Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision

Guoming Wang[1], Daniel Stilck-França[2], Ruizhe Zhang[1,3], Shuchen Zhu[1,4], and Peter D. Johnson[5]

[1]Zapata Computing Canada Inc., Toronto, ON, Canada M5V 2Y1
[2]Univ Lyon, ENS Lyon, UCBL, CNRS, Inria, LIP, F-69342, Lyon Cedex 07, France
[3]Department of Computer Science, University of Texas at Austin, Austin, TX 78712, USA
[4]Department of Computer Science, Georgetown University, Washington, DC 20057, USA
[5]Zapata Computing Inc., Boston, MA 02110 USA

September 15, 2022

Abstract

A milestone in the field of quantum computing will be solving problems in quantum chemistry and materials faster than state of the art classical methods. The current understanding is that achieving quantum advantage in this area will require some degree of error correction. While hardware is improving towards this milestone, optimizing quantum algorithms also brings it closer to the present. Existing methods for ground state energy estimation require circuit depths that scale as $\mathcal{O}(1/\epsilon \cdot \text{polylog}(1/\epsilon))$ to reach accuracy ϵ. In this work, we develop and analyze ground state energy estimation algorithms that use just one auxilliary qubit and for which the circuit depths scale as $\mathcal{O}(1/\Delta \cdot \text{polylog}(\Delta/\epsilon))$, where $\Delta \geq \epsilon$ is a lower bound on the energy gap of the Hamiltonian. With this $\widetilde{\mathcal{O}}(\Delta/\epsilon)$ reduction in circuit depth, relative to recent resource estimates of ground state energy estimation for the industrially-relevant molecules of ethelyne-carbonate and PF_6^-, the estimated gate count and circuit depth is reduced by a factor of 43 and 78, respectively. Furthermore, the algorithm can take advantage of larger available circuit depths to reduce the total runtime. By setting $\alpha \in [0, 1]$ and using depth proportional to $\epsilon^{-\alpha}\Delta_{\text{true}}^{-1+\alpha}$, the resulting total runtime is $\widetilde{\mathcal{O}}(\epsilon^{-2+\alpha}\Delta_{\text{true}}^{1-\alpha})$, where Δ_{true} is the true energy gap of the Hamiltonian. These features make our algorithm a promising candidate for realizing quantum advantage in the era of early fault-tolerant quantum computing.

1 Introduction

When will quantum computers solve valuable problems that are out of reach for state-of-the-art classical approaches? To understand this future moment of quantum advantage, we must know what computational problems are most apt and what useful quantum algorithms will be able solve them in the nearest time frame. Despite some recent challenges being illuminated [1], estimating the ground state energy of quantum systems [2] remains one of the leading contenders for the first realization of quantum advantage. Solving this problem efficiently with a quantum computer would be of high value to areas including combustion [3], batteries [4, 5], and catalysts [6]. Considering that the progress in quantum hardware has led to steady improvement [7], we are urged to investigate: what are the minimal quantum resources needed to realize quantum advantage with ground state energy estimation?

The bottleneck quantum resource in implementing powerful quantum algorithms is the number of operations that need to be executed in a single quantum circuit. Standard methods for ground state energy estimation (GSEE) [2, 8] require hundreds of logical qubits and many operations (e.g. greater than 10^{10} T gates [4]) and, therefore, deep quantum circuits (i.e. number of layers of parallel operations). Accordingly, each operation must be implemented with very low error rate (e.g. less than 10^{-10} [4]). This places fault-tolerant overhead demands on the quantum architecture that cannot be realized on today's hardware. The timeline of realizing such architectures is uncertain.

One path to realizing quantum advantage sooner for the problem of ground state energy estimation is through the development of quantum algorithms. In this work, we aim to develop quantum algorithms for ground state energy estimation that reduce the required number of operations and circuit depth compared to previous methods. To understand how this is possible, we must consider the parameters that govern the cost of existing quantum algorithms.

The cost of ground state energy estimation algorithms is typically measured as a function of several key parameters determined by the problem instance and solution specifications. The problem instance is given by a Hamiltonian H and we assume that there is a means of preparing a quantum state ρ that approximates the ground state of H. Three parameters relevant to costing GSEE algorithms are ϵ, the target accuracy of the ground state energy estimate, Δ, a lower bound on the energy gap (i.e. the difference between the smallest and next-smallest eigenvalue of H), and η, a lower bound on the overlap of ρ with the ground state[1]. In terms of asymptotic scaling, existing methods require circuit depths that scale as $\tilde{O}(\epsilon^{-1}\eta^{-2})$[2] [10]. There are methods which improve the circuit depth scaling down to $\tilde{O}(\epsilon^{-1})$, however, there are additional costs that need to be paid in terms of ancilla qubits and multi-qubit control operations [11]. Such methods were costed in [4] and were responsible for the large gate counts of greater than 10^{10}. As mentioned above, the timeline is uncertain for realizing the hardware architectures that are needed to implement these methods for useful problem instances.

In anticipation of these high circuit depth costs, the field has developed methods for GSEE that do not require as deep of quantum circuits. The most popular approach has been the variational quantum eigensolver (VQE) algorithm [12]. However, recent work [3] has shown that, for chemical system sizes of industrial relevance, the time requirements of VQE are too high to make the method practical. Even when more efficient methods are used to speed up the bottleneck subroutine of energy estimation [13], the cost of the algorithm remains too high. It is important to note that we encounter such roadblocks even if we ignore other challenges for VQE such as the classical optimization [14, 15].

