonian H and its ground state. Motivated by the widely used variational quantum eigensolver (VQE) [PMS+14b, MRBAG16] and the Hartree-Fock method [SO12] in quantum chemistry, it is reasonable to assume that we get access to an initial state $|\phi_0\rangle$ that has a nontrivial overlap with the ground state of H. Moreover, we assume that the Hamiltonian H has a nontrivial spectral gap. We note that almost all Hamiltonians in practice satisfy this condition. With these assumptions, we formally define the ground state property estimation problem as follows:

Problem 2.3 (Ground state property estimation)**.** *Given a Hamiltonian H with spectral gap γ and ground state $|\psi_0\rangle$, an observable O, a unitary U_I such that it prepares an initial state $|\phi_0\rangle$ with $|\langle\phi_0|\psi_0\rangle|^2 \geq \eta$, and $\epsilon \in (0,1)$, estimate $\langle\psi_0| O |\psi_0\rangle$ with an additive/multiplicative error at most ϵ with $\Omega(1)$ probability.*

Remark 2.4. *We notice that when $O = H$, Problem 2.3 becomes the ground state energy estimation problem. Moreover, our assumptions on the initial state and the spectral gap are also used in many ground state energy estimation algorithms [GTC19, LT20a, LT21].*

3 An Overview of the Low-Depth Ground State Energy Estimation

In this section, we provide a brief overview of the low-depth ground state energy estimation algorithm proposed by Lin and Tong [LT21]. Our algorithms are inspired by this algorithm and use it as a subroutine.

More specifically, they showed that:

Theorem 3.1 ([LT21])**.** *Given a Hamiltonian H with eigenvalues in the interval $[-\pi/3, \pi/3]$ and its ground state $|\psi_0\rangle$ has energy λ_0. And suppose we can prepare an initial state $|\phi_0\rangle$ such that $p_0 \geq \eta$ for some known η, where $p_0 := |\langle\phi_0|\psi_0\rangle|^2$. Then, for any $\epsilon, \nu \in (0,1)$, there exists an algorithm that estimates λ_0 with an additive error ϵ with probability $1-\nu$, by running a parameterized quantum circuit with the maximum quantum evolution time $\widetilde{O}(\epsilon^{-1})$ and the expected total quantum evolution time $\widetilde{O}(\epsilon^{-1}\eta^{-2})$.*

The pseudo-code of their algorithm is given in Algorithm 1.

The main technique of their algorithm is a classical post-processing procedure that extracts information from the following Hadamard test circuit (Figure 1).

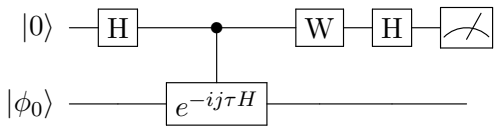


Figure 1: Quantum circuit parameterized by j. H is the Hadamard gate and W is either I or a phase gate.

Suppose the initial state $|\phi_0\rangle$ can be expanded as $|\phi_0\rangle = \sum_k \alpha_k |\psi_k\rangle$ in the eigen-basis of H and let $p_k := |\alpha_k|^2$ be the overlap with the k-th eigenstate. They considered the overlaps $p_0, p_1, \ldots$ as a density function:

$$p(x) := \sum_k p_k \delta(x - \lambda_k). \tag{1}$$

Then, the CDF $C(x)$ can be defined by the convolution of $p(x)$ and the 2π-periodic Heaviside function $H(x)$, which is 0 in $[(2k-1)\pi, 2k\pi)$ and 1 in $[2k\pi, (2k+1)\pi)$ for any $k \in \mathbb{Z}$. Thus, $C(x)$ is also a periodic function, which makes it convenient to apply the Fourier approximation. They showed that $H(x)$ can be approximated by a low-Fourier degree function $F(x)$ in the intervals

Algorithm 1 Ground State Energy Estimation

1: **procedure** EstimateGSE($\epsilon, \tau, \eta, \nu$)
2: ▷ Initialization
3: $d \leftarrow O(\delta^{-1}\log(\delta^{-1}\eta^{-1}))$, $\delta \leftarrow \tau\epsilon$
4: **for** $i \leftarrow -d, \ldots, d$ **do**
5: $\widehat{F}_i \leftarrow \widehat{F}_{d,\delta,i}$
6: Compute θ_i, the phase angle of $\widehat{F}_i$
7: **end for**
8: $\mathcal{F} \leftarrow \sum_{|i|\le d} |\widehat{F}_i|$
9: $N_b \leftarrow \Omega(\log(1/\nu) + \log\log(1/\delta))$, $N_s \leftarrow O(\eta^{-2}\log^2(d))$
10: ▷ Sampling from the quantum circuit
11: **for** $k \leftarrow 1, \ldots, N_b N_s$ **do**
12: Independently sample $J_k \sim [-d, d]$ with $\Pr[J_k = j] \propto |\widehat{F}_j|$
13: Measure (X_k, Y_k) by running the quantum circuit with (Figure. 1) parameter k
14: $Z_k \leftarrow X_k + iY_k$
15: **end for**
16: ▷ Classical post-processing
17: $x_L \leftarrow -\pi/3$, $X_R \leftarrow \pi/3$
18: **while** $x_R - x_L > 2\delta$ **do** ▷ Invert CDF
19: $x_M \leftarrow (x_L + x_R)/2$
20: **for** $r \leftarrow 1, \ldots, N_b$ **do**
21: $\overline{G}_r \leftarrow \frac{\mathcal{F}}{N_s} \sum_{k=(r-1)N_s+1}^{rN_s} Z_k e^{i(\theta_{J_k} + J_k x_M)}$ ▷ Multi-level Monte Carlo method
22: **end for**
23: **if** $|\{r : \overline{G}_r \ge (3/4)\eta\}| \le N_b/2$ **then**
24: $x_R \leftarrow x_M + (2/3)\delta$
25: **else**
26: $x_L \leftarrow x_M - (2/3)\delta$
27: **end if**
28: **end while**
29: **return** $(x_L + x_R)/2$
30: **end procedure**

$[-\pi + \delta, -\delta]$ and $[\delta, \pi - \delta]$. Then, they defined the approximated cumulative distribution function (ACDF) as $\widetilde{C}(x) := (F \star p)(x)$ and proved that

$$C(x - \delta) - \eta/8 \le \widetilde{C}(x) \le C(x + \delta) + \eta/8 \quad \forall x \in [-\pi/3, \pi/3]. \tag{2}$$

Moreover, for each x, we have

$$\widetilde{C}(x) = \sum_{|j|\le d} \widehat{F}_j e^{ijx} \cdot \langle\phi_0| e^{-ijH} |\phi_0\rangle, \tag{3}$$

where $\widehat{F}_j$ is the Fourier coefficient of $F(x)$. Note that $\langle\phi_0| e^{-ijH} |\phi_0\rangle$ can be estimated via the parameterized quantum circuit (Figure 1). Hence, we can estimate the ACDF at every point in $[-\pi/3, \pi/3]$. Moreover, they showed that the multi-level Monte Carlo method can be applied here to save the number of samples needed to achieve a high-accuracy estimation (Line 21).

Therefore, we can estimate the ground state energy λ_0 by locating the first non-zero point of the CDF $C(x)$, which is $\eta/8$-approximated by the ACDF $\widetilde{C}(x)$. Since we assume that $p_0 \ge \eta$, the

approximation error and the estimation error of $\widetilde{C}(x)$ can be tolerated, and we can find λ_0 via a robust binary search (Line 18).

We note that the maximal evolution time of this algorithm corresponds to the Fourier degree of $F(x)$, which is $\widetilde{O}(\epsilon^{-1})$ by the construction. More details of this algorithm and the proofs are given in Appendix A.

4 Algorithm for Commutative Case

In this section, we consider a easier case that O is unitary and commutes with the Hamiltonian H, and give a two-step quantum-classical hybrid algorithm for Problem 2.3. More specifically, suppose the initial state can be expanded in the eigenbasis as follows: $\phi_0 = \sum_k c_k |\psi_k\rangle$ with $p_k := |c_k|^2$. We note that $\{|\psi_k\rangle\}$ is also an eigenbasis of O since O and H commute. In Step 1, we run [LT21]'s algorithm to estimate the ground state energy λ_0 and the overlap between the initial state and the ground state p_0. In Step 2, we construct a similar CDF function for the density $\sum_k O_k p_k \delta(x - \lambda_k)$, where $O_k := \langle\phi_k| O |\phi_k\rangle$. If we evaluate the CDF at λ_0, we can obtain an estimate of O_0.

4.1 Step 1: estimate the initial overlap

We first run the procedure EstimateGSE (Algorithm 1) to estimate the ground state energy λ_0 with an additive error ϵ. Let $x^\star$ be the output. We remark that $x^\star$ satisfy $C(x^\star + \tau\epsilon) \geq p_0$ and $C(x^\star - \tau\epsilon) = 0$. However, we can only extract p_0 from the ACDF $\widetilde{C}(x)$, which satisfies:

$$C(x - \tau\epsilon) - \eta/8 \leq \widetilde{C}(x) \leq C(x + \tau\epsilon) + \eta/8 \quad \forall x \in [-\pi/3, \pi/3]. \tag{4}$$

If $[x - \tau\epsilon, x + \tau\epsilon]$ contains a "jump" of $C(x)$, i.e., an eigenvalue λ_k, then the approximation error of $\widetilde{C}(x)$ will be large.

Hence, we say a point x is "good" for λ_k if $[x - \tau\epsilon, x + \tau\epsilon]$ is contained in $[\tau\lambda_k, \tau\lambda_{k+1})$. It is easy to see that $\widetilde{C}(x)$ will be an $\eta/8$-additive approximation of $\sum_{j\leq k} p_k$ if x is good. Our goal is to find an x_{good} that is good for λ_0, and estimating $\widetilde{C}(x_{\mathsf{good}})$ gives the overlap p_0. The following claim gives a way to construct x_{good} using the spectral gap of H.

Claim 4.1 (Construct x_{good})**.** *Let γ be the spectral gap of the Hamiltonian H. For any $\epsilon \in (0, \gamma/4)$, $x^\star + \tau\gamma/2$ is good for λ_0, where $x^\star$ is the output of* EstimateGSE(ϵ, η) *(Algorithm 1).*

Proof. We know that $x^\star$ satisfies:

$$x^\star - \tau\epsilon < \tau\lambda_0 \leq x^\star + \tau\epsilon. \tag{5}$$

Then, we have

$$x^\star + \tau\gamma/2 > \tau\lambda_0 - \tau\epsilon + \tau\gamma/2 > \tau\lambda_0 + \tau\epsilon. \tag{6}$$

We also have

$$\begin{aligned} x^\star + \tau\gamma/2 &< \tau\lambda_0 + \tau\epsilon + \tau\gamma/2 && (7)\\ &\leq \tau(\lambda_1 - \gamma) + \tau\epsilon + \tau\gamma/2 \\ &= \tau\lambda_1 + \tau(\epsilon - \gamma/2) \\ &< \tau\lambda_1 - \tau\epsilon. && (8) \end{aligned}$$

Therefore, $x^\star$ is good for λ_0. □

We note that in [LT21], the ACDF's approximation error is chosen to be $\eta/8$. We may directly change it to $\epsilon\eta/8$ without significantly changing the circuit depth, since by Lemma A.8 the degree of F can only blowup by a log factor of ϵ.

Lemma 4.2 (Estimating the overlap)**.** *For any $\epsilon_0, \nu \in (0,1)$, the overlap $p_0 := |\langle\phi_0|\psi_0\rangle|^2$ can be estimated with multiplicative error $1 \pm O(\epsilon_0)$ with probability $1-\nu$ by runs the quantum circuit (Figure 1) $\widetilde{O}(\epsilon_0^{-2}\eta^{-2})$ times with expected total evolution time $\widetilde{O}(\gamma^{-1}\epsilon^{-2}\eta^{-2})$ and maximal evolution time $O(\gamma^{-1})$.*

Proof. By Claim 4.1, if we set the additive error of ground state energy λ_0 to be $O(\gamma)$, then we can construct an x_{good} that is good for λ_0. By Theorem 3.1, it can be done with maximum quantum evolution time $\widetilde{O}(\gamma^{-1})$ and the expected total quantum evolution time $\widetilde{O}(\gamma^{-1}\eta^{-2})$. Notice that we need to take $d = O(\delta^{-1}\log(\delta^{-1}\epsilon_0^{-1}\eta^{-1}))$ (Line 3 in Algorithm 1) to make $\widetilde{C}(x_{\mathsf{good}})$ be an $O(\epsilon_0\eta)$-approximation of p_0, where $\delta = \tau\gamma$.

Next, we estimate $\widetilde{C}(x_{\mathsf{good}})$ with additive error $\eta\epsilon$ with probability $1-\nu$. We have an unbiased estimator

$$\overline{G}(x; \mathbf{Z}, \mathbf{J}) = \mathcal{F}\mathbf{Z}e^{i\theta_{\mathbf{J}}+\mathbf{J}x} \tag{9}$$

with variance $O(\log^2(d))$ and expected evolution time $\widetilde{O}(\tau d/\log(d))$. If we naively take the sample mean as the estimator, then by Chebyshev's inequality, the sample complexity is $\widetilde{O}(\epsilon_0^{-2}\eta^{-2}\nu^{-2})$ to have an additive error $O(\epsilon_0\eta)$ with probability $1-\nu$.

Instead, we can use the so-called "median-of-means" trick to reduce the sample complexity. More specifically, let $N_g = O(\log(1/\nu))$ and $K = O(\epsilon_0^{-2})$. We first partition $m = N_g K$ samples $(Z_1, J_1), \ldots, (Z_m, J_m)$ into N_g groups of size K. Then, for any $i \in [N_g]$, the i-th group mean is

$$\overline{G}_i := \frac{1}{K}\sum_{j=1}^{K} \overline{G}(x; Z_{(i-1)K+j}, J_{(i-1)K+j}). \tag{10}$$

The final estimator is given by the median of these group means, i.e.,

$$\overline{G}(x) := \mathrm{median}(\overline{G}_1, \ldots, \overline{G}_{N_g}). \tag{11}$$

By Chernoff bound, it is easy to see that $\overline{G}(x)$ has an additive error at most $(\eta\epsilon_0)$ with probability $1-\nu$. It will imply that multiplicative error is at most $1 \pm O(\epsilon_0)$ since $p_0 = \Theta(\eta)$. And the sample complexity of $\overline{G}(x)$ is $\widetilde{O}(\epsilon_0^{-2}\eta^{-2})$. Hence, the expected total evolution time is $\widetilde{O}(\gamma^{-1}\epsilon_0^{-2}\eta^{-2})$. Since we run the same quantum circuit to estimate $\overline{G}(x)$, the maximal evolution time is still $\widetilde{O}(\gamma^{-1})$. □

4.2 Step 2: estimate the O-weighted CDF

To estimate the expectation value of O, consider the following quantum circuit:

Define the random variables X_j, Y_j be as follows: for $W = I$, $X_j := 1$ if the outcome is 0, and $X_j := -1$ if the outcome is 1. For $W = S$, $Y_j := -1$ if the outcome is 0, and $Y_j := 1$ if the outcome is 1.

Then, we have the following claim on the expectation of the random variables X_j, Y_j:

Claim 4.3 (A variant of Hadamard test)**.** *For any $j \in \mathbb{Z}$, the random variable $X_j + iY_j$ is an un-biased estimator for $\langle\phi_0| Oe^{-ij\tau H} |\phi_0\rangle$.*

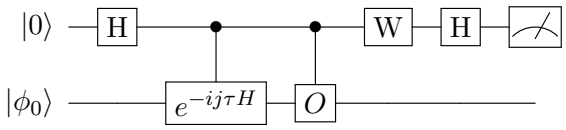


Figure 2: Quantum circuit parameterized by j. H is the Hadamard gate and W is either I or a phase gate S.

We can expand $\langle\phi_0| Oe^{-ij\tau H} |\phi_0\rangle$ in the eigenbasis of H (which is also an eigenbasis of O):

$$\begin{aligned}\langle\phi_0| Oe^{-ij\tau H} |\phi_0\rangle &= \sum_{k,k'} c_k^* c_{k'} e^{-ij\tau\lambda_k} \langle\psi_k| O |\psi_k'\rangle \\ &= \sum_k p_k O_k e^{-ij\tau\lambda_k}, \end{aligned} \tag{12}$$

where the last step follows from the simultaneous diagonalization of O and H, and $O_k := \langle\psi_k| O |\psi_k\rangle$. We may assume that $|O_k| \le 1$ for any $k \in \mathbb{N}$.

Inspired by the ground state energy estimation algorithm in [LT21], we define the O-weighted "density function" for the observable as follows:

$$p_O(x) := \sum_k p_k O_k \delta(x - \tau\lambda_k). \tag{13}$$

Note that $p_O(x)$ can be negative at some points.

Suppose the eigenvalues of τH is within $[-\pi/3, \pi/3]$. Then, we define the O-weighted CDF and ACDF for $p_O(x)$ similar to [LT21]:

$$C_O(x) := (H * p_O)(x), \quad \widetilde{C_O}(x) := (F * p_O)(x), \tag{14}$$

where H is the 2π-periodic Heaviside function and $F = F_{d,\delta}$ is the Fourier approximation of H constructed by Lemma A.8. It is easy to verify that $C_O(x)$ equals to $\sum_k p_k O_k \mathbf{1}_{x \ge p_k O_k}$ for any $x \in [-\pi/3, \pi/3]$.

The following lemma gives an unbiased estimator for the O-weighted ACDF.

Lemma 4.4 (Estimating the O-weighted ACDF)**.** *For any $x \in [-\pi, \pi]$, there exists an unbiased estimator $\overline{G_O}(x)$ for the O-weighted ACDF $\widetilde{C_O}(x)$ with variance $\widetilde{O}(1)$.*

Furthermore, $\overline{G_O}(x)$ runs the quantum circuit (Figure 2) with expected total evolution time $O(\tau d/\log(d))$, where d is the Fourier degree of F .

Proof. $\widetilde{C_O}(x)$ can be expanded in the following way:

$$\begin{aligned}
\widetilde{C_O}(x) &= (F * p_O)(x) \qquad (15)\\
&= \int_{-\pi}^{\pi} F(x-y)p_O(y)\mathrm{d}y \\
&= \sum_{|j|\le d} \int_{-\pi}^{\pi} \widehat{F}_j e^{ij(x-y)} p_O(y)\mathrm{d}y \\
&= \sum_{|j|\le d} \widehat{F}_j e^{ijx} \int_{-\pi}^{\pi} p_O(y) e^{-ijy}\mathrm{d}y \\
&= \sum_{|j|\le d} \widehat{F}_j e^{ijx} \sum_k p_k O_k e^{-ij\tau\lambda_k} \\
&= \sum_{|j|\le d} \widehat{F}_j e^{ijx} \cdot \langle\phi_0| O e^{-ij\tau H} |\phi_0\rangle , \qquad (16)
\end{aligned}$$

where the third step follows from the Fourier expansion of $F(x-y)$, the fifth step follows from the property of Dirac's delta function, and the last step follows from the definition of p_k and the eigenvalues of matrix exponential.

Define an estimator $G(x;\mathbf{J},\mathbf{Z})$ as follows:

$$G(x;\mathbf{J},\mathbf{Z}) := \mathcal{F}\cdot \mathbf{Z} e^{i(\theta_{\mathbf{J}}+\mathbf{J}x)}, \qquad (17)$$

where θ_j is defined by $\widehat{F}_j = |\widehat{F}_j| e^{i\theta_j}$, $\mathbf{Z} = X + iY$ measured from the quantum circuit (Figure 2) with parameter $j = \mathbf{J}$, and $\mathcal{F} = \sum_{|j|\le d} |\widehat{F}_j|$.

Then, we show that $G(x;\mathbf{J},\mathbf{Z})$ is un-biased:

$$\begin{aligned}
\mathbb{E}[G(x;\mathbf{J},\mathbf{Z})] &= \sum_{|j|\le d} \mathbb{E}\left[(X_j + iY_j) e^{i(\theta_j+jx)} |\widehat{F}_j|\right] \qquad (18)\\
&= \sum_{|j|\le d} \widehat{F}_j e^{ijx} \cdot \mathbb{E}\left[X_j + iY_j\right] \\
&= \sum_{|j|\le d} \widehat{F}_j e^{ijx} \cdot \langle\phi_0| O e^{-ij\tau H} |\phi_0\rangle \\
&= \widetilde{C}(x), \qquad (19)
\end{aligned}$$

where the third step follows from Claim 4.3. Moreover, the variance of G can be upper-bounded by:

$$\begin{aligned}
\mathbf{Var}[G(x;\mathbf{J},\mathbf{Z})] &= \mathbb{E}[|G(x;\mathbf{J},\mathbf{Z})|^2] - |\,\mathbb{E}[G(x;\mathbf{J},\mathbf{Z})]|^2 \qquad (20)\\
&\le \mathbb{E}[|G(x;\mathbf{J},\mathbf{Z})|^2] \\
&\le 2\mathcal{F}^2, \qquad (21)
\end{aligned}$$

where the third step follows from $|e^{i(\theta_J+Jx)}| = 1$, and the last step follows from $X_j, Y_j \in \{\pm 1\}$. By Lemma A.8, we know that $|\widehat{F}_j| = O(1/|j|)$. Hence, we have $\mathcal{F} = \sum_{|j|\le d} O(1/|j|) = O(\log d)$. Thus, $\mathbf{Var}[G(x;\mathbf{J},\mathbf{Z})] = O(\log^2(d))$.

The expected total evolution time is

$$\mathcal{T}_{\mathsf{tot}} := \mathbb{E}[|J|] = \tau \sum_{|j|\le d} |j| \cdot \frac{|\widehat{F}_j|}{\mathcal{F}} = O(\tau d/\log(d)). \qquad (22)$$

The lemma is then proved. □

The following lemma shows that the O-weighted CDF $C_O(x)$ can be approximated by the O-weighted ACDF $\widetilde{C_O}(x)$:

Lemma 4.5 (Approximating the O-weighted CDF)**.** *For any $\epsilon > 0$, $0 < \delta < \pi/6$, let $F(x) := F_{d,\delta}(x)$ constructed by Lemma A.8 with approximation error $\eta\epsilon/8$. Then, for any $x \in [-\pi/3, \pi/3]$, it holds that:*

$$C_O(x-\delta) - \eta\epsilon/8 \le \widetilde{C_O}(x) \le C_O(x+\delta) + \eta\epsilon/8. \tag{23}$$

The proof is very similar to Lemma A.9, so we omit it here.

We can take $\delta := \tau\gamma/5$ and let $x_{\mathsf{good}} := x^\star + \tau\gamma/2$. Then, by Claim 4.1, we know that x_{good} is good for λ_0, i.e., $[x_{\mathsf{good}} - \tau\gamma, x_{\mathsf{good}} + \tau\gamma] \subset (\tau\lambda_0, \tau\lambda_1)$. Hence, $\widetilde{C}_O(x_{\mathsf{good}})$ satisfies

$$\left|\widetilde{C}_O(x_{\mathsf{good}}) - p_0O_0\right| \le \eta\epsilon/8. \tag{24}$$

The following lemma shows how to estimate $\widetilde{C_O}(x_{\mathsf{good}})$, which is very similar to Lemma 4.2.

Lemma 4.6 (Estimating p_0O_0)**.** *For any $\epsilon_1, \nu \in (0,1)$, p_0O_0 can be estimated with multiplicative error $1 \pm O(\epsilon_1)$ with probability $1-\nu$ by runs the quantum circuit (Figure 1) $\widetilde{O}(\epsilon_1^{-2}\eta^{-2})$ times with expected total evolution time $\widetilde{O}(\gamma^{-1}\epsilon_1^{-2}\eta^{-2})$ and maximal evolution time $O(\gamma^{-1})$.*

4.3 Putting it all together

In this section, we will put the components together and prove the following main theorem, which gives an algorithm for the ground state property estimation.

Theorem 4.7 (Ground state property estimation with commutative observable, restate)**.** *Suppose $p_0 \ge \eta$ for some known η. Then, for any $\epsilon, \nu \in (0,1)$, the ground state property $\langle\psi_0| O |\psi_0\rangle$ can be estimated within additive error at most ϵ with probability $1-\nu$, such that:*

1. *the number of times running the quantum circuits (Figure 1 and 2) is $\widetilde{O}(\epsilon^{-2}\eta^{-2})$,*
2. *the expected total evolution time is $\widetilde{O}(\gamma^{-1}\epsilon^{-2}\eta^{-2})$,*
3. *the maximal evolution time is $\widetilde{O}(\gamma^{-1})$.*

Proof. By Lemma 4.2, we obtain $\overline{p_0}$ such that

$$\left|\overline{p_0} - p_0\right| \le O(\eta\epsilon_0), \tag{25}$$

where ϵ_0 will be chosen shortly.

By Lemma 4.6, we obtain $\overline{p_0O_0}$ such that

$$\left|\overline{p_0O_0} - p_0O_0\right| \le O(\eta\epsilon_1), \tag{26}$$

where ϵ_1 will be chosen shortly.

Then, we have

$$\left|\frac{\overline{p_0O_0}}{\overline{p_0}} - O_0\right| = \left|\frac{\overline{p_0O_0}}{\overline{p_0}} - \frac{p_0O_0}{\overline{p_0}} + \frac{p_0O_0}{\overline{p_0}} - \frac{p_0O_0}{p_0}\right| \tag{27}$$

$$\begin{aligned}
&\leq \frac{|\overline{p_0O_0} - p_0O_0|}{\overline{p_0}} + |p_0O_0|\left|\frac{1}{\overline{p_0}} - \frac{1}{p_0}\right| \\
&\leq \frac{O(\eta\epsilon_1)}{p_0 - O(\eta\epsilon_0)} + |p_0O_0|\left|\frac{1}{p_0 - O(\eta\epsilon_0)} - \frac{1}{p_0}\right| \\
&\leq \frac{O(\eta\epsilon_1)}{\eta - O(\eta\epsilon_0)} + |p_0O_0|\left|\frac{1}{p_0 - p_0O(\epsilon_0)} - \frac{1}{p_0}\right| \\
&\leq O(\epsilon_1)(1 - O(\epsilon_0)) + |O_0|(1 + O(\epsilon_0) - 1) \\
&\leq O(\epsilon_0 + \epsilon_1),
\end{aligned} \tag{28}$$

where the second step follows from the triangle inequality, the third step follows from Eqs. (25) and (26), the third step follows from $p_0 \geq \eta$, the fifth step follows from $\frac{1}{1-x} \leq 1 + O(x)$ for $x \in (0, 1)$.

Hence, if we take $\epsilon_0 = \epsilon_1 = O(\epsilon)$, we will achieve additive error at most ϵ.

For the success probability, we can make Eq.(25) hold with probability $1 - \nu/2$ in Lemma 4.2 and Eq.(26) hold with probability $1 - \nu/2$ in Lemma 4.6. Then, by the union bound, we get a good estimate with probability at least $1 - \nu$.

The computation costs follow directly from Lemma 4.2 and Lemma 4.6. And the proof of the theorem is then completed. □

5 Algorithm for General Unitary Observables

In this section, we will prove the following theorem for unitary observables in the general case:

Theorem 5.1 (Ground state property estimation with general unitary observable)**.** *Suppose $p_0 \geq \eta$ for some known η and the spectral gap of the Hamiltonian H is at least γ. For any $\epsilon, \nu \in (0, 1)$, there exists an algorithm for estimating the ground state property $\langle\psi_0| O |\psi_0\rangle$ within additive error at most ϵ with probability at least $1 - \nu$, such that:*

1. *the expected total evolution time is $\widetilde{O}(\gamma^{-1}\epsilon^{-2}\eta^{-2})$*
2. *the maximal evolution time is $\widetilde{O}(\gamma^{-1})$.*

In the following parts, we will first introduce the 2-d O-weighted density function and CDF, which extend the commuting observables to the general case. Then, we will show how to combine them with the overlap estimation in Section 4.1 for proving Theorem 5.1.

5.1 2-d O-weighted density function and CDF

Let $|\phi_0\rangle = \sum_k c_k |\psi_k\rangle$ where $|c_k|^2 = p_k$. In general, O and H may not commute. Hence, we consider a more symmetric form of expectation: $\langle\phi_0| e^{-ij\tau H} O e^{-ij'\tau H} |\phi_0\rangle$, which can be expanded in the eigenbasis of H as follows:

$$\begin{aligned}
\langle\phi_0| e^{-ij\tau H} O e^{-ij'\tau H} |\phi_0\rangle &= \sum_{k,k'} c_k^* c_{k'} e^{-ij\tau\lambda_k} e^{-ij'\tau\lambda_{k'}} \langle\psi_k| O |\psi'_k\rangle \\
&= \sum_{k,k'} c_k^* c_{k'} e^{-ij\tau\lambda_{k'}} e^{-ij'\tau\lambda_{k'}} \langle\psi_k| O |\psi_{k'}\rangle
\end{aligned} \tag{29}$$

Algorithm 2 Ground State Property Estimation (Commutative Case)

procedure EstimateGSProp($\epsilon, \tau, \eta, \gamma, \nu$)
  $\delta \leftarrow O(\tau\gamma)$, $d \leftarrow O(\delta^{-1}\log(\delta^{-1}\epsilon^{-1}\eta^{-1}))$
  for $j \leftarrow -d, \ldots, d$ **do**
    Compute $\widehat{F}_j := \widehat{F}_{d,\delta,j}$ and θ_j
  end for
  ▷ Estimate the ground state energy
  $x^\star \leftarrow$ EstimateGSE($\gamma/8, \tau, \eta, \nu/10$)
  $x_{\mathsf{good}} \leftarrow x^\star + \tau\gamma/2$
  ▷ Generate samples from the Hadamard test circuits
  $N_g \leftarrow O(\log(1/\nu))$, $K \leftarrow O(\epsilon^{-2})$
  for $k \leftarrow 1, \ldots, N_g K$ **do**
    Sample (Z_k, J_k) from the quantum circuit (Figure 1)
    Sample (Z'_k, J'_k) from the quantum circuit (Figure 2)
  end for
  ▷ Estimate p_0
  for $i \leftarrow 1, \ldots, N_g$ **do**
    $\overline{G}_i \leftarrow \frac{1}{K}\sum_{j=1}^{K} \overline{G}(x_{\mathsf{good}}; Z_{(i-1)K+j}, J_{(i-1)K+j})$
  end for
  $\overline{p_0} \leftarrow \mathrm{median}(\overline{G}_1, \ldots, \overline{G}_{N_g})$
  ▷ Estimate $p_0 O_0$
  for $i \leftarrow 1, \ldots, N_g$ **do**
    $\overline{G}'_i \leftarrow \frac{1}{K}\sum_{j=1}^{K} \overline{G}(x_{\mathsf{good}}; Z'_{(i-1)K+j}, J'_{(i-1)K+j})$
  end for
  $\overline{p_0 O_0} \leftarrow \mathrm{median}(\overline{G}'_1, \ldots, \overline{G}'_{N_g})$
  return $\overline{p_0 O_0}/\overline{p_0}$
end procedure

Similar to the commutative case, we define a 2-d O-weighted density function:

$$p_{O,2}(x,y) := \sum_{k,k'} c_k^* c_{k'} O_{k,k'} \delta(x - \tau\lambda_k)\delta(y - \tau\lambda_{k'}), \tag{30}$$

where $O_{k,k'} := \langle\psi_k| O |\psi_{k'}\rangle$. Then, define the corresponding 2-d O-weighted CDF function as follows:

$$C_{O,2}(x) := (H_2 * p_{O,2})(x,y), \tag{31}$$

where $H_2(x,y) := H(x) \cdot H(y)$, the 2-d 2π-periodic Heaviside function.

We first justify that $C_{O,2}$ is indeed a CDF of $p_{O,2}$ in $[-\pi/3, \pi/3]$:

$$\begin{aligned} C_2(x,y) &= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} H_2(x-u, y-v)p(u,v)\mathrm{d}u\mathrm{d}v && (32)\\ &= \sum_{k,k'} c_k^* c_{k'} O_{k,k'} \cdot \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} H_2(x-u, y-v)\delta(u-\tau\lambda_k)\delta(v-\tau\lambda_{k'})\mathrm{d}u\mathrm{d}v \\ &= \sum_{k,k'} c_k^* c_{k'} O_{k,k'} \cdot H(x-\tau\lambda_k)H(y-\tau\lambda_{k'}) \\ &= \sum_{k,k'} c_k^* c_{k'} O_{k,k'} \cdot \mathbf{1}_{x\geq\tau\lambda_k, y\geq\tau\lambda_{k'}} \\ &= \sum_{\substack{k:\tau\lambda_k\leq x,\\ k':\tau\lambda_{k'}\leq y}} c_k^* c_{k'} O_{k,k'}. && (33) \end{aligned}$$

Hence, the definition of $C_{O,2}$ is reasonable.

Then, we show that $C_{O,2}$ can be approximated similar to the 1-d case. Let $F_2(x)$ be the 2-d approximated Heaviside function, i.e.,

$$F_2(x,y) := F(x)\cdot F(y). \tag{34}$$

The 2-d O-weighted approximated CDF (ACDF) is defined to be

$$\widetilde{C_{O,2}}(x,y) := (F_2 * p_{O,2})(x,y). \tag{35}$$

The following lemma shows that $\widetilde{C_{O,2}}(x,y)$ is close to $C_{O,2}(x',y')$ for some (x',y') close to (x,y).