In the quest to discover the first realization of quantum advantage, the above challenges have motivated us to explore a "goldilocks" region for ground state energy estimation. We aim to develop ground state energy estimation algorithms which require modest circuit depths while also having provable performance guarantees. The two main components of most ground state energy estimation circuits are the initial state preparation circuit $V : |0\rangle \langle 0| \to \rho$ and the energy estimation circuit U (e.g., the series of controlled time evolutions used in the quantum phase estimation algorithm). For state-of-the-art methods, the circuit depth scaling for ground state preparation is $\tilde{O}(1/\Delta)$ and for ground state estimation is $\tilde{O}(1/\epsilon)$. Accordingly, the total circuit depth is $\tilde{O}(1/\Delta + 1/\epsilon)$. In many problem instances, and especially those in quantum chemistry [8], the spectral gap Δ_{true} (and in many cases the estimated lower bound Δ) is substantially larger than the required accuracy ϵ. Therefore, to reduce the circuit depth of ground state energy estimation, we are

---

[1]An important caveat for all known ground state energy estimation methods (c.f Table I of [9]) is that if η is extremely small, then we have little hope of accurately estimating the ground state energy. Thus, it is common to assume that the Hamiltonian of interest admits a good ground state approximation.

[2]We will use $\tilde{O}(g)$ to abbreviate $O(g \cdot \text{polylog}(g))$.

motivated to improve the scaling in the energy estimation component of the algorithm, rather than the state preparation component. As it might be expected, this circuit depth savings will come at the expense of a higher sample complexity and additional classical post-processing [16].

Recent work [17] has developed ground state energy estimation algorithms in which the energy estimation circuit depths scale as $\tilde{O}(1/\epsilon)$, require just a single ancilla qubit, and involve no costly circuit operations beyond controlled time evolutions. As mentioned in the footnote above, a caveat of this method and all energy estimation algorithms is that a good ground state approximation is needed as input (i.e. η is not too small). This algorithm requires running the circuits of depth $\tilde{O}(\epsilon^{-1})$ multiple times, leading to a total runtime of $\tilde{O}(\epsilon^{-1}\eta^{-2})$. More recently, [9] improved on this result, developing an algorithm with similar characteristics, yet achieving a runtime of $\tilde{O}(\epsilon^{-1}\eta^{-1})$. Both of these methods achieve the so-called Heisenberg limit scaling in the runtime with respect to ϵ. Other recent work [18] combined the ideas in [17] with the principles of QDRIFT [19] to propose a ground state energy estimation algorithm which trades off between number of non-Clifford gates and runtime.

All of these methods employ quantum circuits whose depth scales inversely with the target accuracy ϵ. This circuit depth cost may put some important problem instances out of reach for early fault-tolerant quantum computers. The question addressed by this work is: *do there exist ground state energy estimation methods for which the circuit depth provably scales better than $\tilde{O}(1/\epsilon)$?*

1.1 Main results

In this work, we develop and analyze low-depth ground state energy estimation (GSEE) algorithms with high accuracy for which the circuit depth scales more favorably than $\tilde{O}(1/\epsilon)$. As is typical, the circuit depth and quantum runtime of the algorithm is measured in terms of the number of controlled evolution operations c-exp$(-2\pi \mathbf{i}H)$ referred to as *Hamiltonian evolution time*.

Theorem 1.1 (Low-depth GSEE, informal version of Theorem 5.1). *Let H be a Hamiltonian with spectral gap at least Δ. Suppose we can prepare an initial state ρ such that the overlap with the ground state satisfies $\langle E_0|\rho|E_0\rangle \geq \eta$. Given Δ, η and sufficiently small ϵ, there exists an algorithm to estimate the ground state energy within accuracy ϵ with high probability such that:*

- *The circuit depth, measured in maximal Hamiltonian evolution time, is*

$$\mathcal{T}_{\max} = \mathcal{O}(\Delta^{-1} \cdot \text{poly} \log\left(\epsilon^{-1}\eta^{-1}\Delta\right)). \tag{1}$$

- *The quantum runtime, measured in total Hamiltonian evolution time, is*

$$\mathcal{T}_{\text{tot}} = \mathcal{O}(\eta^{-2}\epsilon^{-2}\Delta \cdot \text{poly} \log\left(\epsilon^{-1}\eta^{-1}\Delta\right)). \tag{2}$$

Remark 1.2. *In our work we measure circuit depth and runtime in terms of the costs needed to implement a unit time evolution operation c-exp$(-2\pi \mathbf{i}H)$. The cost of this operation depends on the Hamiltonian itself and the method used for implementing the time evolution. Therefore, the absolute costs of our method will inherit the costs associated with the Hamiltonian and choice of simulation method. State-of-the-art methods for simulating a Hamiltonian H for a given time t have circuit depths linear in $\|H\|t$, where $\|\cdot\|$ is a matrix norm (e.g., the method in [20] based on quantum signal processing has cost linear in $\|H\|_{\max} t$, while the method in [21] based on quantum singular value transformation has cost linear in $\|H\|_2 t$). The resource costs of our GSEE algorithm depend mainly on the relative accuracy ϵ/Δ, and simultaneously re-scaling the Hamiltonian H (and thus Δ and ϵ) will not alter this crucial factor.*

We emphasize that in Theorem 1.1, η and Δ are merely lower bounds of the values p_0 and $\Delta_{\text{true}} = E_1 - E_0$, respectively. These are common assumptions in the analysis of quantum algorithms whose performance is governed by these quantities [22]. In practice, there are surely instances of Hamiltonians for which these quantities can only be loosely lower bounded with very small values. As with methods in classical quantum chemistry, we anticipate that practitioners of these quantum algorithms will exploit structure in the Hamiltonian to help run such algorithms. For example, for many Hamiltonians of interest the Hartree-Fock state serves as an initial state with large overlap p_0 and can be prepared with low depth (proportional to the system size) on a quantum computer[23]. Previously we mentioned that for many systems in quantum chemistry the true gap is large relative to chemical accuracy. Beyond knowing this, it is important to have some sufficiently accurate estimate of Δ_{true} so that the lower bound Δ can be as large as possible. Classical quantum chemistry methods can be used to obtain estimates of the gap. The method we use to get estimates on the gaps of ethelyne carbonate (EC) and PF_6^- is EOM-CCSD, a gold standard in quantum chemistry. Three standard deviations in accuracy for such estimates is roughly ± 20 mHa (see STEOM-CCSD results in Table 12 of [24]), which we use to inform our lower bounds on the estimate of the gap.