Lemma 5.2 (Approximation ratio of the 2-d O-weighted ACDF)**.** *For any $\epsilon > 0$, $0 < \delta < \pi/6$, let $F_2(x,y) := F_{d,\delta}(x)\cdot F_{d,\delta}(y)$ constructed by Lemma A.8. Then, for any $x, y \in [-\pi/3, \pi/3]$, the 2-d O-weighted ACDF $\widetilde{C_{O,2}}(x,y) = (F_2 * p_{O,2})(x,y)$ satisfies:*

$$C_{O,2}(x-\delta, y-\delta) - 2\epsilon \leq \widetilde{C_{O,2}}(x,y) \leq C_{O,2}(x+\delta, y+\delta) + 2\epsilon. \tag{36}$$

Proof. By (2) in Lemma A.8, we have

$$|F(x) - H(x)| \leq \epsilon \quad \forall x \in [-\pi+\delta, -\delta] \cup [\delta, \pi-\delta], \tag{37}$$

which implies that for all $x, y \in [-\pi+\delta, -\delta] \cup [\delta, \pi-\delta]$,

$$\begin{aligned} |F_2(x,y) - H_2(x,y)| &\leq |F(x)F(y) - H(x)H(y)| && (38)\\ &= |F(x)F(y) - F(x)H(y) + F(x)H(y) - H(x)H(y)| \\ &\leq F(x)|F(y) - H(y)| + H(y)|F(x) - H(x)| \\ &\leq (F(x) + H(y))\epsilon \\ &\leq 2\epsilon, && (39) \end{aligned}$$

where the last step follows from $F(x) \in [0,1]$ by (1) in Lemma A.8. Furthermore, we also have for $x \in [-\delta, \delta]$, $y \in [-\pi+\delta, -\delta]$,

$$\begin{aligned} |F_2(x,y) - H_2(x,y)| &\leq |F(x)F(y) - H(x)H(y)| && (40)\\ &= |F(x)F(y)| && (H(y) = 0) \\ &\leq F(y) \\ &\leq \epsilon. && (41) \end{aligned}$$

Similarly, for $x \in [-\pi+\delta, -\delta]$, $y \in [-\delta, \delta]$,

$$|F_2(x,y) - H_2(x,y)| \le \epsilon. \tag{42}$$

Define $F_{L,2} := F_2(x-\delta, y-\delta)$ such that

$$\begin{aligned} |F_{L,2}(x) - H_2(x)| \le 2\epsilon \quad \forall (x,y) \in & [-\pi+2\delta, 0] \times [-\pi+2\delta, \pi] \\ & \cup [-\pi+2\delta, \pi] \times [-\pi+2\delta, 0] \\ & \cup [2\delta, \pi] \times [2\delta, \pi]. \end{aligned} \tag{43}$$

For $\widetilde{C_{L,2}}(x,y) := (F_{L,2} * p_{O,2})(x,y)$, we have $\widetilde{C_{L,2}}(x,y) = \widetilde{C_{O,2}}(x-\delta, y-\delta)$.

Let $p_2 := p_{O,2}$. Then, for any $x, y \in [-\pi/3, \pi/3]$, we have

$$\begin{aligned} &\left|C_{O,2}(x,y) - \widetilde{C_{L,2}}(x,y)\right| = \left|\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} p_2(x-u, y-v)(H_2(u,v) - F_{L,2}(u,v))\mathrm{d}u\mathrm{d}v\right| \qquad (44) \\ \le\ & \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} p_2(x-u, y-v)|H_2(u,v) - F_{L,2}(u,v)|\mathrm{d}u\mathrm{d}v \\ =\ & \left(\int_{-\pi}^{0}\int_{-\pi}^{\pi} + \int_{0}^{\pi}\int_{-\pi}^{0} + \int_{2\delta}^{\pi}\int_{2\delta}^{\pi}\right) p_2(x-u, y-v))|H_2(u,v) - F_{L,2}(u,v)|\mathrm{d}u\mathrm{d}v \\ & + \left(\int_{0}^{2\delta}\int_{0}^{\pi} + \int_{0}^{\pi}\int_{0}^{2\delta} - \int_{0}^{2\delta}\int_{0}^{2\delta}\right) p_2(x-u, y-v))|H_2(u,v) - F_{L,2}(u,v)|\mathrm{d}u\mathrm{d}v \\ \le\ & 2\epsilon \cdot \left(\int_{-\pi}^{0}\int_{-\pi}^{\pi} + \int_{0}^{\pi}\int_{-\pi}^{0} + \int_{2\delta}^{\pi}\int_{2\delta}^{\pi}\right) p_2(x-u, y-v)\mathrm{d}u\mathrm{d}v \\ & + \left(\int_{0}^{2\delta}\int_{0}^{\pi} + \int_{0}^{\pi}\int_{0}^{2\delta} - \int_{0}^{2\delta}\int_{0}^{2\delta}\right) p_2(x-u, y-v))|H_2(u,v) - F_{L,2}(u,v)|\mathrm{d}u\mathrm{d}v \\ \le\ & 2\epsilon + \left(\int_{0}^{2\delta}\int_{0}^{\pi} + \int_{0}^{\pi}\int_{0}^{2\delta} - \int_{0}^{2\delta}\int_{0}^{2\delta}\right) p_2(x-u, y-v))|H_2(u,v) - F_{L,2}(u,v)|\mathrm{d}u\mathrm{d}v \\ \le\ & 2\epsilon + \left(\int_{0}^{2\delta}\int_{0}^{\pi} + \int_{0}^{\pi}\int_{0}^{2\delta} - \int_{0}^{2\delta}\int_{0}^{2\delta}\right) p_2(x-u, y-v)\mathrm{d}u\mathrm{d}v \\ =\ & 2\epsilon + \left(\int_{x-2\delta}^{x}\int_{y-\pi}^{y} + \int_{x-\pi}^{x}\int_{y-2\delta}^{y} - \int_{x-2\delta}^{x}\int_{y-2\delta}^{y}\right) p_2(u,v)\mathrm{d}u\mathrm{d}v \qquad (45) \\ =\ & 2\epsilon + C_{O,2}(x,y) - C_{O,2}(x-2\delta, y-2\delta), \end{aligned}$$

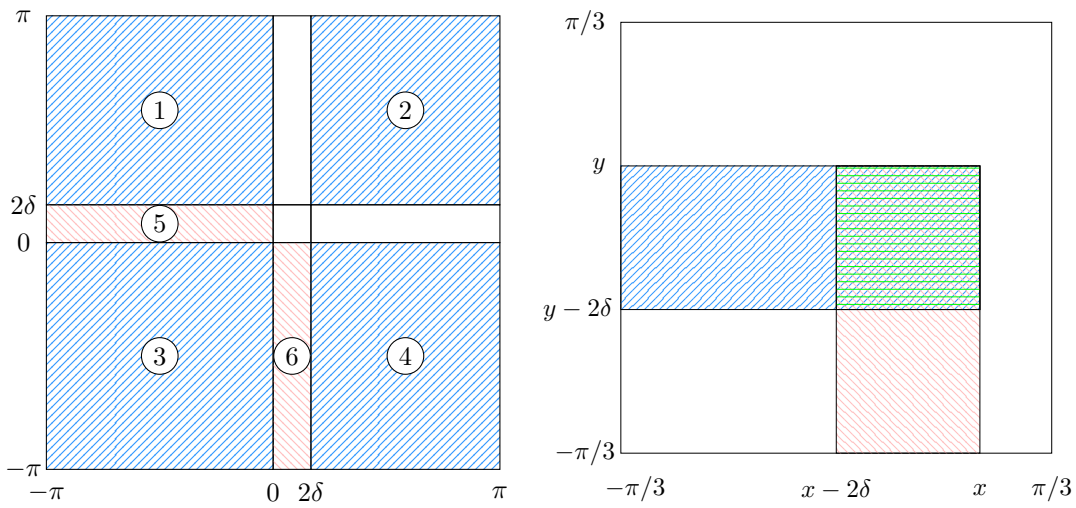
where the second step follows from Cauchy-Schwarz inequality, the third step follows from partitioning the integration region, the forth step follows from Eq. (43) and the fact that $p(x,y)$ is supported in $[-\pi/3, \pi/3] \times [-\pi/3, \pi/3]$ and $\delta < \pi/6$ (see Figure 3 (a)), the fifth step follows from $p_{O,2}(x)$ is a density function, the last step follows from $C_{O,2}(x)$ is the CDF of $p_{O,2}(x)$ in $[-\pi, \pi] \times [-\pi, \pi]$ and $x, y \in [-\pi/3, \pi/3]$ (see Figure 3 (b)).

Hence, we have

$$\begin{aligned} \widetilde{C_{L,2}}(x,y) &\ge C_{O,2}(x,y) - (2\epsilon + C_{O,2}(x,y) - C_{O,2}(x-2\delta, y-2\delta)) \\ &= C_{O,2}(x-2\delta, y-2\delta) - 2\epsilon, \end{aligned} \tag{46}$$

which proves the first inequality:

$$\widetilde{C_{O,2}}(x-\delta, y-\delta) \ge C_{O,2}(x-2\delta, y-2\delta) - 2\epsilon. \tag{47}$$



(a) (b)

Figure 3: (a) is the integral region for Eq. (44), where the integral in regions 1-6 can be upper bounded by Eq. (43). (b) is the integral region for Eq. (45).

Similarly, we can define $F_{R,2} := F_2(x+\delta, y+\delta)$ and $\widetilde{C_{R,2}}(x,y) := (F_{R,2} * p_2)(x,y)$. We can show that

$$\left|C_{O,2}(x,y) - \widetilde{C_{R,2}}(x,y)\right| \leq 2\epsilon + C_{O,2}(x+2\delta, y+2\delta) - C_{O,2}(x,y), \tag{48}$$

which gives

$$\widetilde{C_{O,2}}(x+\delta, y+\delta) \leq C_{O,2}(x+2\delta, y+2\delta) + 2\epsilon. \tag{49}$$

The lemma is then proved. □

5.2 Estimating the 2-d ACDF

We use the following parameterized quantum circuit to estimate the 2-d O-weighted ACDF $\widetilde{C_{O,2}}(x,y)$.

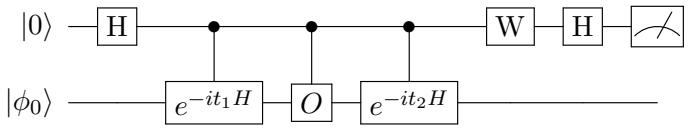


Figure 4: Quantum circuit parameterized by t_1, t_2. H is the Hadamard gate and W is either I or a phase gate S.

Lemma 5.3 (Estimate 2-d O-weighted ACDF)**.** *For any $x, y \in [-\pi/3, \pi/3]$, for any $\epsilon_2, \nu \in (0,1)$, we can estimate $\widetilde{C_{O,2}}(x,y)$ with additive error $\eta\epsilon$ with probability $1-\nu$ by running the quantum circuit (Figure 4) $O(\epsilon_2^{-2}\eta^{-2}\log(1/\nu))$ times with maximal evolution time $\widetilde{O}(\gamma^{-1})$ and total expected evolution time $\widetilde{O}(\gamma^{-1}\epsilon_2^{-1}\eta^{-1})$.*

Proof. $\widetilde{C_{O,2}}(x,y)$ can be expanded in the following way:

$$\begin{aligned}
\widetilde{C_{O,2}}(x,y) &= (F_2 * p_2)(x,y) && (50)\\
&= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} F_2(x-u,y-v)p_2(u,v)\mathrm{d}u\mathrm{d}v\\
&= \sum_{|j|\le d,|j'|\le d}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \widehat{F}_j\widehat{F}_{j'}e^{ij(x-u)}e^{ij'(y-v)}p_2(u,v)\mathrm{d}u\mathrm{d}v\\
&= \sum_{|j|\le d,|j'|\le d} \widehat{F}_j\widehat{F}_{j'}e^{i(jx+j'y)}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} p_2(u,v)e^{-iju}e^{-ij'v}\mathrm{d}u\mathrm{d}v\\
&= \sum_{|j|\le d,|j'|\le d} \widehat{F}_j\widehat{F}_{j'}e^{i(jx+j'y)}\sum_{k,k'} c_k^* c_k O_{k,k'}e^{-ij\tau\lambda_k}e^{-ij'\tau\lambda_{k'}}\\
&= \sum_{|j|\le d,|j'|\le d} \widehat{F}_j\widehat{F}_{j'}e^{i(jx+j'y)}\cdot\langle\phi_0|\, e^{-ij\tau H}Oe^{-ij'\tau H}\,|\phi_0\rangle\,, && (51)
\end{aligned}$$

To estimate $\langle\phi_0|\, e^{-ij\tau H}Oe^{-ij'\tau H}\,|\phi_0\rangle$, we use the multi-level Monte Carlo method. Define a random variables J, J' with support $\{-d,\cdots,d\}$ such that

$$\Pr[J=j,J'=j'] = \frac{|\widehat{F}_j||\widehat{F}_{j'}|}{\mathcal{F}^2}, \tag{52}$$

where $\mathcal{F} := \sum_{|j|\le d}|\widehat{F}_j|$. Then, let $Z := X_{J,J'}+iY_{J,J'} \in \{\pm 1\pm i\}$. Define an estimator $\overline{G_2}(x;J,J',Z)$ as follows:

$$\overline{G_2}(x,y;J,Z) := \mathcal{F}^2\cdot Ze^{i(\theta_J+Jx)}e^{i(\theta_{J'}+J'y)}, \tag{53}$$

where θ_j is defined by $\widehat{F}_j = |\widehat{F}_j|e^{i\theta_j}$, and similar definition for $\theta_{j'}$. Then, we show that $\overline{G_2}(x,y;J,Z)$ is un-biased:

$$\begin{aligned}
\mathbb{E}[\overline{G_2}(x,y;J,J',Z)] &= \sum_{|j|\le d,|j'|\le d}\mathbb{E}\left[(X_{j,j'}+iY_{j,j'})e^{i(\theta_j+jx)}e^{i(\theta_{j'}+j'y)}|\widehat{F}_j||\widehat{F}_{j'}|\right] && (54)\\
&= \sum_{|j|\le d,|j'|\le d}\widehat{F}_j\widehat{F}_{j'}e^{ijx}e^{ij'y}\cdot\mathbb{E}\left[X_{j,j'}+iY_{j,j'}\right]\\
&= \sum_{|j|\le d,|j'|\le d}\widehat{F}_j\widehat{F}_{j'}e^{ijx}e^{ij'y}\cdot\langle\phi_0|\, e^{-ij\tau H}Oe^{-ij'\tau H}\,|\phi_0\rangle\\
&= \widetilde{C_2}(x,y), && (55)
\end{aligned}$$

where the third step follows from Claim A.1. Moreover, the variance of $\overline{G_2}$ can be upper-bounded by:

$$\begin{aligned}
\mathbf{Var}[\overline{G_2}(x,y;J,J',Z)] &= \mathbb{E}[|\overline{G_2}(x,y;J,J',Z)|^2] - |\,\mathbb{E}[\overline{G_2}(x,y;J,J',Z)]|^2 && (56)\\
&\le \mathbb{E}[|\overline{G_2}(x,y;J,J',Z)|^2]\\
&= \mathcal{F}^4\cdot\mathbb{E}[|X_{J,J'}+iY_{J,J'}|^2]\\
&= 2\mathcal{F}^4, && (57)
\end{aligned}$$

where the third step follows from $|e^{i(\theta_J+Jx)}| = |e^{i(\theta_{J'}+J'y)}| = 1$, and the last step follows from $X_{j,j'}, Y_{j,j'} \in \{\pm 1\}$.

By Lemma A.8, we know that $\mathcal{F} = \widetilde{O}(1)$. Hence, we have for all $x, y \in [-\pi/3, \pi/3]$,

$$\mathbb{E}[\overline{G_2}(x,y)] = \widetilde{C_{O,2}}(x,y), \quad \text{and} \quad \mathbf{Var}[\overline{G_2}(x,y)] = \widetilde{O}(1). \tag{58}$$

Then, using median-of-means estimator, we can obtain an ϵ_2-additive error estimate of $\widetilde{C_{O,2}}(x,y)$ with probability $1-\nu$ using $O(\epsilon_2^{-2}\eta^{-2}\log(1/\nu))$ samples.

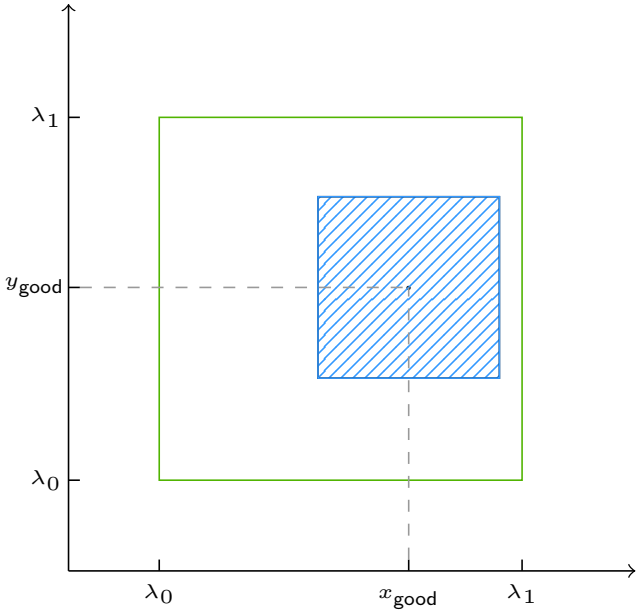
The maximal evolution time is $2d = \widetilde{O}(\gamma^{-1})$. And the expected evolution time for one trial is

$$\tau \sum_{|j|,|j'|\le d} (j+j')\frac{|\widehat{F}_j||\widehat{F}_{j'}|}{\mathcal{F}^2} = 2\tau \sum_{|j|\le d} j\frac{|\widehat{F}_j|}{\mathcal{F}} = O(\tau d/\log(d)). \tag{59}$$

Hence, the total expected evolution time is $\widetilde{O}(\gamma^{-1}\epsilon_2^{-2}\eta^{-2})$.

The lemma is then proved. □

(a)



(b)

Figure 5: (a) shows a point that is *good* for λ_0, where the blue interval is the approximation region such that $\widetilde{C_O}(x_{\text{good}})$ is close to $C(x)$ for some x in this interval. (b) shows a good point in the 2-d case, where in the green square, the 2-d O-weighted CDF $C_{O,2}$ takes the same value $C_{O,2}(\lambda_0,\lambda_0)$. And the blue square is the approximation region of $(x_{\text{good}}, y_{\text{good}})$ such that $\widetilde{C_{O,2}}(x_{\text{good}}, y_{\text{good}})$ is close to some $C_{O,2}(x,y)$ in this region.

Similar to the 1-d case, we can construct a "good" point for (λ_0,λ_0) via the following claim.

Claim 5.4 (Construct 2-d good point)**.** *Let γ be the spectral gap of the Hamiltonian H. Let $x_{\text{good}} := x^\star + \tau\gamma/2$ where $x^\star$ is the output of* EstimateGSE$(\gamma/8, \tau, \eta, \nu/10)$ *(Algorithm 1). Then,*

$(x_{\mathsf{good}}, x_{\mathsf{good}})$ *is good for* (λ_0, λ_0). *In particular, for any* $\epsilon \in (0,1)$, *if the approximation error of* $F(x)$ *is set to be* $\epsilon\eta$, *then*

$$\left|\widetilde{C_{O,2}}(x_{\mathsf{good}}, x_{\mathsf{good}}) - C_{O,2}(\lambda_0, \lambda_0)\right| \leq 2\epsilon\eta. \tag{60}$$

Proof. By Claim 4.1, we know that x_{good} is good for λ_0, i.e., $[x_{\mathsf{good}} - \delta, x_{\mathsf{good}} + \delta]$ is contained in $[\lambda_0, \lambda_1)$. It also holds in the 2-d case for $(x_{\mathsf{good}}, x_{\mathsf{good}})$. Then, by Lemma 5.3, we have

$$C_{O,2}(x_{\mathsf{good}} - \delta, x_{\mathsf{good}} - \delta) - 2\epsilon\eta \leq \widetilde{C_{O,2}}(x_{\mathsf{good}}, x_{\mathsf{good}}y) \leq C_{O,2}(x_{\mathsf{good}} + \delta, x_{\mathsf{good}} + \delta) + 2\epsilon\eta. \tag{61}$$

The claim then follows from $C_{O,2}(x,y) = C_{O,2}(\lambda_0, \lambda_0)$ for any $(x,y) \in [\lambda_0, \lambda_1) \times [\lambda_0, \lambda_1)$. □

5.3 Putting it all together

The main algorithm for the ground state property estimation will first estimate the ground state energy λ_0 and the overlap p_0, which are described in Section 4.1. Then, by Lemma 5.3 and Claim 5.4, the weighted expectation $p_0 O_0$ can also be estimated. Hence, we will obtain an estimate for $O_0 = \langle\psi_0| O |\psi_0\rangle$.

Algorithm 3 Ground State Property Estimation (General Case)

procedure EstimateGSProp$(\epsilon, \tau, \eta, \gamma, \nu)$
  $\delta \leftarrow O(\tau\gamma)$, $d \leftarrow O(\delta^{-1}\log(\delta^{-1}\epsilon^{-1}\eta^{-1}))$
  for $j \leftarrow -d, \ldots, d$ **do**
    Compute $\widehat{F}_j := \widehat{F}_{d,\delta,j}$ and θ_j
  end for
  ▷ Estimate the ground state energy
  $x^\star \leftarrow$ EstimateGSE$(\gamma/8, \tau, \eta, \nu/10)$
  $x_{\mathsf{good}} \leftarrow x^\star + \tau\gamma/2$
  ▷ Generate samples from the Hadamard test circuits
  $B \leftarrow O(\log(1/\nu))$, $K \leftarrow \widetilde{O}(\epsilon^{-2})$
  for $k \leftarrow 1, \ldots, BK$ **do**
    Sample (Z_k, J_k) from the quantum circuit (Figure 1)
    Sample $(Z''_k, J''_{k,1}, J''_{k,2})$ from the quantum circuit (Figure 4)
  end for
  ▷ Estimate p_0
  for $i \leftarrow 1, \ldots, B$ **do**
    $\overline{G}_i \leftarrow \frac{1}{K}\sum_{j=1}^{K} \overline{G}(x_{\mathsf{good}}; Z_{(i-1)K+j}, J_{(i-1)K+j})$
  end for
  $\overline{p_0} \leftarrow \mathrm{median}(\overline{G}_1, \ldots, \overline{G}_B)$
  ▷ Estimate $p_0 O_0$
  for $i \leftarrow 1, \ldots, B$ **do**
    $\overline{G}''_i \leftarrow \frac{1}{K}\sum_{j=1}^{K} \overline{G_2}(x_{\mathsf{good}}, x_{\mathsf{good}}; Z''_{(i-1)K+j}, J''_{(i-1)K+j,1}, J''_{(i-1)K+j,2})$ ▷ Eq. (53)
  end for
  $\overline{p_0 O_0} \leftarrow \mathrm{median}(\overline{G}''_1, \ldots, \overline{G}''_B)$
  return $\overline{p_0 O_0}/\overline{p_0}$
end procedure

Proof of Theorem 5.1. We first analyze the estimation error of Algorithm 3. By Lemma 4.2, $\overline{p_0}$ (Line 19) has additive error at most $O(\eta\epsilon)$. By Lemma 5.3 and Claim 5.4, $\overline{p_0 O_0}$ (Line 24) has additive error at most $O(\eta\epsilon)$. Then, by a similar error propagation analysis in Theorem 4.7, we get that

$$\left|\frac{\overline{p_0 O_0}}{\overline{p_0}} - O_0\right| \leq O(\epsilon). \tag{62}$$

For the success probability, Algorithm 3 has three components: estimate ground state energy, estimate p_0, and estimate $p_0 O_0$. By our choice of parameters, each of them will fail with probability at most $\nu/3$. Hence, Algorithm 3 will succeed with probability at least $1-\nu$.

The maximal evolution time and the total expected evolution time follows from Theorem 3.1, Lemma 4.2, and Lemma 5.3. □

6 Handling non-unitary observables

One may notice that Algorithm 3 works only for unitary observables because it needs to use the circuit in Figure 4 to estimate $\langle\phi_0| e^{-it_2 H} O e^{-it_1 H} |\phi_0\rangle$ for certain $t_1, t_2 \in \mathbb{R}$, in which controlled-O must be a unitary operation. In this section, we show that under reasonable assumptions this algorithm can be modified to estimate the ground state property $\langle\psi_0| O |\psi_0\rangle$ where O is a general observable.

Before we present this result, one may wonder why it is necessary. After all, we can always decompose O into a linear combination of Pauli strings $O = \sum_{\vec{s}} w_{\vec{s}} P_{\vec{s}}$, and use Algorithm 3 to estimate each term $\mu_{\vec{s}} := \langle\psi_0| P_{\vec{s}} |\psi_0\rangle$ individually, and return $\sum_{\vec{s}} w_{\vec{s}} \mu_{\vec{s}}$ as the result. While this strategy works in principle, it might be not efficient enough to be practical, depending on the weights $w_{\vec{s}}$'s of Pauli strings in the linear expansion of O.

Alternatively, one can fix the issue of Algorithm 3 by designing a procedure for estimating $\langle\phi_0| e^{-it_2 H} O e^{-it_1 H} |\phi_0\rangle$ for arbitrary non-unitary O. Such quantities are utilized in the same way as before. We have followed this approach and found that it is possible when there is a block-encoding of O. Namely, suppose O is an n-qubit observable with $\|O\| \leq 1$ and U is an $(n+m)$-qubit unitary operator such that

$$(\langle 0^m| \otimes I) U (|0^m\rangle \otimes I) = \alpha^{-1} O \tag{63}$$

for some $\alpha \geq \|O\|$. Then we can still perform Hadamard test for U to estimate $\langle\phi_0| e^{-it_2 H} O e^{-it_1 H} |\phi_0\rangle$ for arbitrary $t_1, t_2 \in \mathbb{R}$. The main theorem of this section is stated below:

Theorem 6.1 (Ground state property estimation with block-encoded observable)**.** *Suppose $p_0 \geq \eta$ for some known η and the spectral gap of the Hamiltonian H is at least γ. Suppose we have access to the α-block-encoding of the observable O. For any $\epsilon, \nu \in (0,1)$, there exists an algorithm for estimating the ground state property $\langle\psi_0| O |\psi_0\rangle$ within additive error at most ϵ with probability at least $1-\nu$, such that:*

1. *the expected total evolution time is $\widetilde{O}(\gamma^{-1}\epsilon^{-2}\eta^{-2}\alpha^2)$,*
2. *the maximal evolution time is $\widetilde{O}(\gamma^{-1})$.*

Proof sketch of Theorem 6.1. The algorithm for handling non-unitary block-encoded observables is quite similar to Algorithm 3 for handling unitary observables, except that it relies on a different

procedure to estimate $\langle\phi_0| e^{-it_2H} O e^{-it_1H} |\phi_0\rangle$ for arbitrary $t_1, t_2 \in \mathbb{R}$. Here we briefly describe this procedure and defer the detailed analysis to Appendix B.

Let $C\text{-}V := |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes V$ be the controlled-V operation for arbitrary unitary operator V. Let $|\phi_0\rangle$ be an arbitrary n-qubit state. Consider the following procedure (as illustrated in Figure 6:

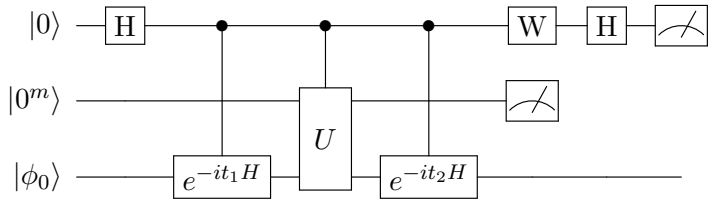


Figure 6: Quantum circuit parameterized by t_1, t_2. H is the Hadamard gate and W is either I or a phase gate S. U is the block-encoding of the non-unitary observable O.

1. Prepare the state $|0\rangle |0^m\rangle |\phi_0\rangle$.
2. Apply a Hadamard gate on the first register.
3. Apply a $C\text{-}e^{-iHt_1}$ on the first and third registers.
4. Apply $C\text{-}U$ on the current state, obtaining
$$\frac{1}{\sqrt{2}} \left(|0\rangle |0^m\rangle |\phi_0\rangle + |1\rangle U |0^m\rangle e^{-iHt_1} |\phi_0\rangle \right). \tag{64}$$
5. Measure the second register in the standard basis. If the outcome is not 0^m, then this procedure fails; otherwise, continue. The probability of this step succeeding is
$$p_{succ} = \frac{1 + \alpha^{-2} \langle\phi_0| e^{iHt_1} O^2 e^{-iHt_1} |\phi_0\rangle}{2}, \tag{65}$$
and when this event happens, the state becomes
$$\frac{1}{\sqrt{2p_{succ}}} \left[|0\rangle |\phi_0\rangle + \alpha^{-1} |1\rangle O e^{-iHt_1} |\phi_0\rangle \right]. \tag{66}$$
6. Apply a $C\text{-}e^{-iHt_2}$ on the first and third registers. The state becomes
$$\frac{1}{\sqrt{2p_{succ}}} \left[|0\rangle |\phi_0\rangle + \alpha^{-1} |1\rangle e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle \right]. \tag{67}$$
7. Apply $W = I$ or phase gate S on the first register.
8. Apply a Hadamard gate on the first register.

9. Measure the first register in the standard basis. Then if $W = I$, the (conditional) probability of getting outcome 0 is

$$\mathbb{P}[0|succ] = \frac{p_{succ} + \alpha^{-1} \operatorname{Re}[\langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle]}{2p_{succ}}; \tag{68}$$

if $W = S$, this probability is

$$\mathbb{P}[0|succ] = \frac{p_{succ} - \alpha^{-1} \operatorname{Im}[\langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle]}{2p_{succ}}. \tag{69}$$

Now we define two random variables X and Y as follows. First, we run the above procedure with $W = I$ in step 7. If step 5 fails, $X = 0$; otherwise, if the measurement outcome is 0 or 1 in step 9, then $X = \alpha$ or $-\alpha$, respectively. One can show that X is an unbiased estimator of $\operatorname{Re}[\langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle]$, i.e.

$$\mathbb{E}[X] = \operatorname{Re}[\langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle]. \tag{70}$$

Y is defined similarly. We run the above procedure with $W = S$ in step 7. If step 5 fails, $Y = 0$; otherwise, if the measurement outcome is 1 or 0 in step 9, then $Y = \alpha$ or $-\alpha$, respectively. Then Y is an unbiased estimator of $\operatorname{Im}[\langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle]$, i.e.

$$\mathbb{E}[Y] = \operatorname{Im}[\langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle]. \tag{71}$$

It follows that $Z := X + iY$ is an unbiased estimator of $\langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle$, i.e.

$$\mathbb{E}[Z] = \langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle . \tag{72}$$

Note that $|Z|^2 = |X|^2 + |Y|^2 \leq 2\alpha^2$ with certainty.

Equipped with the above method for estimating $\langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle$ for arbitrary $t_1, t_2 \in \mathbb{R}$, we can now use the same strategy as in Lemma 5.3 to estimate $\widetilde{C_{O,2}}(x, y)$. The other components of Algorithm 3 remain intact. The analysis of this modified algorithm is almost the same as before, except that now we have

$$\mathbf{Var}[\overline{G_2}(x, y)] = \widetilde{O}(\alpha^2). \tag{73}$$

As a consequence, compared to Theorem 5.1, the total evolution time of this modified algorithm is larger by a factor of $O(\alpha^2)$, while its maximal evolution time is of the same order. □

7 Applications

In this section, we discuss some applications of our ground state property estimation algorithm.

7.1 Charge density

The primary application of the technique is the estimation of ground state properties of physical systems. Here we describe how to compute the charge density of a molecule, which can be used to compute properties like electric dipole moments of a molecule [RGM+21]. From a second-quantized representation of the electronic system (assuming fixed positions of the nuclear positions), the charge density is determined from the one-particle reduced density matrix as,

$$\rho(\vec{r}) = -e \sum_{p,q} D_{p,q} \phi_p(\vec{r}) \phi_q(\vec{r}), \tag{74}$$

where e is the electric constant, $D_{p,q}$ is the one-electron reduced density matrix (1RDM) of the ground state, and $\phi_q(\vec{r})$ are the basis wave functions chosen for the second-quantized representation of the electronic system [HJO14]. The 1RDM of the ground state is a matrix of properties of the ground state with each entry defined as

$$D_{p,q} = \langle\psi_0| a_p^\dagger a_q |\psi_0\rangle , \tag{75}$$

where a_p are annihilation operators. The operators involved in the 1RDM can each be expressed as a linear combination of unitary operators using the Majorana representation $a_p^\dagger = \frac{1}{2}(\gamma_{2p-1} + i\gamma_{2p})$, where the γ_k are hermitian and unitary[i],

$$D_{p,q} = \frac{1}{4}\left(\langle\psi_0| \gamma_{2p-1}\gamma_{2q-1} |\psi_0\rangle + i \langle\psi_0| \gamma_{2p-1}\gamma_{2q} |\psi_0\rangle - i \langle\psi_0| \gamma_{2p}\gamma_{2q-1} |\psi_0\rangle - \langle\psi_0| \gamma_{2p}\gamma_{2q} |\psi_0\rangle\right). \tag{76}$$

Accordingly, we may use the method of Section 5 to estimate each entry of the 1RDM and then obtain the charge density function of the ground state. As a point of comparison, we could alternatively use the variational quantum eigensolver algorithm to prepare an approximation to the ground state and then directly estimate each of the Pauli expectation values. However, there is no guarantee on whether a target accuracy for the ground state approximation can be achieved. Remarkably, the methods introduced in this paper can be used to ensure a target accuracy in the estimation regardless of the quality of ground state approximation, though possibly at the cost of an increase in runtime.