The fact that Δ is simply a lower bound in our algorithm also presents an opportunity. Consider the case where a reduced runtime is desired at the expense of increased circuit depth. One can set Δ to a value between $\widetilde{\mathcal{O}}(\epsilon)$ and Δ_{true} (though this is possible in principle, in practice one might have to use a lower bound on Δ_{true} for one end of the interpolation). By tuning Δ between these two extremes, we achieve an interpolation between an algorithm with low circuit depth and an algorithm that reaches the Heisenberg limit. Specifically, we have:

Corollary 1.3. *Let H be a Hamiltonian with spectral gap Δ_{true}. Suppose we can prepare an initial state ρ such that the overlap with the ground state satisfies $\langle E_0 | \rho | E_0 \rangle \geq \eta$. Then for arbitrary $\alpha \in [0, 1]$, given Δ_{true}, η and sufficiently small ϵ, there exists an algorithm to estimate the ground state energy within accuracy ϵ with high probability such that:*

- *The circuit depth, measured in maximal Hamiltonian evolution time, is*

$$\mathcal{T}_{\max} = \widetilde{\mathcal{O}}(\epsilon^{-\alpha} \Delta_{\text{true}}^{-1+\alpha}). \tag{3}$$

- *The quantum runtime, measured in total Hamiltonian evolution time, is*

$$\mathcal{T}_{\text{tot}} = \widetilde{\mathcal{O}}(\eta^{-2} \epsilon^{-2+\alpha} \Delta_{\text{true}}^{1-\alpha}). \tag{4}$$

Before giving a more detailed overview of the algorithm and the detailed technical results we discuss the implications of the main results. The algorithm we develop achieves a reduction in circuit depth of $\tilde{O}(\Delta/\epsilon)$ compared to state-of-the-art ground state energy estimation algorithms [9]. This reduces the quantum resources needed to solve the GSEE problem, opening the possibility of achieving quantum advantage for GSEE sooner.

As an example, consider the ethelyne carbonate (EC) and PF_6^- molecules analyzed in [4]. For these molecules in the cc-pVDZ basis, we can estimate the energy gaps using EOM-CCSD calculations with ORCA software [25, 26], finding them to be $\Delta_{\text{EC}} = 264 \pm 20$ mHa and $\Delta_{\text{PF}_6^-} = 468 \pm 20$ mHa, respectively. Accordingly, we take the lower bounds on these gaps to be 244 mHa and 448 mHa, respectively. The target accuracy considered in [4] was $\epsilon = 1$ mHa. The standard approach to quantum phase estimation (ignoring the cost due to imperfect ground state preparation) uses a circuit with $2/\epsilon$ applications of $c\text{-}\exp(2\pi \mathbf{i} H)$ to achieve an ϵ accurate estimate with high probability. Including the various logarithmic factors (c.f. Algorithm 5) and setting a
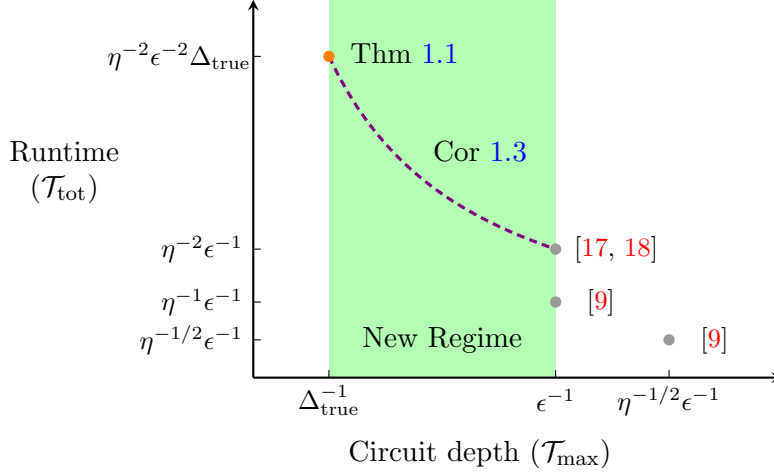
4

Figure 1: This figure shows the landscape of early fault-tolerant GSEE algorithms plotted according to their runtime and circuit depth measured in terms of total evolution time ($\mathcal{T}_{\text{tot}}$) and maximal evolution time ($\mathcal{T}_{\text{max}}$), respectively. The green region indicates the new low-depth regime introduced in this work. The orange dot corresponds to the Δ^{-1}-depth GSEE algorithm (Theorem 1.1) when $\Delta = \Delta_{\text{true}}^{-1}$, and the curve shows the smooth trade off between $\mathcal{T}_{\text{max}}$ and $\mathcal{T}_{\text{tot}}$ described in Corollary 1.3. We also remark that the right-most dot which shows an algorithm in [9] requires multiple ancilla qubits and multi-qubit controlled operations, whereas the algorithms in this work and [17, 18] only use a single ancilla qubit. For simplicity, we have ignored all the poly-logarithmic factors.

conservative value of $\eta = 1/1000$, we find that with our method it is possible to reduce the gate count and circuit depths of ground state energy estimation by a factor of 43 and 78, respectively.