7.2 Quantum linear system solver

In the seminal [HHL09] paper, one of the motivations of solving linear systems using quantum computers is that in many cases, we only need to know $\langle x| M |x\rangle$, where $|x\rangle$ is the solution of a linear system $A|x\rangle = |b\rangle$, and M is a linear operator. For example, in quantum mechanics, many features of $|x\rangle$ can be extracted in this way, including normalization, moments, etc. One approach to solve this problem is first solving the linear system using any quantum linear system solver [HHL09, CKS17, CGJ18, GSLW19] to obtain the state $|x\rangle$ and then performing the measurement of M. However, a shortcoming of this method is that most of the quantum linear system solvers require deep quantum circuits. And hence, the needed quantum resources may not be accessible in the near future.

Recently, a few quantum algorithms [BPLC+19, HBR19, SSO19] were developed to solve linear systems of equations by encoding such a system into an effective Hamiltonian

$$H_G := A^\dagger(I - |b\rangle\langle b|)A, \tag{77}$$

whose ground state corresponds to the solution vector $|x\rangle$. We can combine this idea with our ground state property estimation algorithm to get a low-depth algorithm for estimating the properties of linear system solution. More specifically, suppose we can simulate the Hamiltonian H_G for some specified time and we know the normalization factor τ such that the eigenvalues of τH_G are in $[-\pi/3, \pi/3]$. For the operator M, we can assume that M can be decomposed into a linear combination of Pauli operators $M = \sum_{\ell=1}^{L} c_\ell \sigma_\ell$, or we assume that M is given in the block-encoding form. The estimation algorithm has two steps:

[i] To implement this application on a quantum computer we must represent the unitaries as operations on qubits. For an n-electron system, using the Jordan-Wigner or Bravyi-Kitaev transformation [SRL12], each Majorana operator, and products thereof, can be represented as a Pauli string.

1. Run a quantum linear system algorithm (e.g. [SSO19], [AL19], or [LT20b]) with constant precision to prepare an initial state $|\phi_0\rangle$ such that $|\langle\phi_0|x\rangle|^2$ is $\Omega(1)$.

2. Using $|\phi_0\rangle$ from step 1 as the initial state, run Algorithm 3 to estimate $\langle x|M|x\rangle$ within ϵ-additive error for any $\epsilon \in (0,1)$.

Step 1 takes $\widetilde{O}(\kappa)$ time, where κ is the condition number of A. To analyze the computation cost of the second step, we need a lower-bound on the spectral gap of H_G. Since $\langle x|A^\dagger(I-|b\rangle\langle b|)A|x\rangle = 0$, we have $\lambda_0(H_G) = 0$. For the second smallest eigenvalue, since $H_G = A^\dagger A - A^\dagger|b\rangle\langle b|A$, by Weyl's inequality, we have

$$\begin{aligned}\lambda_1(H_G) &\geq \lambda_0(A^\dagger A) - \lambda_1(A^\dagger|b\rangle\langle b|A)\\ &= \lambda_0(A^\dagger A),\end{aligned} \tag{78}$$

where the second step follows from $A^\dagger|b\rangle\langle b|A$ is rank-1. Due to the normalization, the smallest (normalized) singular value of A is $\Omega(\kappa^{-1})$. Hence, we have $\gamma = \Omega(\kappa^{-2})$.

By Theorem 5.1, the maximal evolution time of the Hamiltonian will be $\widetilde{O}(\kappa^2)$. To further improve the circuit depth, we may apply the gap amplification technique [SB13, SSO19] to quadratically increase the spectral gap of H_G. Consider the following Hamiltonian:

$$H'_G := \sigma^+ \otimes A^\dagger(I-|b\rangle\langle b|) + \sigma^- \otimes (I-|b\rangle\langle b|)A, \tag{79}$$

where $\sigma^\pm = (X \pm iY)/2$ are single-qubit (raising and lowering) operators. Then, we have

$$(H'_G)^2 = \begin{bmatrix} H_G & 0 \\ 0 & (I-|b\rangle\langle b|)AA^\dagger(I-|b\rangle\langle b|)\end{bmatrix}. \tag{80}$$

As shown by [SSO19], the eigenvalues of H'_G are

$$\left\{0, 0, \pm\sqrt{\lambda_1(H_G)}, \pm\sqrt{\lambda_2(H_G)}, \dots\right\}. \tag{81}$$

Hence, the spectral gap of H'_G will be $\Omega(\kappa^{-1})$. However, the smallest eigenvalue of H'_G has multiplicity 2, that is, the ground state space is spanned by $\{|0\rangle|x\rangle, |1\rangle|b\rangle\}$. Thus, for the initial state $|\phi\rangle$ outputted by VQE or QAOA, we should first project it into the space space $(|1\rangle|b\rangle)^\perp$. In this way, the initial state only has overlap with $|0\rangle|x\rangle$ and the remaining part of the ground state property estimation will work.

Then, we can run Algorithm 3 to estimate $\langle x|M|x\rangle$. Notice that since we know the ground state energy of H'_G is zero, we do not need to first estimate the ground state energy using Algorithm 1. Instead, we directly evaluate the O-weighted CDF at zero.

Therefore, by Theorem 6.1, we immediately get the following result:

Corollary 7.1 (Quantum linear system solution property estimation)**.** *For a linear system $A|x\rangle = |b\rangle$, suppose we can simulate the Hamiltonian H'_G (Eq. (79)) for some period of time. Furthermore, suppose A is Hermitian with eigenvalues in $[-1,-1/\kappa]\cup[1/\kappa,1]$ for $\kappa > 1$ and the eigenvalues of H'_G are in $[-\pi/3,\pi/3]$.*

Then, for any linear operator M given by its α-block encoding unitary U_M, and for any $\epsilon \in (0,1)$, the expectation value $\langle x|M|x\rangle$ can be estimated with ϵ-additive error with high probability such that:

- *the maximal evolution time of H'_G is $\widetilde{O}(\kappa)$.*
- *the expected total evolution time is $\widetilde{O}(\kappa\epsilon^{-2}\alpha^2)$.*

For comparison, the adiabatic algorithm in [SSO19] needs $\widetilde{O}(\kappa/\epsilon)$ evolution time of H'_G to obtain a state that is ϵ-close to $|x\rangle$, which is larger than our maximal evolution time. Moreover, to estimate $\langle x| M |x\rangle$, even using amplitude amplification and estimation, they still need $\Omega(\epsilon^{-1})$ copies of the state to achieve ϵ-additive error. Hence, their total evolution time will be $\widetilde{O}(\kappa\epsilon^{-2})$, nearly matching our result.

8 Discussion and Outlook

We have shown a quantum-classical hybrid algorithm for estimating properties of the ground state of a Hamiltonian, such that the quantum circuit depth is relatively small and only poly-logarithmically depends on ϵ^{-1}. Therefore, the algorithm has a significant advantage in high-accuracy estimation, and it is possible to be implemented in early fault-tolerant devices. In practice, our algorithm can solve many important tasks by combining with some initial state preparation methods (e.g., VQE or QAOA). In this paper, we provide two examples, one in quantum chemistry and another in solving linear systems. And we believe more applications will be explored in the future.

Another important direction is to improve the total evolution time of our algorithm which quadratically depends on ϵ^{-1}. The blowup comes from evaluating the O-weighted CDF in high precision and a trade-off between maximal evolution time and total evolution time. However, this does not meet the Heisenberg-limit of linear dependence on ϵ^{-1} for generic Hamiltonians [AA17]. In our main result (Theorem 5.1), the $\epsilon^{-2}\eta^{-2}$ factor comes from the number of samples needed to reduce the estimator's error to $O(\epsilon\eta)$. Amplitude estimation can be used to reduce this number of samples and the total evolution time. However, this comes at the cost of significantly increasing the maximal evolution time, which could require large fault-tolerant overheads for reliable implementation. A strategy to achieve improved performance that is more amenable to early fault tolerant quantum computers is to use recently introduced "enhanced sampling" techniques [WKJC21]. If λ characterizes the fidelity decay rate of the circuit as deeper circuits are used, then we would expect to need a maximal evolution time of $O(\lambda^{-1}\gamma^{-1})$ and an total evolution time of $O(\lambda\gamma^{-1}\epsilon^{-2}\eta^{-2})$. Note that because this approach incorporates the impact of error into the algorithm, the maximal evolution time is of no concern. Rather than being a cost that needs monitoring, the maximal evolution time is chosen by the algorithm to minimize the total evolution time. With this, we expect that as the quality of devices is improved, the performance of the algorithm improves proportionally. We note that a similar approach can also be applied to improve the total evolution time in [LT21] from $\widetilde{O}(\epsilon^{-1}\eta^{-2})$ to $\widetilde{O}(\lambda\epsilon^{-1}\eta^{-2})$.

This work fits into the paradigm of "beyond the ground state energy" and studies more general properties of the ground state. Can we go further beyond the ground state? Some prior works have explored the estimation of such kind of properties of Hamiltonian. For example, Brown, Flammia, and Schuch [BFS11] studied the density of states. Jordan, Gosset, and Love [JGL10] focused on the energy of excited states. Gharibian and Sikora [GS15] identified the energy barriers. In general, for an unknown Hamiltonian, these estimation problems will be hard. An interesting open problem is, given some prior knowledge of the Hamiltonian, can we design efficient or low-depth quantum algorithms for estimating Hamiltonian properties beyond ground state?

Acknowledgements We thank Yu Tong, Phillip Jensen, Max Radin, Jérôme Gonthier, Alex Kunitsa, and Aram Harrow for useful input and feedback on this work.

References

[AA17] Yosi Atia and Dorit Aharonov. Fast-forwarding of hamiltonians and exponentially precise measurements. *Nature communications*, 8(1):1–9, 2017.

[AGDLHG05] Alán Aspuru-Guzik, Anthony D Dutoi, Peter J Love, and Martin Head-Gordon. Simulated quantum computation of molecular energies. *Science*, 309(5741):1704–1707, 2005.

[AL19] Dong An and Lin Lin. Quantum linear system solver based on time-optimal adiabatic quantum computing and quantum approximate optimization algorithm. *arXiv preprint arXiv:1909.05500*, 2019.

[Amb14] Andris Ambainis. On physical problems that are slightly more difficult than qma. In *2014 IEEE 29th Conference on Computational Complexity (CCC)*, pages 32–43. IEEE, 2014.

[BFS11] Brielin Brown, Steven T Flammia, and Norbert Schuch. Computational difficulty of computing the density of states. *Physical review letters*, 107(4):040501, 2011.

[BPLC+19] Carlos Bravo-Prieto, Ryan LaRose, Marco Cerezo, Yigit Subasi, Lukasz Cincio, and Patrick J Coles. Variational quantum linear solver. *arXiv preprint arXiv:1909.05820*, 2019.

[CGJ18] Shantanav Chakraborty, András Gilyén, and Stacey Jeffery. The power of block-encoded matrix powers: improved regression techniques via faster hamiltonian simulation. *arXiv preprint arXiv:1804.01973*, 2018.

[CKS17] Andrew M Childs, Robin Kothari, and Rolando D Somma. Quantum algorithm for systems of linear equations with exponentially improved dependence on precision. *SIAM Journal on Computing*, 46(6):1920–1950, 2017.

[CRAG18] Yudong Cao, Jhonathan Romero, and Alán Aspuru-Guzik. Potential of quantum computing for drug discovery. *IBM Journal of Research and Development*, 62(6):6–1, 2018.

[CRO+19] Yudong Cao, Jonathan Romero, Jonathan P Olson, Matthias Degroote, Peter D Johnson, Mária Kieferová, Ian D Kivlichan, Tim Menke, Borja Peropadre, Nicolas PD Sawaya, et al. Quantum chemistry in the age of quantum computing. *Chemical reviews*, 119(19):10856–10915, 2019.

[GPY20] Sevag Gharibian, Stephen Piddock, and Justin Yirka. Oracle complexity classes and local measurements on physical hamiltonians. In *37th International Symposium on Theoretical Aspects of Computer Science (STACS 2020)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2020.

[GRB+20] Jérôme F Gonthier, Maxwell D Radin, Corneliu Buda, Eric J Doskocil, Clena M Abuan, and Jhonathan Romero. Identifying challenges towards practical quantum

advantage through resource estimation: the measurement roadblock in the variational quantum eigensolver. *arXiv preprint arXiv:2012.04001*, 2020.

[GS15] Sevag Gharibian and Jamie Sikora. Ground state connectivity of local hamiltonians. In *International Colloquium on Automata, Languages, and Programming*, pages 617–628. Springer, 2015.

[GSLW19] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 193–204, 2019.

[GTC19] Yimin Ge, Jordi Tura, and J Ignacio Cirac. Faster ground state preparation and high-precision ground energy estimation with fewer qubits. *Journal of Mathematical Physics*, 60(2):022202, 2019.

[GY19] Sevag Gharibian and Justin Yirka. The complexity of simulating local measurements on quantum systems. *Quantum*, 3:189, 2019.

[HBR19] Hsin-Yuan Huang, Kishor Bharti, and Patrick Rebentrost. Near-term quantum algorithms for linear systems of equations. *arXiv preprint arXiv:1909.07344*, 2019.

[HHL09] Aram W Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Physical review letters*, 103(15):150502, 2009.

[HJO14] Trygve Helgaker, Poul Jorgensen, and Jeppe Olsen. *Molecular electronic-structure theory*. John Wiley & Sons, 2014.

[Jen17] Frank Jensen. *Introduction to computational chemistry*. John wiley & sons, 2017.

[JGL10] Stephen P Jordan, David Gosset, and Peter J Love. Quantum-merlin-arthur–complete problems for stoquastic hamiltonians and markov matrices. *Physical Review A*, 81(3):032331, 2010.

[LT20a] Lin Lin and Yu Tong. Near-optimal ground state preparation. *Quantum*, 4:372, 2020.

[LT20b] Lin Lin and Yu Tong. Optimal polynomial based quantum eigenstate filtering with application to solving quantum linear systems. *Quantum*, 4:361, 2020.

[LT21] Lin Lin and Yu Tong. Heisenberg-limited ground state energy estimation for early fault-tolerant quantum computers. *arXiv preprint arXiv:2102.11340*, 2021.

[MMS+19] Sam McArdle, Alexander Mayorov, Xiao Shan, Simon Benjamin, and Xiao Yuan. Digital quantum simulation of molecular vibrations. *Chemical science*, 10(22):5725–5735, 2019.

[MRBAG16] Jarrod R McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18(2):023023, 2016.

[MW92] Yigal Meir and Ned S Wingreen. Landauer formula for the current through an interacting electron region. *Physical review letters*, 68(16):2512, 1992.

[OSS+19] Thomas E O'Brien, Bruno Senjean, Ramiro Sagastizabal, Xavier Bonet-Monroig, Alicja Dutkiewicz, Francesco Buda, Leonardo DiCarlo, and Lucas Visscher. Calculating energy derivatives for quantum chemistry on a quantum computer. *npj Quantum Information*, 5(1):1–12, 2019.

[PMS+14a] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O'brien. A variational eigenvalue solver on a photonic quantum processor. *Nature communications*, 5(1):1–7, 2014.

[PMS+14b] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O'brien. A variational eigenvalue solver on a photonic quantum processor. *Nature communications*, 5(1):1–7, 2014.

[RGM+21] Julia E Rice, Tanvi P Gujarati, Mario Motta, Tyler Y Takeshita, Eunseok Lee, Joseph A Latone, and Jeannette M Garcia. Quantum computation of dominant products in lithium–sulfur batteries. *The Journal of Chemical Physics*, 154(13):134115, 2021.

[SB13] Rolando D Somma and Sergio Boixo. Spectral gap amplification. *SIAM Journal on Computing*, 42(2):593–610, 2013.

[SO12] Attila Szabo and Neil S Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Courier Corporation, 2012.

[Som19] Rolando D Somma. Quantum eigenvalue estimation via time series analysis. *New Journal of Physics*, 21(12):123025, 2019.

[SRL12] Jacob T Seeley, Martin J Richard, and Peter J Love. The bravyi-kitaev transformation for quantum computation of electronic structure. *The Journal of chemical physics*, 137(22):224109, 2012.

[SSO19] Yiğit Subaşı, Rolando D Somma, and Davide Orsucci. Quantum algorithms for systems of linear equations inspired by adiabatic quantum computing. *Physical review letters*, 122(6):060504, 2019.

[WBG20] James D Watson, Johannes Bausch, and Sevag Gharibian. The complexity of translationally invariant problems beyond ground state energies. *arXiv preprint arXiv:2012.12717*, 2020.

[WKJC21] Guoming Wang, Dax Enshan Koh, Peter D Johnson, and Yudong Cao. Minimizing estimation runtime on noisy quantum computers. *PRX Quantum*, 2(1):010346, 2021.

A Ground State Energy Estimation

In this section, we review the techniques in [LT20a], which proposed a hybrid quantum/classical algorithm for estimating the ground state energy of a Hamiltonian. Compared with the algorithms in previous works, the algorithm in [LT21] uses fewer quantum resources and does not need to access the block-encoding of the Hamiltonian.

First of all, they assumed that the given initial state $|\phi_0\rangle$[ii] has a nontrivial overlap with the ground state of H.

[ii] In [LT21], they allowed the initial state to be a mixed state. For simplicity, we still denote it as $|\phi_0\rangle$.

A.1 Quantum part of the algorithm

Fix $j \in \mathbb{Z}$. Suppose we want to estimate $\Re(\langle\phi_0| e^{-ij\tau H} |\phi_0\rangle)$. Then, we set $\mathrm{W} = I$ and define a random variable X_j as follows:

$$X_j := \begin{cases} 1 & \text{if the outcome is } 0 \\ -1 & \text{if the outcome is } 1 \end{cases}.$$

Since the state before the measurement is

$$\frac{1}{2}(|0\rangle \otimes (I + e^{-ij\tau H}) |\phi_0\rangle + |1\rangle \otimes (I - e^{-ij\tau H}) |\phi_0\rangle), \tag{82}$$

we have

$$\begin{aligned} \mathbb{E}[X_j] =& \Pr[X_j = 0] - \Pr[X_j = 1] \\ =& \frac{1}{4} \langle\phi_0| (I + e^{ij\tau H})(I + e^{-ij\tau H}) |\phi_0\rangle - \frac{1}{4} \langle\phi_0| (I - e^{ij\tau H})(I - e^{-ij\tau H}) |\phi_0\rangle \\ =& \frac{1}{2} \langle\phi_0| (e^{ij\tau H} + e^{-ij\tau H}) |\phi_0\rangle \\ =& \Re(\langle\phi_0| e^{-ij\tau H} |\phi_0\rangle). \end{aligned} \tag{83}$$

For the imaginary part $\Im(\langle\phi_0| e^{-ij\tau H} |\phi_0\rangle)$, we can set W to be the phase gate $\begin{bmatrix} 1 & 0 \\ 0 & -i \end{bmatrix}$ and define the random variable Y_j similarly. Then, we have

$$\mathbb{E}[Y_j] = \Im(\langle\phi_0| e^{-ij\tau H} |\phi_0\rangle). \tag{84}$$

Therefore, Eqs. (83) and (84) implies the following claim:

Claim A.1 (Estimator of the Hamiltonian expectation)**.** *For any $j \in \mathbb{Z}$, the random variable $X_j + iY_j$ is an un-biased estimator for $\langle\phi_0| e^{-ij\tau H} |\phi_0\rangle$.*

A.2 Classical part of the algorithm

Let τ be a normalization factor such that $\|\tau H\| \leq \pi/3$. Suppose the initial state $|\phi_0\rangle$ can be decomposed in the eigenspace of H as $|\phi_0\rangle = \sum_k \sqrt{p_k} |\psi_k\rangle$. Let $p(x)$ be the following density function (spectral measure):

$$p(x) := \sum_k p_k \delta(x - \tau\lambda_k) \quad \forall x \in [-\pi, \pi]. \tag{85}$$

That is, $p(x)$ is the distribution of the state energy with respect to τH after we measure $|\phi_0\rangle$ in the eigenbasis of H.

Define the 2π-periodic Heaviside function by

$$H(x) = \begin{cases} 1 & x \in [2k\pi, (2k+1)\pi) \\ 0 & x \in [(2k-1)\pi, 2k\pi) \end{cases} \quad \forall k \in \mathbb{Z}. \tag{86}$$

Then, we define the 2π-periodic CDF of p as the convolution of H and p:

$$C(x) := (H * p)(x). \tag{87}$$

For any $x \in [-\pi/3, \pi/3]$, for any $w \in \mathbb{Z}$, we have

$$\begin{aligned} C(x + 2w\pi) &= \int_{-\pi}^{\pi} H(x + 2w\pi - t)p(t)\mathrm{d}t \qquad (88) \\ &= \sum_k p_k \cdot \int_{-\pi}^{\pi} H(x + 2w\pi - t)\delta(t - \tau\lambda_k)\mathrm{d}t \\ &= \sum_k p_k \cdot H(x + 2w\pi - \tau\lambda_k) \\ &= \sum_k p_k \cdot \mathbf{1}_{x \geq \tau\lambda_k} \\ &= \sum_{k:\lambda_k \leq x} p_k, \qquad (89) \end{aligned}$$

where the first step follows from the definition of convolution, the second step follows from Dirac delta function's property, and the third step follows from H has period 2π. We note that $C(x)$ is right continuous and non-decreasing in $[-\pi/3, \pi/3]$.

However, we cannot directly evaluate $C(x)$, but we can approximate it! Define the approximate CDF (ACDF) as

$$\widetilde{C}(x) := (F * p)(x), \qquad (90)$$

where $F(x) = \sum_{|j| \leq d} \widehat{F}_j e^{ijx}$ is a low Fourier-degree approximation of the Heaviside function $H(x)$ such that

$$|F(x) - H(x)| \leq \epsilon \quad \forall x \in [-\pi + \delta, -\delta] \cup [\delta, \pi - \delta]. \qquad (91)$$

The construction of F is given by Lemma A.8. Furthermore, the approximation error of $\widetilde{C}(x)$ is bounded by

$$C(x - \delta) - \epsilon \leq \widetilde{C}(x) \leq C(x + \delta) + \epsilon, \qquad (92)$$

for any $x \in [-\pi/3, \pi/3]$, $\delta \in (0, \pi/6)$ and $\epsilon > 0$.

A.2.1 Estimating the ACDF

The goal of this section is to prove Lemma A.2, which constructs an estimator for $\widetilde{C}(x)$ (defined by Eq. (90)).

Lemma A.2 (Estimating the ACDF)**.** *For any $\sigma > 0$, for any $x \in [-\pi, \pi]$, there exists an un-biased estimator $\overline{G}(x)$ for the ACDF $\widetilde{C}(x)$ with variance at most σ^2.*

Furthermore, $\overline{G}(x)$ runs the quantum circuit (Figure 1) $O(\frac{\log^2 d}{\sigma^2})$ times with expected total evolution time $O(\frac{\tau d \log d}{\sigma^2})$.

Proof. $\widetilde{C}(x)$ can be expanded in the following way:

$$\begin{aligned}
\widetilde{C}(x) &= (F * p)(x) \qquad (93)\\
&= \int_{-\pi}^{\pi} F(x-y)p(y)\mathrm{d}y \\
&= \sum_{|j|\le d} \int_{-\pi}^{\pi} \widehat{F}_j e^{ij(x-y)} p(y)\mathrm{d}y \\
&= \sum_{|j|\le d} \widehat{F}_j e^{ijx} \int_{-\pi}^{\pi} p(y) e^{-ijy}\mathrm{d}y \\
&= \sum_{|j|\le d} \widehat{F}_j e^{ijx} \sum_k p_k e^{-ij\tau\lambda_k} \\
&= \sum_{|j|\le d} \widehat{F}_j e^{ijx} \cdot \langle\phi_0| e^{-ij\tau H} |\phi_0\rangle , \qquad (94)
\end{aligned}$$

where the third step follows from the Fourier expansion of $F(x-y)$, the fifth step follows from the property of Dirac's delta function, and the last step follows from the definition of p_k and the eigenvalues of matrix exponential.

To estimate $\langle\phi_0| e^{-ij\tau H} |\phi_0\rangle$, we use the multi-level Monte Carlo method. Define a random variable J with support $\{-d, \cdots, d\}$ such that

$$\Pr[J = j] = \left|\widehat{F}_j\right| / \mathcal{F}, \qquad (95)$$

where $\mathcal{F} := \sum_{|j|\le d} |\widehat{F}_j|$. Then, let $Z := X_J + iY_J \in \{\pm 1 \pm i\}$. Define an estimator $G(x; J, Z)$ as follows:

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

where θ_j is defined by $\widehat{F}_j = |\widehat{F}_j| e^{i\theta_j}$. Then, we show that $G(x; J, Z)$ is un-biased:

$$\begin{aligned}
\mathbb{E}[G(x; J, Z)] &= \sum_{|j|\le d} \mathbb{E}\left[(X_j + iY_j) e^{i(\theta_j + jx)} |\widehat{F}_j|\right] \\
&= \sum_{|j|\le d} \widehat{F}_j e^{ijx} \cdot \mathbb{E}\left[X_j + iY_j\right] \\
&= \sum_{|j|\le d} \widehat{F}_j e^{ijx} \cdot \langle\phi_0| e^{-ij\tau H} |\phi_0\rangle \\
&= \widetilde{C}(x),
\end{aligned}$$

where the third step follows from Claim A.1. Moreover, the variance of G can be upper-bounded by:

$$\begin{aligned}
\mathrm{Var}[G(x; J, Z)] &= \mathbb{E}[|G(x; J, Z)|^2] - |\,\mathbb{E}[G(x; J, Z)]|^2 \\
&\le \mathbb{E}[|G(x; J, Z)|^2] \\
&= \mathcal{F}^2 \cdot \mathbb{E}[|X_J + iY_J|^2] \\
&= 2\mathcal{F}^2,
\end{aligned}$$

where the third step follows from $|e^{i(\theta_J+Jx)}|=1$, and the last step follows from $X_j, Y_j \in \{\pm 1\}$.

Hence, we can take $N_s := \frac{2\mathcal{F}^2}{\sigma^2}$ independent samples of (J,Z), denoted by $\{(J_k,Z_k)\}_{k\in[N_s]}$ and compute

$$\overline{G}(x) := \frac{1}{N_s}\sum_{k=1}^{N_s} G(x;J_k,Z_k).$$

Then, we have

$$\mathbb{E}[\overline{G}(x)] = \widetilde{C}(x), \quad \text{and} \quad \mathrm{Var}[\overline{G}(x)] \le \sigma^2.$$

The expected total evolution time is

$$\mathcal{T}_{\mathsf{tot}} := N_s\tau\,\mathbb{E}[|J|] = \frac{2\mathcal{F}^2}{\sigma^2}\tau\sum_{|j|\le d}|j|\cdot\frac{|\widehat{F}_j|}{\mathcal{F}} = \frac{2\mathcal{F}\tau}{\sigma^2}\sum_{|j|\le d}|j||\widehat{F}_j|.$$

By Lemma A.8, we know that $|\widehat{F}_j| = O(1/|j|)$. Hence, we have $\mathcal{F} = \sum_{|j|\le d} O(1/|j|) = O(\log d)$. Thus, the number of samples is

$$N_s = O\left(\frac{\log^2 d}{\sigma^2}\right).$$

And the expected total evolution time is

$$\mathcal{T}_{\mathsf{tot}} = O\left(\frac{\tau d\log d}{\sigma^2}\right).$$

The lemma is then proved. □

A.2.2 Inverting the CDF

We first define the CDF inversion problem:

Definition A.3 (The CDF inversion problem)**.** *For $0<\delta<\pi/6$, $0<\eta<1$, find $x^\star \in (-\pi/3,\pi/3)$ such that*

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

Remark A.4. *The condition in Definition A.3 is weaker than $\eta/2 < C(x) < \eta$ due to the discontinuity of $C(x)$. For any CDF $C(x)$, such an $x^\star$ must exist: let $a := \sup\{x\in(-\pi/3,\pi/3) : C(x)\le\eta/2\}$ and $b := \inf\{x\in(-\pi/3,\pi/3) : C(x)\ge\eta\}$. Since $C(x)$ is non-decreasing, we have $a\le b$. And any $x\in(a-\delta,b+\delta)$ satisfies the condition in Definition A.3.*

Then, we give an algorithm that solves the CDF inversion problem.

Lemma A.5 (Inverting the CDF, Theorem 2 in [LT21])**.** *There exists an algorithm that solves the CDF inversion problem (Definition A.3) with probability at least $1-\nu$ such that:*

1. *the number of independent samples of (J,Z) is*

$$O\left(\eta^{-2}\cdot(\log(\nu^{-1})+\log\log(\delta^{-1}))\cdot(\log(\delta^{-1})+\log\log(\delta^{-1}\eta^{-1}))^2\right)$$

Algorithm 4 Inverting the CDF

```
1: procedure InvertCDF(η, δ, {J_k, Z_k})
2:     x_L ← −π/3, X_R ← π/3
3:     while x_R − x_L > 2δ do
4:         x_M ← (x_L + x_R)/2
5:         u ← Certify(x_M, (2/3)δ, η, {J_k, Z_k})
6:         if u = 0 then
7:             x_R ← x_M + (2/3)δ
8:         else
9:             x_L ← x_M − (2/3)δ
10:        end if
11:    end while
12:    return (x_L + x_R)/2
13: end procedure
```

2. *the expected total evolution time is*

$$O\left(\tau\eta^{-2}\cdot\delta^{-1}\log(\delta^{-1}\eta^{-1})\cdot(\log(\delta^{-1})+\log\log(\delta^{-1}\eta^{-1}))\cdot(\log(\nu^{-1})+\log\log(\delta^{-1}))\right)$$

3. *the maximal evolution time is*

$$O\left(\tau\delta^{-1}\log(\delta^{-1}\eta^{-1})\right)$$

4. *the classical running time is*

$$O\left(\eta^{-2}\log(\delta^{-1})\cdot(\log(\nu^{-1})+\log\log(\delta^{-1}))\cdot(\log(\delta^{-1})+\log\log(\delta^{-1}\eta^{-1}))^2\right).$$

Proof. For any $x\in[-\pi/3,\pi/3]$, at least one of the following conditions will hold:

$$C(x+\delta)>\eta/2,\quad\text{or}\quad C(x-\delta)<\eta. \tag{96}$$

Suppose we have a sub-routine $\textsc{Certify}(x,\delta,\eta,\{J_k,Z_k\})$ such that if $C(x+\delta)>\eta/2$, it returns 0; otherwise, it returns 1.

Then, we can solve the CDF inversion problem via the binary search (Algorithm 4).

In Line 3, x_L and x_R always satisfy the following conditions:

$$C(x_L)<\eta,\quad C(x_R)>\eta/2,$$

which is guaranteed by $\textsc{Certify}(x_M,(2/3)\delta,\eta,\{J_k,Z_k\})$. Then, when the while-loop ends, we have $x_R-x_L\le 2\delta$. Let $x^\star:=(x_L+x_R)/2$ be the output of Algorithm 4. Then, we get that

$$\begin{aligned}C(x^\star+\delta)&\ge C(x_R)>\eta/2,\\ C(x^\star-\delta)&\le C(x_L)<\eta.\end{aligned}$$

And it is easy to see that Algorithm 4 will call $\textsc{Certify}$ $L:=O(\log(1/\delta))$ times. Then, by Lemma A.6 and union bound, Algorithm 4 will be correct with probability at least $1-\nu$. We note that different runs of $\textsc{Certify}$ can share a same set of samples $\{J_k,Z_k\}$, which does not affect the union bound. Hence, the number of samples and the total evolution time follows directly from Lemma A.6 and $d=O(\delta^{-1}\log(\delta^{-1}\eta^{-1}))$. □

Algorithm 5 Distinguish the two cases in Eq. (96)

procedure CERTIFY($x, \eta, \delta, \{J_k, Z_k\}$)
  $c \leftarrow 0$, $N_b \leftarrow \Omega(\log(1/\nu) + \log\log(1/\delta))$
  for $1 \le r \le N_b$ **do**
    Compute $\overline{G}(x)$ using $\{J_k, Z_k\}_{k\in[(r-1)N_s+1, rN_s]}$ ▷ Lemma A.2
    if $\overline{G}(x) \ge (3/4)\eta$ **then**
      $c \leftarrow c + 1$
    end if
  end for
  return $\mathbf{1}_{c\le N_b/2}$
end procedure

Lemma A.6 (CERTIFY sub-routine)**.** *For any $\nu > 0$, there exists an algorithm that distinguishes the two cases in Eq. (96) for any $x \in [-\pi/3, \pi/3]$ with probability at least $1 - O(\nu/L)$ using*

$$O\left(\eta^{-2}\log^2(d)(\log(1/\nu) + \log\log(1/\delta))\right)$$

independent samples of (J, Z), and total evolution time

$$O\left(\eta^{-2}\tau d\log(d)(\log(1/\nu) + \log\log(1/\delta))\right)$$

in expectation.

Proof. To decide which one of the conditions holds for x, we can estimate the ACDF $\widetilde{C}(x)$. If we take $\epsilon = \eta/8$ in Lemma A.8, then the constructed ACDF satisfies

$$C(x-\delta) - \eta/8 \le \widetilde{C}(x) \le C(x+\delta) + \eta/8.$$

Thus,

$$\begin{aligned}\widetilde{C}(x) > (5/8)\eta \quad &\Rightarrow \quad C(x+\delta) > \eta/2,\\ \widetilde{C}(x) < (7/8)\eta \quad &\Rightarrow \quad C(x-\delta) < \eta.\end{aligned}$$

Then, we can distinguish $\widetilde{C}(x) > (5/8)\eta$ or $\widetilde{C}(x) < (7/8)\eta$ by the estimator in Lemma A.2.