Compared to recent methods [17], which use $2/\pi\epsilon$ applications of $c\text{-}\exp(2\pi\mathbf{i}H)$, the circuit depths are reduced by a factor of 16 and 28, respectively. For molecules with larger gaps, these savings would improve proportionally. Accordingly, the fault-tolerant overhead required to implement ground state energy estimation is reduced. Such reductions may help to bring such problem instances within reach of earlier fault-tolerant quantum architectures, potentially realizing quantum advantage sooner.

It might be the case, however, that the runtime of this low depth algorithm is too high to outperform state-of-the-art classical methods for solving the same problem. Our second main result (c.f. Corollary 5.2) is that we can trade circuit depth for total runtime reduction. This gives a means of speeding up the overall algorithm. Through the era of early fault-tolerant quantum computing, as quantum architectures are able to realize deeper quantum circuits, there may be a crossing point into quantum advantage.

2   Overview of the GSEE algorithm

In this section, we give an overview of our low-depth GSEE algorithm. We first give a formal stating of the GSEE problem. Then we discuss how to overcome the barrier of the quantum circuit depth required by previous algorithms, and propose a general approach for efficiently evaluating any convolution of the spectral density (c.f. Equation 7).
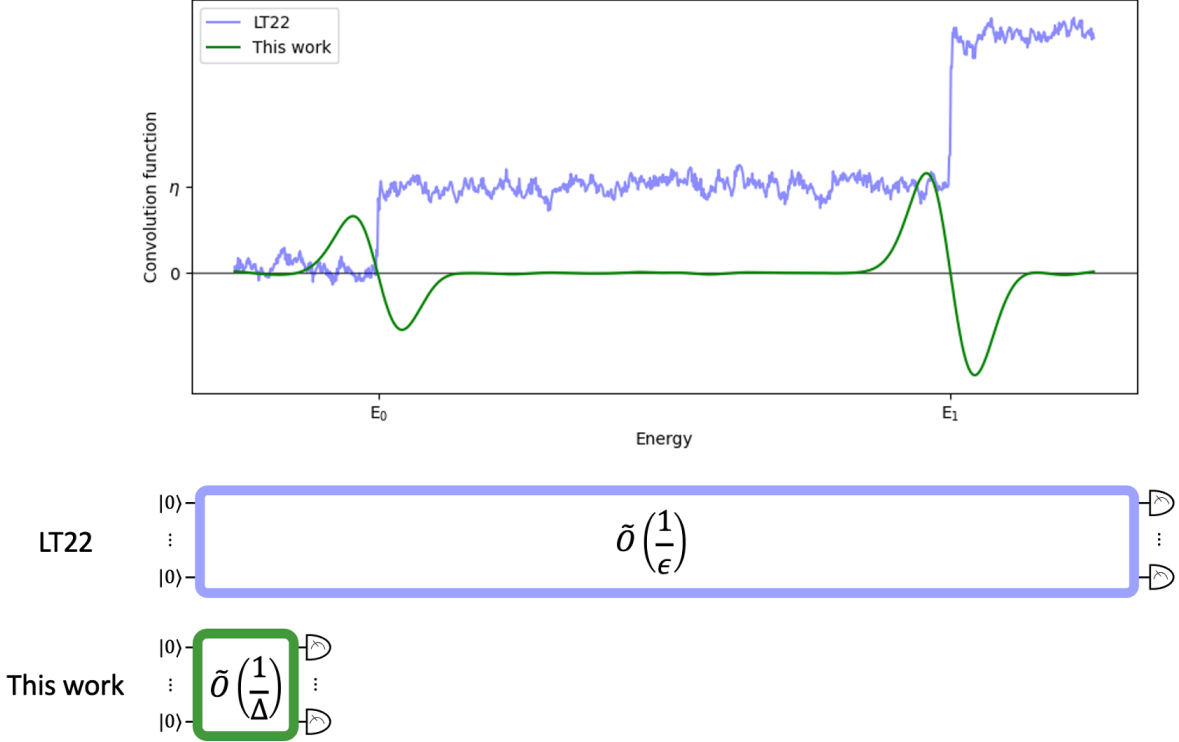
Figure 2: This figure compares the convolution functions and circuit depths used in the ground state energy estimation method of LT22 [17] and the method developed here. The LT22 method uses a Heaviside convolution, while our method uses a Gaussian derivative convolution. Their method requires a steep jump in the convolution function, which necessitates $\tilde{O}(1/\epsilon)$-depth circuits. Our method only requires that contribution of the excited state energies to the convolution function does not interfere too much with that of the ground state energy. This affords the use of a less-steep convolution function, which only requires $\tilde{O}(1/\Delta)$-depth circuits. The trade-off is that our method requires more samples, leading to an increased total runtime.

Problem formulation. Suppose we are given a classical description of a quantum Hamiltonian H. This Hamiltonian has (unknown) spectral decomposition $H = \sum_{j=0}^{N-1} E_j |E_j\rangle \langle E_j|$, where $E_0 < E_1 \leq E_2 \leq ... \leq E_{N-1}$ are the eigenvalues of H, and the $|E_j\rangle$'s are orthonormal eigenstates of H. Let ρ be an easy-to-prepare state (of the same dimension as H). Let $p_j := \langle E_j| \rho |E_j\rangle$ be the overlap between ρ and $|E_j\rangle$, for $0 \leq j \leq N - 1$. We assume that two numbers $\eta \in (0,1)$ and $\Delta > 0$ are given such that $p_0 = \langle E_0| \rho |E_0\rangle \geq \eta$ and $E_1 - E_0 \geq \Delta$. Our goal is to estimate E_0 with accuracy ϵ and confidence $1 - \delta$, i.e. to output a sample from a random variable $\hat{E}_0$ such that

$$\mathbb{P}\left[|\hat{E}_0 - E_0| > \epsilon\right] < \delta, \tag{5}$$

for given small $\epsilon > 0$ and $\delta \in (0,1)$. Furthermore, we want to achieve this by using only Hadamard tests (in which the unitary operation is controlled-$e^{\mathbf{i}Ht}$ for some small $t \in \mathbb{R}$) as well as classical post-processing.

The quantum circuit of the Hadamard test is given in Figure 3. Let $\mathbf{b} \in \{0,1\}$ denote the measurement outcome of the circuit. Then one can show that the expectation $\mathbb{E}[(-1)^{\mathbf{b}}]$ equals the real or imaginary part of $\mathrm{tr}\left[\rho e^{-\mathbf{i}H\tau}\right]$ depending on whether $W = I$ or $W = S^\dagger$ where S is the phase gate.
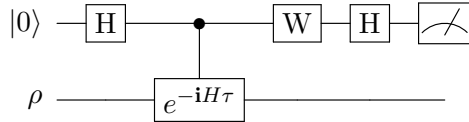
6

Figure 3: Hadamard test circuit parameterized by the Hamiltonian evolution time τ. H is the Hadamard gate and W is either I or $S^\dagger$, where S is the phase gate.

Intuitive introduction to the methods. Before giving an overview of previous methods, we will provide an intuitive description of why one might expect to be able to solve the GSEE problem given data from these Hadamard tests. Our hope is that this will help the reader to anticipate the subsequent methods that will be discussed and to appreciate what is achieved by the algorithms developed in this work.

It is helpful to view the quantity $\mathrm{tr}\left[\rho e^{-\mathbf{i}H\tau}\right]$ as a complex-valued time signal, with τ being the time. This time signal encodes information about the eigenvalues of H and the density operator ρ. In particular, if we can determine how this signal depends on the ground state energy E_0, then we might be able to estimate E_0 from the time signal. Although we are unable to exactly determine the time signal, we can estimate the real and imaginary parts of the signal at any time τ to within any desired accuracy using sufficiently many Hadamard test measurement outcomes as described above. The time cost of each Hadamard test is proportional to τ and the total time cost will depend on how many Hadamard tests, or samples, we take over the different chosen times τ.

To understand how the ground state energy can be estimated from the signal, we can express the quantities inside the trace in the energy eigenbasis of H,

$$\mathrm{tr}\left[\rho e^{-\mathbf{i}H\tau}\right] = \sum_{j=0}^{N-1} \langle E_j|\, \rho\, |E_j\rangle\, e^{-\mathbf{i}E_j\tau}. \tag{6}$$

The signal is composed of a mixture of "pure tones" $e^{-\mathbf{i}E_j\tau}$, one for each eigenvalue and each with weight $\langle E_j|\, \rho\, |E_j\rangle$. Observe that if $\langle E_0|\, \rho\, |E_0\rangle = 0$, then the signal has no dependence on E_0 and there is no hope of accurately estimating the ground state energy. Also, if $\langle E_0|\, \rho\, |E_0\rangle$ is too small, then the ground state energy will be difficult to distinguish from any statistical fluctuations in the signal that are due to finite sampling. Intuitively, the performance of any algorithm for GSEE should depend on $\langle E_0|\, \rho\, |E_0\rangle$, the weight of the ground state energy pure tone in the signal. Generally, it is difficult to estimate $\langle E_0|\, \rho\, |E_0\rangle$ exactly. A weaker and more reasonable assumption is that a lower bound $\eta \leq \langle E_0|\, \rho\, |E_0\rangle$ is given.

Finally, with a promise on the weight of the ground state energy pure tone $e^{-\mathbf{i}E_0\tau}$, how are we to estimate the value of E_0? Observe that the Fourier modes of the signal $\mathrm{tr}\left[\rho e^{-\mathbf{i}H\tau}\right]$ *are* the pure tones corresponding to each eigenvalue. Therefore, the Fourier transform of the true signal would reveal a series of peaks located at frequencies that correspond to the eigenvalues. The challenge is that we cannot learn the exact signal. At best we can estimate the value of the signal at a finite subset of times yielding a noisy, discretized version of the signal. Although we cannot determine the exact Fourier transform output, or *frequency signal*, we can calculate the discrete Fourier transform of the noisy, discretized signal estimate. Remarkably, with a sufficiently accurate estimate of the discrete time signal, this noisy, discretized frequency signal exhibits enough resemblance to the exact frequency signal to pick out the ground state energy.

In this framework, the bulk of any GSEE algorithm design and analysis amounts to choosing how accurately to estimate the signal, choosing how to extract the ground state energy from this signal estimate, and proving that, with high probability, an accurate estimate is obtained. When choosing how accurately to estimate the signal, we must choose a finite time window in which to estimate the time signal. As is common in signal processing, a longer time window translates into a higher-resolution frequency signal. For example, imagine two swings swinging in unison. If their frequencies are different, but very close to one another, we would have to wait a long time before we could determine which was swinging faster. Alternatively, say we only let the swings swing for a short period of time. Is it possible to determine which is swinging faster? The answer is yes, but at a cost. We would need to very accurately measure the position of the two swings.