In Algorithm 5, we compute the estimator $\overline{G}(x)$ N_b times independently, where each time we use N_s samples of (J, Z). We note that an error occurs when $\widetilde{C}(x) > (7/8)\eta$ but $\overline{G}(x) < (3/4)\eta$, or $\widetilde{C}(x) < (5/8)\eta$ but $\overline{G}(x) > (3/4)\eta$ (when $(5/8)\eta \le \widetilde{C}(x) \le (7/8)\eta$, any output is correct). By Chebyshev's inequality, we have

$$\begin{aligned}\Pr[\overline{G}(x) \text{ has an error}] &\le \Pr\left[\overline{G}(x) < \frac{3}{4}\eta \;\middle|\; \widetilde{C}(x) > \frac{7}{8}\eta\right] + \Pr\left[\overline{G}(x) > \frac{3}{4}\eta \;\middle|\; \widetilde{C}(x) < \frac{5}{8}\eta\right]\\ &\le 2\cdot\frac{\sigma^2}{\eta^2/64}\\ &\le \frac{1}{4},\end{aligned}$$

if we take $\sigma^2 = O(\eta^2)$ in Lemma A.2.

Then, by the Chernoff bound, we have

$$\Pr[\text{CERTIFY makes an error}] \le \exp(-\Omega(N_b)) \le \nu/L,$$

if we take $N_b := \Omega(\log(L/\nu)) = \Omega(\log(1/\nu) + \log\log(1/\delta))$. Thus, the total number of samples is

$$N_b N_s = O\left(\eta^{-2}\log^2(d)(\log(1/\nu) + \log\log(1/\delta))\right),$$

and the expected total evolution time is

$$O\left(\eta^{-2}\tau d\log(d)(\log(1/\nu) + \log\log(1/\delta))\right),$$

which complete the proof of the lemma. □

A.2.3 Estimating the ground state energy

Corollary A.7 (Ground state energy estimation, Corollary 3 in [LT21]). *If $p_0 \geq \eta$ for some known η, then with probability at least $1-\nu$, the ground state energy λ_0 can be estimated within additive error ϵ, such that:*

1. *the number of times running the quantum circuit (Figure 1) is $\widetilde{O}(\eta^{-2})$.*
2. *the expected total evolution time is $\widetilde{O}(\epsilon^{-1}\eta^{-2})$.*
3. *the maximal evolution time is $\widetilde{O}(\epsilon^{-1})$.*
4. *the classical running time is $\widetilde{O}(\eta^{-2})$.*

Proof. Suppose we can solve the CDF inversion problem (Definition A.3) for $\delta = \tau\epsilon$ and η, i.e., we find an $x^\star$ such that

$$C(x^\star + \tau\epsilon) > \eta/2 > 0, \quad C(x^\star - \tau\epsilon) < \eta \leq p_0.$$

Since $C(x)$ cannot take value between 0 and p_0, we have

$$x^\star + \tau\epsilon \geq \tau\lambda_0, \quad x^\star - \tau\epsilon < \tau\lambda_0,$$

which is

$$|x^\star/\tau - \lambda_0| \leq \epsilon.$$

The costs of this algorithm follows from Lemma A.5. □

A.3 Low Fourier degree approximation of the Heaviside function

We construct the low degree approximation of the Heaviside function in this section.[iii]

Lemma A.8 (Constructing low degree approximation of H). *Let $H(x)$ be the 2π-period Heaviside function (Eq. (86)). For any $\delta \in (0, \pi/2)$ such that $\tan(\delta/2) \leq 1 - 1/\sqrt{2}$, there exists a $d = O(\delta^{-1}\log(\delta^{-1}\epsilon^{-1}))$ and a 2π-period function $F_{d,\delta}(x)$ of the form:*

$$F_{d,\delta}(x) = \frac{1}{\sqrt{2\pi}} \sum_{j=-d}^{d} \widehat{F}_{d,\delta,j} \cdot e^{ijx} \tag{97}$$

such that

[iii] The construction in [LT21] is not enough to prove Lemma A.9 because the range of $F_{d,\delta}$ is $[-\epsilon/2, 1+\epsilon]$ while Lemma A.9 requires the range to be $[0,1]$. We fix this issue in Lemma A.8.

1. $F_{d,\delta}(x) \in [0,1]$ *for all* $x \in \mathbb{R}$.
2. $|F_{d,\delta}(x) - H(x)| \le \epsilon$ *for* $x \in [-\pi+\delta, -\delta] \cup [\delta, \pi-\delta]$.
3. $|\widehat{F}_{d,\delta,j}| = \Theta(1/|j|)$ *for* $j \neq 0$.

Proof. We first construct $F'_{d,\delta}(x)$ by mollifying the Heaviside function with $M_{d,\delta}(x)$ in Lemma A.10:

$$F'_{d,\delta}(x) := (M_{d,\delta} * H)(x) = \int_{-\pi}^{\pi} M_{d,\delta}(y) H(x-y) \mathrm{d}y. \tag{98}$$

We can verify that $F'_{d,\delta}$ has Fourier degree at most d. It follows from the Chebyshev polynomial $T_d(x)$ is of degree d. Hence, the Fourier coefficients of $M_{d,\delta}(x)$:

$$\widehat{M}_{d,\delta,j} = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} M_{d,\delta}(x) e^{-ijx} \mathrm{d}x \neq 0 \tag{99}$$

only if $j \in \{-d, \ldots, d\}$. Since $F_{d,\delta}$ is a convolution of $M_{d,\delta}$ and H, we have

$$\widehat{F'}_{d,\delta,j} = \sqrt{2\pi} \widehat{M}_{d,\delta,j} \widehat{H}_j \quad \forall |j| \le d. \tag{100}$$

Then, we define

$$F_{d,\delta}(x) := \frac{1}{\sqrt{2\pi}} \sum_{j=-d}^{d} \widehat{F}_{d,\delta,j} \cdot e^{ijx}, \tag{101}$$

where

$$\widehat{F}_{d,\delta,j} = \begin{cases} \frac{1}{1+(5/4)\epsilon} \left(\widehat{F'}_{d,\delta,j} + \sqrt{2\pi}\epsilon/4 \right) & \text{if } j = 0, \\ \frac{1}{1+(5/4)\epsilon} \widehat{F'}_{d,\delta,j} & \text{otherwise.} \end{cases} \tag{102}$$

It is easy see that

$$F_{d,\delta}(x) = \frac{F'_{d,\delta}(x) + \epsilon/4}{1 + (5/4)\epsilon} \quad \forall x \in \mathbb{R}. \tag{103}$$

Then, we will show that taking $d = O(\delta^{-1} \log(\delta^{-1}\epsilon^{-1}))$ is enough to satisfy (1)-(3).

Part (1): We first compute the range of $F'_{d,\delta}(x)$:

$$F'_{d,\delta}(x) \le \int_{-\pi}^{\pi} |M_{d,\delta}(y)| \mathrm{d}y \le 1 + \frac{4\pi}{\mathcal{N}_{d,\delta}}, \tag{104}$$

where the second step follows from (2) in Lemma A.10. On the other hand,

$$F'_{d,\delta}(x) \ge -\frac{1}{\mathcal{N}_{d,\delta}} \int_{-\pi}^{\pi} H(y) \mathrm{d}y = \frac{-\pi}{\mathcal{N}_{d,\delta}}.$$

Hence, if we take $d = O(\delta^{-1} \log(\delta^{-1}\epsilon^{-1}))$ such that

$$\mathcal{N}_{d,\delta} \ge C_1 e^{d\delta/\sqrt{2}} \sqrt{\frac{\delta}{d}} \cdot \mathrm{erf}(C_2 \sqrt{d\delta}) \ge \frac{4\pi}{\epsilon} \tag{105}$$

holds, we will have

$$-\epsilon/4 \le F'_{d,\delta} \le 1 + \epsilon. \tag{106}$$

Therefore, for all $x \in \mathbb{R}$,

$$F_{d,\delta}(x) = \frac{F'_{d,\delta}(x) + \epsilon/4}{1 + (5/4)\epsilon} \in [0,1]. \tag{107}$$

Part (2): The approximation error of $F'_{d,\delta}$ is

$$|F'_{d,\delta}(x) - H(x)| \le \left| \int_{-\pi}^{\pi} M_{d,\delta}(y)(H(x-y) - H(x))\mathrm{d}y \right| \le \int_{-\pi}^{\pi} |M_{d,\delta}(y)||H(x-y) - H(x)|\mathrm{d}y, \tag{108}$$

where the first step follows from (2) in Lemma A.10, and the second step follows from the triangle inequality.

Fix $x \in [-\pi+\delta, -\delta] \cup [\delta, \pi-\delta]$. If $y \in (-\delta, \delta)$, then $H(x-y) = H(x)$ and

$$\int_{-\delta}^{\delta} |M_{d,\delta}(y)||H(x-y) - H(x)|\mathrm{d}y = 0. \tag{109}$$

If $|y| \ge \delta$, by (1) in Lemma A.10, we have $|M_{d,\delta}| \le \frac{1}{\mathcal{N}_{d,\delta}}$. Since $|H(x-y) - H(x)| \le 1$, we have

$$\left(\int_{-\pi}^{-\delta} + \int_{\delta}^{\pi} \right) |M_{d,\delta}(y)||H(x-y) - H(x)|\mathrm{d}y \le \frac{2\pi}{\mathcal{N}_{d,\delta}} \le \epsilon/2, \tag{110}$$

where the last step follows from Eq. (105). Therefore,

$$|F'_{d,\delta}(x) - H(x)| \le \epsilon/2 \quad \forall |x| \in [\delta, \pi-\delta]. \tag{111}$$

Thus,

$$|F_{d,\delta}(x) - H(x)| = \left| \frac{F'_{d,\delta}(x) + \epsilon/4}{1 + (5/4)\epsilon} - H(x) \right| \tag{112}$$

$$\le |F'_{d,\delta}(x) - H(x)| + \frac{(5/4)\epsilon}{1+(5/4)\epsilon}|F'_{d,\delta}(x)| + \frac{\epsilon/4}{1+(5/4)\epsilon}$$

$$\le \epsilon/2 + \frac{(5/4)\epsilon}{1+(5/4)\epsilon}(1+\epsilon) + \frac{\epsilon/4}{1+(5/4)\epsilon}$$

$$\le 2\epsilon, \tag{113}$$

where the second step follows from the triangle inequality, the third step follows from Eq. (106). By scaling for ϵ, we can make the approximation error at most ϵ.

Part (3): Since $|\widehat{F'}_{d,\delta,j}| = \sqrt{2\pi}|\widehat{M}_{d,\delta,j}||\widehat{H}_j|$, we first bound $|\widehat{M}_{d,\delta,j}|$:

$$\left|\widehat{M}_{d,\delta,j}\right| \le \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} |M_{d,\delta}(x)|\mathrm{d}x \le \frac{1}{\sqrt{2\pi}} \left(1 + \frac{4\pi}{\mathcal{N}_{d,\delta}}\right) \le \frac{1+\epsilon}{\sqrt{2\pi}}, \tag{114}$$

where the second step follows from (2) in Lemma A.10 and the last step follows from Eq. (105).

For $|\widehat{H}_j|$, if $j \neq 0$, we have

$$\widehat{H}_j = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} H(x) e^{-ijx}\mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_{0}^{\pi} e^{-ijx}\mathrm{d}x = \begin{cases} \frac{\sqrt{2}}{i\sqrt{\pi}j} & \text{if } j \text{ is odd,} \\ 0 & \text{if } j \text{ is even.} \end{cases} \tag{115}$$

Hence, for $j \neq 0$,

$$|\widehat{F'}_{d,\delta,j}| \le \sqrt{2\pi} \cdot \frac{1+\epsilon}{\sqrt{2\pi}} \cdot \sqrt{\frac{2}{\pi}} \frac{1}{|j|} = \frac{1+\epsilon}{\sqrt{\pi/2}|j|}. \tag{116}$$

Then, by definition, we get that

$$|\widehat{F}_{d,\delta,j}| \leq \frac{1+\epsilon}{\sqrt{\pi/2}(1+(5/4)\epsilon)|j|} = \Theta(1/|j|). \tag{117}$$

The proof of the lemma is completed. □

The following lemma shows the approximation ratio of the ACDF $\widetilde{C}(x)$ constructed from the low degree approximated Heaviside function $F(x)$ by Lemma A.8.

Lemma A.9 (Approximation ratio of the ACDF)**.** *For any $\epsilon > 0$, $0 < \delta < \pi/6$, let $F(x) := F_{d,\delta}(x)$ constructed by Lemma A.8. Then, for any $x \in [-\pi/3, \pi/3]$, the ACDF $\widetilde{C}(x) = (F * p)(x)$ satisfies:*

$$C(x-\delta) - \epsilon \leq \widetilde{C}(x) \leq C(x+\delta) + \epsilon.$$

Proof. By (2) in Lemma A.8, we have

$$|F(x) - H(x)| \leq \epsilon \quad \forall x \in [-\pi+\delta, -\delta] \cup [\delta, \pi-\delta]. \tag{118}$$

Define $F_L := F(x-\delta)$ such that

$$|F_L(x) - H(x)| \leq \epsilon \quad \forall x \in [-\pi+2\delta, 0] \cup [2\delta, \pi]. \tag{119}$$

For $\widetilde{C}_L(x) := (F_L * p)(x)$, we have $\widetilde{C}_L(x) = \widetilde{C}(x-\delta)$, and for $x \in [-\pi/3, \pi/3]$,

$$\begin{aligned}
|C(x) - \widetilde{C}_L(x)| &= \left| \int_{-\pi}^{\pi} p(x-y)(H(y) - F_L(y)) \mathrm{d}y \right| \qquad (120)\\
&\leq \int_{-\pi}^{\pi} p(x-y)|H(y) - F_L(y)| \mathrm{d}y \\
&= \left(\int_{-\pi}^{0} + \int_{2\delta}^{\pi} \right) p(x-y)|H(y) - F_L(y)| \mathrm{d}y + \int_{0}^{2\delta} p(x-y)|H(y) - F_L(y)| \mathrm{d}y \\
&\leq \epsilon \cdot \left(\int_{-\pi}^{0} + \int_{2\delta}^{\pi} \right) p(x-y) \mathrm{d}y + \int_{0}^{2\delta} p(x-y)|H(y) - F_L(y)| \mathrm{d}y \\
&\leq \epsilon + \int_{0}^{2\delta} p(x-y)|H(y) - F_L(y)| \mathrm{d}y \\
&\leq \epsilon + \int_{0}^{2\delta} p(x-y) \mathrm{d}y \\
&= \epsilon + \int_{x-2\delta}^{x} p(y) \mathrm{d}y \\
&= \epsilon + C(x) - C(x-2\delta), \qquad (121)
\end{aligned}$$

where the second step follows from Cauchy-Schwarz inequality, the forth step follows from Eq. (119), the fifth step follows from $p(x)$ is a density function, the sixth step follows from $H(y) = 1$ and $F_L(y) \in [0,1]$ for $y \in [0, 2\delta]$, the last step follows from $C(x)$ is the CDF of $p(x)$ in $[-\pi, \pi]$.

Hence, we have

$$\widetilde{C}_L(x) \geq C(x) - (\epsilon + C(x) - C(x-2\delta)) = C(x-2\delta) - \epsilon, \tag{122}$$

which proves the first inequality:

$$\widetilde{C}(x-\delta) \geq C(x-2\delta) - \epsilon. \tag{123}$$

Similarly, we can define $F_R := F(x+\delta)$ and $\widetilde{C}_R(x) := (F_R * p)(x)$. We can show that

$$|C(x) - \widetilde{C}_R(x)| \leq \epsilon + C(x+2\delta) - C(x), \tag{124}$$

which gives

$$\widetilde{C}(x+\delta) \leq C(x+2\delta) + \epsilon. \tag{125}$$

The lemma is then proved. □

A.3.1 Technical lemma

Lemma A.10 (Mollifier, Lemma 5 in [LT21]). *Define $M_{d,\delta}(x)$ to be*

$$M_{d,\delta} := \frac{1}{\mathcal{N}_{d,\delta}} T_d\left(1 + 2\frac{\cos(x) - \cos(\delta)}{1+\cos(\delta)}\right) \tag{126}$$

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

$$\mathcal{N}_{d,\delta} := \int_{-\pi}^{\pi} T_d\left(1 + 2\frac{\cos(x) - \cos(\delta)}{1+\cos(\delta)}\right) \mathrm{d}x. \tag{127}$$

Then

1. $|M_{d,\delta}(x)| \leq \frac{1}{\mathcal{N}_{d,\delta}}$ *for* $x \in [-\pi, -\delta] \cup [\delta, \pi]$, *and* $M_{d,\delta}(x) \geq \frac{1}{\mathcal{N}_{d,\delta}}$ *for* $x \in [-\delta, \delta]$.
2. $\int_{-\pi}^{\pi} M_{d,\delta}(x)\mathrm{d}x = 1$, $1 \leq \int_{-\pi}^{\pi} |M_{d,\delta}(x)|\mathrm{d}x \leq 1 + \frac{4\pi}{\mathcal{N}_{d,\delta}}$.
3. *When* $\tan(\delta/2) \leq 1 - 1/\sqrt{2}$, *we have*

$$\mathcal{N}_{d,\delta} \geq C_1 e^{d\delta/\sqrt{2}} \sqrt{\frac{\delta}{d}} \cdot \mathrm{erf}(C_2\sqrt{d}\delta), \tag{128}$$

for some universal constant C_1, C_2.