This trade-off is analogous to the one explored in our work. Instead of the maximum swing duration, we have the maximum evolution time τ, which is proportional to quantum circuit depth. And, instead of accuracy in measuring the position of the swings, we have the accuracy of estimating the signal. The accuracy of estimating the signal is determined by the number of Hadamard test samples. Previous GSEE methods used deeper quantum circuits and fewer Hadamard tests, while our method uses shallower quantum circuits and more Hadamard tests.

Overview of previous early fault-tolerant GSEE methods. Here we give a high-level overview of Lin and Tong's early fault-tolerant algorithms for GSEE in [17], which involve similar techniques to ours. The main idea is to process the time signal described above such that the output frequency signal is a mixture of step functions (i.e. Heaviside functions). The ground state energy is then estimated by locating the first step of the frequency signal. More specifically, let

$$p(x) := \sum_{j=0}^{N-1} p_j \delta(x - E_j) \tag{7}$$

be the spectral measure of H associated with the initial state ρ. Then consider the convolution of $p(x)$ and $\Theta(x)$, the 2π-periodic Heaviside function[3]:

$$C(x) := (\Theta * p)(x) = \sum_{j=0}^{N-1} p_j \cdot \mathbf{1}_{x \geq E_j} \quad \forall x \in (-\pi, \pi), \tag{8}$$

assuming that the eigenvalues of H are between $(-\pi/3, \pi/3)$. Since $p_0 > 0$ and $p_j \geq 0$ for $1 \leq j \leq N-1$, we know that E_0 is the first non-zero point of $C(x)$ for $x \in (-\pi, \pi)$. Moreover, $C(x)$ is non-decreasing in $(-\pi, \pi)$. Thus, if we could evaluate the function $C(x)$ at different points x, then we could narrow in on the first step of $C(x)$ (located at E_0) to any desired accuracy using a binary search: each step of the search determines whether x is to the left or to the right of E_0 by checking whether $C(x) = 0$ or $C(x) > 0$. Unfortunately, we cannot directly compute $C(x)$, but we can estimate:

$$C(x) = \int_{-\infty}^{\infty} \hat{\Theta}(t) e^{2\pi \mathbf{i} x t} \operatorname{tr}\left[\rho e^{-2\pi \mathbf{i} H t}\right] dt, \tag{9}$$

with the help of a quantum computer. Here $\hat{\Theta}(t)$ is the Fourier coefficient of $\Theta(x)$. Note that $\operatorname{tr}\left[\rho e^{-2\pi \mathbf{i} H t}\right]$ can be estimated via the Hadamard-test circuit with evolution time $2\pi t$. Thus, Eq. (9) gives a way to estimate $C(x)$. One issue is that the Fourier spectrum of $\Theta(x)$ is unbounded, which

---

[3]$\Theta(x) = \begin{cases} 1 & \text{if } x \in [2k\pi, (2k+1)\pi) \\ 0 & \text{if } x \in [(2k-1)\pi, 2k\pi) \end{cases} \quad \forall k \in \mathbb{Z}.$

means we need to evolve the Hamiltonian for infinite time. Lin and Tong [17] resolved this problem by designing a low Fourier-degree function $F(x)$ that approximates $\Theta(x)$ in $[-\pi + \epsilon, -\epsilon] \cup [\epsilon, \pi - \epsilon]$. Then we have

$$\widetilde{C}(x) := (F * p)(x) = \sum_{j=-d}^{d} \hat{F}(j) e^{2\pi \mathbf{i} j x} \operatorname{tr}\left[\rho e^{-2\pi \mathbf{i} H j}\right].$$

If suffices to estimate $2d + 1$ terms to approximately evaluate $\widetilde{C}(x)$. They also proved that $C(x) \approx \widetilde{C}(y)$ for some y that is ϵ-close to x. Therefore, the first non-zero point of $C(x)$ can be approximated within ϵ-accuracy by a *robust binary search* on $\widetilde{C}(x)$.

We now describe the costs of this algorithm. The maximal Hamiltonian evolution time (i.e., circuit depth) is proportional to the Fourier degree d of $F(x)$. The fact that the Heaviside function $\Theta(x)$ is discontinuous at 0 imposes challenges to approximating it by a low Fourier-degree function or other low-degree smooth functions in a neighborhood of 0 [27, 28]. To the best of our knowledge, the state-of-the-art result is by Wan, Berta and Campbell [18], which obtained an $\mathcal{O}(\epsilon^{-1})$-Fourier degree function that approximates $\Theta(x)$ up to ϵ in $[-\pi, \pi] \backslash (-\epsilon, \epsilon)$. Hence, the maximal evolution time of this approach is $\mathcal{O}(1/\epsilon)$. As to the expected total evolution time (i.e., quantum runtime), they used the *multi-level Monte Carlo method* to achieve $\widetilde{\mathcal{O}}(\epsilon^{-1} \eta^{-2})$-time.

Overcoming the $1/\epsilon$-depth barrier. We observe that the bottleneck of the maximal evolution time in previous methods is due to the high Fourier approximation degree of the Heaviside function. It is natural to ask: can we choose another filter function that can still help us estimate the minimum eigenvalue while having a lower Fourier approximate degree (or band-limit)?

Indeed, we have a key observation for the design of filter functions. If a filter function satisfies the following properties, then it can isolate the minimum eigenvalue from the others well and the corresponding convolution can be evaluated easily:

1. The filter function $f(x)$ has an exponentially-decaying tail, i.e., $|f(x)| = \exp\left(-\Omega(|x|)\right)$ for sufficiently large x. This enables the filter function to "almost" eliminate the interference of other eigenvalues to the peak around E_0.

2. The filter function's Fourier transform $\hat{f}(t)$ also has an exponentially-decaying tail, i.e., $|\hat{f}(t)| = \exp\left(-\Omega(|t|)\right)$ for sufficiently large t. This allows f to be well-approximated by a band-limited function, which means that the maximal evolution time in the Hadamard tests will be small.

Based on this observation, a natural choice is the Gaussian filter, defined as:

$$f_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}x^2/\sigma^2}, \tag{10}$$

where $\sigma > 0$ is a parameter to be chosen later. Note that its Fourier transform is another Gaussian kernel (up to some scaling factor):

$$\hat{f}_\sigma(t) = e^{-\frac{1}{2}(\sigma \pi t)^2}.$$

Thus, most of its mass is concentrated within $|t| = \mathcal{O}(\sigma^{-1})$. More importantly, by convolving f_σ with the spectral measure p, we get:

$$(f_\sigma * p)(x) = \sum_{j=0}^{N-1} p_j \cdot f_\sigma(x - E_j),$$
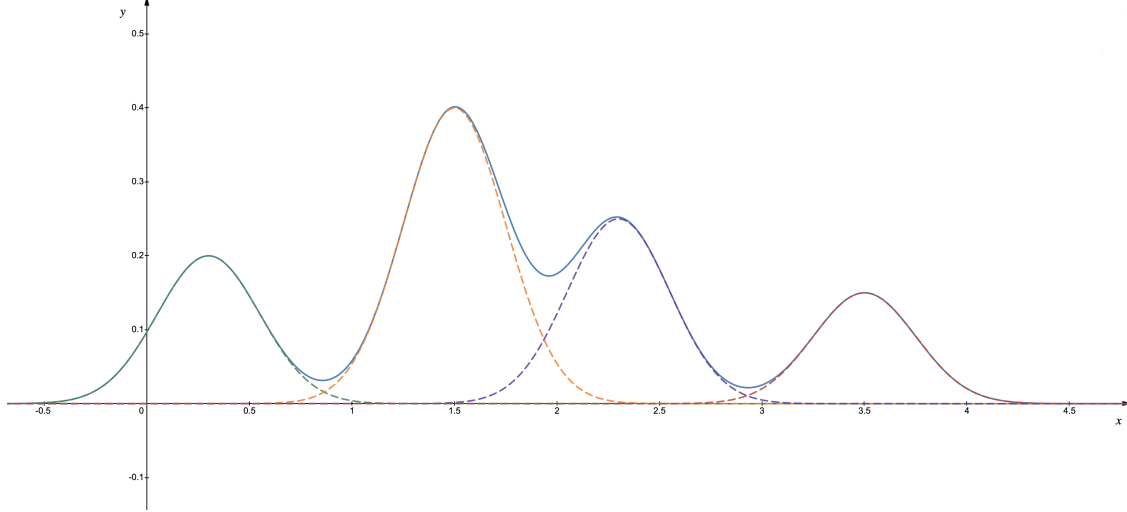
9

Figure 4: This figure illustrates the convolution $(f_\sigma * p)(x) = \sum_{j=0}^{3} p_j f_\sigma(x - E_j)$ for $\sigma = 0.25$, $(p_0, p_1, p_2, p_3) = (0.2, 0.4, 0.25, 0.15)$ and $(E_0, E_1, E_2, E_3) = (0.3, 1.5, 2.3, 3.5)$. The solid curve is $(f_\sigma * p)(x)$, while the dashed curves are $p_j f_\sigma(x - E_j)$ for $j = 0, 1, 2, 3$ respectively. Note that $(f_\sigma * p)(x)$ resembles $p_0 f_\sigma(x - E_0)$ in a neighborhood of $E_0 = 0.3$.

which is a mixture of Gaussians. Figure 4 illustrates an example of $f_\sigma * p$.

Since the Gaussian filter has an exponentially-decaying tail, if we zoom-in to a neighborhood of E_0, the convolution value is dominated by the first Gaussian kernel $p_0 \cdot f_\sigma(x - E_0)$. Therefore, the first significant *peak* of $f_\sigma * p$ will be close to E_0 and the GSEE is then reduced to a *peak finding* problem. Our approach for this problem is to first run the Lin and Tong's algorithm ([17]) with low accuracy (i.e., $\widetilde{E}_0 \in [E_0 - \mathcal{O}(\sigma), E_0 + \mathcal{O}(\sigma)]$). Then we partition the interval $[\widetilde{E}_0 - \sigma, \widetilde{E}_0 + \sigma]$ into a $\mathcal{O}(\epsilon)$-width grid, and estimate the convolution $f_\sigma * p$ at each grid point. Finally, we output the position of the grid point with maximum convolution value. The complexity of this algorithm depends on σ, the width of the Gaussian filter, since we can only truncate its spectrum to $[-T, T]$ for $T = \widetilde{\Theta}(1/\sigma)$ in order to evaluate the convolution with enough precision. We prove that for sufficiently small accuracy-parameter ϵ, we can take $\sigma = \mathcal{O}(\Delta/\mathrm{polylog}(\Delta \epsilon^{-1} \eta^{-1}))$ such that the algorithm can output an estimate for E_0 within ϵ-additive error. It implies that the maximal Hamiltonian evolution time of our algorithm is $\widetilde{\mathcal{O}}(1/\Delta)$. We note that when $\epsilon \ll \Delta$, i.e., in the high-accuracy regime, our algorithm has shallower quantum circuit depth than previous methods [17, 18].