The proof can be found in Appendix A in [LT21], and we omit it here.

B Technical details of the Hadamard test of block-encoded observable

In this section, we give detailed analysis of the Hadamard test for block-encodings which plays a crucial role in the proof of Theorem 6.1.

We first note that the quantum state before the final measurements is as follows:

$$|\phi_1\rangle = \begin{cases} \frac{1}{\sqrt{2}}\left(|+\rangle\,|0^m\rangle\,|\phi_0\rangle + |-\rangle\,(I\otimes e^{-iHt_2})U(I\otimes e^{-iHt_1})\,|0^m\rangle\,|\phi_0\rangle\right) & \text{if } W = I, \\ \frac{1}{\sqrt{2}}\left(|+\rangle\,|0^m\rangle\,|\phi_0\rangle + i\,|-\rangle\,(I\otimes e^{-iHt_2})U(I\otimes e^{-iHt_1})\,|0^m\rangle\,|\phi_0\rangle\right) & \text{if } W = S. \end{cases} \tag{129}$$

Case 1: $W = I$ We measure the first two registers. If the outcome is $(0, 0^m)$, the (un-normalized) remaining state is:

$$\begin{aligned}&(\langle 0|\langle 0^m|\otimes I)\frac{1}{\sqrt{2}}\left(|+\rangle|0^m\rangle|\phi_0\rangle + |-\rangle(I\otimes e^{-iHt_2})U(I\otimes e^{-iHt_1})|0^m\rangle|\phi_0\rangle\right)\\ =\ &\frac{1}{2}|\phi_0\rangle + \frac{1}{2\alpha}e^{-iHt_2}Oe^{-iHt_1}|\phi_0\rangle\end{aligned} \tag{130}$$

Hence, this event happens with the following probability:

$$\begin{aligned}&\Pr[\text{the outcome is } (0,0^m)|W=I] \qquad (131)\\ =\ &\langle\phi_0|\left(\frac{1}{2}I + \frac{1}{2\alpha}e^{iHt_1}O^\dagger e^{iHt_2}\right)\left(\frac{1}{2}I + \frac{1}{2\alpha}e^{-iHt_2}Oe^{-iHt_1}\right)|\phi_0\rangle\\ =\ &\frac{1}{4}\left(1 + \frac{1}{\alpha}\langle\phi_0|e^{iHt_1}O^\dagger e^{iHt_2}|\phi_0\rangle + \frac{1}{\alpha}\langle\phi_0|e^{-iHt_2}Oe^{-iHt_1}|\phi_0\rangle + \frac{1}{\alpha^2}\langle\phi_0|e^{iHt_1}O^\dagger Oe^{-iHt_1}|\phi_0\rangle\right). \qquad (132)\end{aligned}$$

Similarly, if the outcome is $(1, 0^m)$, the remaining (un-normalized) state is

$$\frac{1}{2}|\phi_0\rangle - \frac{1}{2\alpha}e^{-iHt_2}Oe^{-iHt_1}|\phi_0\rangle, \tag{133}$$

and the probability is

$$\begin{aligned}&\Pr[\text{the outcome is } (1,0^m)|W=I]\\ =\ &\frac{1}{4}\left(1 - \frac{1}{\alpha}\langle\phi_0|e^{iHt_1}O^\dagger e^{iHt_2}|\phi_0\rangle - \frac{1}{\alpha}\langle\phi_0|e^{-iHt_2}Oe^{-iHt_1}|\phi_0\rangle + \frac{1}{\alpha^2}\langle\phi_0|e^{iHt_1}O^\dagger Oe^{-iHt_1}|\phi_0\rangle\right). \qquad (134)\end{aligned}$$

Hence, the expectation of X is

$$\begin{aligned}\mathbb{E}[X] &= \alpha\cdot(\Pr[\text{the outcome is } (0,0^m)|W=I] - \Pr[\text{the outcome is } (1,0^m)|W=I]) \qquad (135)\\ &= \frac{1}{2}(\langle\phi_0|e^{-iHt_2}Oe^{-iHt_1}|\phi_0\rangle + \langle\phi_0|e^{iHt_1}O^\dagger e^{iHt_2}|\phi_0\rangle)\\ &= \frac{1}{2}(\langle\phi_0|e^{-iHt_2}Oe^{-iHt_1}|\phi_0\rangle + \overline{\langle\phi_0|e^{-iHt_2}Oe^{-iHt_1}|\phi_0\rangle})\\ &= \Re\langle\phi_0|e^{-iHt_2}Oe^{-iHt_1}|\phi_0\rangle. \qquad (136)\end{aligned}$$

Case 2: $W = S$ Similar to the case 1, we have

$$\begin{aligned}&\Pr[\text{the outcome is } (0,0^m)|W=S] \qquad (137)\\ =\ &\langle\phi_0|\left(\frac{1}{2}I - \frac{i}{2\alpha}e^{iHt_1}O^\dagger e^{iHt_2}\right)\left(\frac{1}{2}I + \frac{i}{2\alpha}e^{-iHt_2}Oe^{-iHt_1}\right)|\phi_0\rangle\\ =\ &\frac{1}{4}\left(1 - \frac{i}{\alpha}\langle\phi_0|e^{iHt_1}O^\dagger e^{iHt_2}|\phi_0\rangle + \frac{i}{\alpha}\langle\phi_0|e^{-iHt_2}Oe^{-iHt_1}|\phi_0\rangle + \frac{1}{\alpha^2}\langle\phi_0|e^{iHt_1}O^\dagger Oe^{-iHt_1}|\phi_0\rangle\right). \qquad (138)\end{aligned}$$

And

$$\begin{aligned}&\Pr[\text{the outcome is } (1,0^m)|W=S]\\ =\ &\frac{1}{4}\left(1 + \frac{i}{\alpha}\langle\phi_0|e^{iHt_1}O^\dagger e^{iHt_2}|\phi_0\rangle - \frac{i}{\alpha}\langle\phi_0|e^{-iHt_2}Oe^{-iHt_1}|\phi_0\rangle + \frac{1}{\alpha^2}\langle\phi_0|e^{iHt_1}O^\dagger Oe^{-iHt_1}|\phi_0\rangle\right). \qquad (139)\end{aligned}$$

Hence,

$$\mathbb{E}[Y] = \alpha \cdot (\Pr[\text{the outcome is } (1,0^m)|W=S] - \Pr[\text{the outcome is } (0,0^m)|W=S]) \tag{140}$$

$$= \frac{i}{2}(-\langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle + \overline{\langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle})$$

$$= \Im \langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle . \tag{141}$$

Therefore,

$$\mathbb{E}[X+iY] = \langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle . \tag{142}$$

B.1 Generalized Hadamard test

In this subsection, we study the generalized Hadamard test for block-encodings and we will show that the estimator's variance can be reduced by replacing the first Hadamard gate with an α-dependent single-qubit gate.

Suppose $W = I$ and we replace the first Hadamard gate with the following single-qubit gate:

$$G(a,b,\theta) := \begin{bmatrix} a & b \\ -e^{i\theta}\overline{b} & e^{i\theta}\overline{a} \end{bmatrix}, \tag{143}$$

where $\theta \in \mathbb{R}$, $a, b \in \mathbb{C}$ with $|a|^2 + |b|^2 = 1$.

Then, we have

$$\begin{aligned}
& |0\rangle |0^m\rangle |\phi_0\rangle \\
\xrightarrow{G(a,b,\theta)} & a|0\rangle |0^m\rangle |\phi_0\rangle - e^{i\theta}\overline{b}|1\rangle |0^m\rangle |\phi_0\rangle \\
\xrightarrow{C\text{-}e^{-iHt_1}} & a|0\rangle |0^m\rangle |\phi_0\rangle - e^{i\theta}\overline{b}|1\rangle (I \otimes e^{-iHt_1}) |0^m\rangle |\phi_0\rangle \\
\xrightarrow{C\text{-}U} & a|0\rangle |0^m\rangle |\phi_0\rangle - e^{i\theta}\overline{b}|1\rangle U(I \otimes e^{-iHt_1}) |0^m\rangle |\phi_0\rangle \\
\xrightarrow{C\text{-}e^{-iHt_2}} & a|0\rangle |0^m\rangle |\phi_0\rangle - e^{i\theta}\overline{b}|1\rangle (I \otimes e^{-iHt_2}) U (I \otimes e^{-iHt_1}) |0^m\rangle |\phi_0\rangle \\
\xrightarrow{G(p,q,\rho)} & a(p|0\rangle - e^{i\rho}\overline{q}|1\rangle) |0^m\rangle |\phi_0\rangle - e^{i\theta}\overline{b}(q|0\rangle + e^{i\rho}\overline{p}|1\rangle)(I \otimes e^{-iHt_2}) U (I \otimes e^{-iHt_1}) |0^m\rangle |\phi_0\rangle \\
=: & |\phi_1\rangle .
\end{aligned}$$

Hence, the un-normalized remaining state after the measurement with outcome $(0,0^m)$ is:

$$(\langle 0| \langle 0^m| \otimes I) |\phi_1\rangle = ap|\phi_0\rangle - \frac{e^{i\theta}\overline{b}q}{\alpha} e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle . \tag{144}$$

It implies that

$$\Pr[\text{the outcome is } (0,0^m)|W=I] \tag{145}$$

$$= \langle\phi_0| \left(\overline{ap} I - \frac{e^{-i\theta}b\overline{q}}{\alpha} e^{iHt_1} O^\dagger e^{iHt_2}\right) \left(apI - \frac{e^{i\theta}\overline{b}q}{\alpha} e^{-iHt_2} O e^{-iHt_1}\right) |\phi_0\rangle$$

$$= |a|^2|p|^2 + \frac{|b|^2|q|^2}{\alpha^2} \langle\phi_0| e^{iHt_1} O^\dagger O e^{-iHt_1} |\phi_0\rangle$$

$$- \frac{e^{i\theta}\overline{a}\overline{b}\overline{p}q}{\alpha} \langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle - \frac{e^{-i\theta}abp\overline{q}}{\alpha} \overline{\langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle}. \tag{146}$$

On the other hand, the un-normalized state for the outcome $(1, 0^m)$ is

$$(\langle 1| \langle 0^m| \otimes I) |\phi_1\rangle = -e^{i\rho} a\overline{q} |\phi_0\rangle - \frac{e^{i(\theta+\rho)}\overline{b}\overline{p}}{\alpha} e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle , \tag{147}$$

and the probability is

$$\begin{aligned}
&\Pr[\text{the outcome is } (1, 0^m) | W = I] \qquad (148)\\
&= \langle\phi_0| \left(-e^{-i\rho}\overline{a}qI - \frac{e^{-i(\theta+\rho)}bp}{\alpha} e^{iHt_1} O^\dagger e^{iHt_2} \right) \left(-e^{i\rho} a\overline{q} I - \frac{e^{i(\theta+\rho)}\overline{b}\overline{p}}{\alpha} e^{-iHt_2} O e^{-iHt_1} \right) |\phi_0\rangle \\
&= |a|^2|q|^2 + \frac{|b|^2|p|^2}{\alpha^2} \langle\phi_0| e^{iHt_1} O^\dagger O e^{-iHt_1} |\phi_0\rangle \\
&+ \frac{e^{i\theta}\overline{a}\overline{b}\overline{p}q}{\alpha} \langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle + \frac{e^{-i\theta}abp\overline{q}}{\alpha} \overline{\langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle} \qquad (149)
\end{aligned}$$

If we choose $|p| = |q| = \frac{1}{\sqrt{2}}$, then we have

$$\begin{aligned}
&\Pr[\text{the outcome is } (1, 0^m) | W = I] - \Pr[\text{the outcome is } (0, 0^m) | W = I] \\
&= \Re \frac{4e^{i\theta}\overline{a}\overline{b}\overline{p}q}{\alpha} \langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle . \qquad (150)
\end{aligned}$$

Notice that to make the Hadamard test work, we need the coefficient $\frac{4e^{i\theta}\overline{a}\overline{b}\overline{p}q}{\alpha}$ to be a real or an imaginary number.

Now, we show how to choose the parameters to minimize the variance. Without loss of generality, we may assume $a, b \in (0, 1)$ such that $a^2 + b^2 = 1$ and use p, q to cancel the phase factor, i.e., $e^{i\theta}\overline{a}\overline{b}\overline{p}q = \frac{1}{2}ab$. It gives that:

$$\begin{aligned}
&\Pr[\text{the outcome is } (1, 0^m) | W = I] - \Pr[\text{the outcome is } (0, 0^m) | W = I] \\
&= \frac{2ab}{\alpha} \Re \langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle , \qquad (151)
\end{aligned}$$

and

$$\begin{aligned}
&\Pr[\text{the outcome is } (1, 0^m) | W = I] + \Pr[\text{the outcome is } (0, 0^m) | W = I] \\
&= a^2 + \frac{b^2}{\alpha^2} \langle\phi_0| e^{iHt_1} O^\dagger O e^{-iHt_1} |\phi_0\rangle \qquad (152)
\end{aligned}$$

Now, define the random variable as follows:

$$X := \begin{cases} \frac{\alpha}{2ab} & \text{if the outcome is } (1, 0^m), \\ -\frac{\alpha}{2ab} & \text{if the outcome is } (0, 0^m), \\ 0 & \text{otherwise.} \end{cases} \tag{153}$$

Then, we have

$$\mathbb{E}[X] = \Re \langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle . \tag{154}$$

And we have

$$\begin{aligned}
\mathbf{Var}[X] =& \ \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
=& \frac{\alpha^2}{4a^2b^2} \left(a^2 + \frac{b^2}{\alpha^2} \langle\phi_0| e^{iHt_1} O^\dagger O e^{-iHt_1} |\phi_0\rangle \right) - \left(\Re \langle\phi_0| e^{-iHt_2} O e^{-iHt_1} |\phi_0\rangle \right)^2 . \qquad (155)
\end{aligned}$$

The second term is fixed for any parameters. And for the first term, we have

$$\frac{\alpha^2}{4a^2b^2}\left(a^2+\frac{b^2}{\alpha^2}\left\langle\phi_0\right|e^{iHt_1}O^\dagger Oe^{-iHt_1}\left|\phi_0\right\rangle\right)=\frac{\alpha^2}{4b^2}+\frac{1}{4a^2}\left\langle\phi_0\right|e^{iHt_1}O^\dagger Oe^{-iHt_1}\left|\phi_0\right\rangle \tag{156}$$

$$=\frac{\alpha^2}{4(1-a^2)}+\frac{\|Oe^{-iHt_1}\left|\phi_0\right\rangle\|^2}{4a^2}$$

$$\geq\frac{1}{4}(\alpha+\|Oe^{-iHt_1}\left|\phi_0\right\rangle\|)^2, \tag{157}$$

where the minimizer is at $a:=\sqrt{\frac{\|Oe^{-iHt_1}|\phi_0\rangle\|}{\alpha+\|Oe^{-iHt_1}|\phi_0\rangle\|}}$. However, since we do not know the value of $\|Oe^{-iHt_1}\left|\phi_0\right\rangle\|$, there are two approaches to resolve this issue: (1) use another quantum circuit to estimate $\|Oe^{-iHt_1}\left|\phi_0\right\rangle\|$ and then set the parameters; (2) just take $a:=\sqrt{\frac{1}{\alpha+1}}$. Notice that when the first gate is the Hadamard gate, i.e., $a=\frac{1}{\sqrt{2}}$, we have

$$\mathbf{Var}\left[X\;\middle|\;a=\frac{1}{\sqrt{2}}\right]=\frac{1}{2}(\alpha^2+\|Oe^{-iHt_1}\left|\phi_0\right\rangle\|^2). \tag{158}$$

When $a=\sqrt{\frac{1}{\alpha+1}}$, we have

$$\mathbf{Var}\left[X\;\middle|\;a=\frac{1}{\sqrt{\alpha+1}}\right]=\frac{1}{4}\alpha(\alpha+1)+\frac{1}{4}\|Oe^{-iHt_1}\left|\phi_0\right\rangle\|^2(\alpha+1) \tag{159}$$

$$=\frac{1}{2}(\alpha^2+\|Oe^{-iHt_1}\left|\phi_0\right\rangle\|^2)-\frac{1}{4}(\alpha-1)(\alpha-\|Oe^{-iHt_1}\left|\phi_0\right\rangle\|^2)$$

$$\leq\mathbf{Var}\left[X\;\middle|\;a=\frac{1}{\sqrt{2}}\right], \tag{160}$$

where the last step follows from $\alpha\geq 1$ and $\|Oe^{-iHt_1}\left|\phi_0\right\rangle\|^2\leq 1$. Therefore, we can reduce the estimator's variance by choosing $a=\sqrt{\frac{1}{\alpha}}$. Moreover, if α is large, the new variance is about half of the variance using the Hadamard gate.

Similar strategy can also be used to reduce the variance of the random variable Y.