A universal approach for the convolution evaluation. It remains to design a sample-efficient method to evaluate the convolution $f_{\sigma,T} * p$, where

$$f_{\sigma,T}(x) := \int_{-T}^{T} \hat{f}_\sigma(t) e^{2\pi \mathbf{i} x t} \mathrm{d}t$$

is the band-limited approximation of f_σ. In this work, we propose a general approach to evaluating such convolution for a large family of filter functions, which may be of independent interest. Let f_T be any function such that $\mathrm{supp}(\hat{f}_T) \subseteq [-T, T]$ and $\hat{f}_T$ is either continuous or a weighted sum of Dirac delta functions. Recall that

$$(f_T * p)(x) = \int_{-T}^{T} \hat{f}_T(t) e^{2\pi \mathbf{i} t x} \mathrm{tr}\left[\rho e^{-2\pi \mathbf{i} H t} \right] \mathrm{d}t.$$

10

The key idea for estimating this integral is to sample $t \in [-T, T]$ from a distribution with the following probability density:

$$\nu(t) := \frac{|\hat{f}_T(t)|}{\|\hat{f}_T\|_1}, \quad \forall t \in [-T, T],$$

where $\|\hat{f}_T\|_1$ is the L_1-norm of $\hat{f}_T$. By a change of variables, we have

$$(f_T * p)(x) = \int_{-T}^{T} \|\hat{f}_T\|_1 e^{2\pi \mathbf{i}(tx + \phi(t))} \operatorname{tr}\left[\rho e^{-2\pi \mathbf{i} H t}\right] \nu(t) \mathrm{d}t,$$

where $\phi(t)$ is the phase of $\hat{f}_T(t)$. Then, for S independent samples $t_1, \ldots, t_S \sim \nu$, our estimator for $(f_T * p)(x)$ is defined as:

$$\overline{Z}(x) := \frac{\|\hat{f}_T\|_1}{S} \sum_{i=1}^{S} e^{2\pi \mathbf{i}(t_i x + \phi(t_i))} Z_{t_i},$$

where Z_{t_i} is an estimate for $\operatorname{tr}\left[\rho e^{-2\pi \mathbf{i} H t_i}\right]$ obtained by running the Hadamard-test circuit twice (one for the real part and another for the imaginary part). To upper bound the sample complexity S, we observe that $|e^{2\pi \mathbf{i}(t_i x + \phi(t_i))} Z_{t_i}| = \mathcal{O}(1)$. Thus, by Hoeffding's bound, it suffices to take $S = \widetilde{\mathcal{O}}(\epsilon_1^{-2} \|\hat{f}_T\|_1^2)$ to achieve ϵ_1-accuracy.

We now sketch how to bound the total evolution time of our algorithm with the Gaussian filter. As we will see, its performance will be sub-optimal, so we will only analyze a more refined version with better performance in detail. For the truncated Gaussian filter $f_{\sigma, T}$, we have $\|\hat{f}_{\sigma, T}\|_1 = \mathcal{O}(\sigma^{-1})$, and we need to evaluate the convolution with accuracy $\epsilon_1 = \widetilde{\mathcal{O}}(\eta \epsilon^2 \sigma^{-3})$. To see this, note that the Taylor expansion for the Gaussian density around 0 up to second order yields

$$f_\sigma(\epsilon) \simeq \frac{1}{\sqrt{2\pi}\sigma} - \frac{\epsilon^2}{\sqrt{2\pi}\sigma^3},$$

and we see that the peak will decrease by a factor of $\mathcal{O}(\eta \epsilon^2 \sigma^{-3})$ around the ground state. Thus, by our choice of $\sigma = \widetilde{\mathcal{O}}(\Delta)$, the sample complexity is $\widetilde{\mathcal{O}}(\eta^{-2} \epsilon^{-4} \sigma^6 \cdot \sigma^{-2}) = \widetilde{\mathcal{O}}(\eta^{-2} \epsilon^{-4} \Delta^4)$. Hence, the total evolution time is $\mathcal{T}_{\mathsf{tot}} \leq \widetilde{\mathcal{O}}(\eta^{-2} \epsilon^{-4} \Delta^4 \cdot T) = \widetilde{\mathcal{O}}(\eta^{-2} \epsilon^{-4} \Delta^3)$ as $T = \widetilde{\mathcal{O}}(1/\sigma) = \widetilde{\mathcal{O}}(1/\Delta)$.

Reducing $\mathcal{T}_{\mathsf{tot}}$ via Gaussian derivative filtering. The bottleneck of our total evolution time is the *normalized* convolution evaluation accuracy $\frac{\epsilon_1}{\|\hat{f}_T\|_1}$, which equals to $\mathcal{O}(\eta \epsilon^2 \sigma^{-2})$ for the Gaussian filter f_σ.

To improve this factor, we switch to the Gaussian derivative filter g_σ which is defined as follows:

$$g_\sigma(x) := -\frac{1}{\sqrt{2\pi}\sigma^3} x e^{-\frac{x^2}{2\sigma^2}}.$$

Figure 5 illustrates an example of $g_\sigma * p$. Since the Gaussian derivative filter has an exponentially-decaying tail, $(g_\sigma * p)(x)$ resembles $p_0 g_\sigma(x - E_0)$ in a neighborhood of E_0. In particular, the unique zero point of $g_\sigma * p$ in this region is close to E_0.

The Gaussian derivative filter allows for a more favorable normalized convolution evaluation accuracy. On the one hand, the separation between the absolute values of the convolution $(g_\sigma * p)(x)$ for x that is $\epsilon/2$-close to E_0 and for x that is ϵ-far from E_0 is $\Omega(\eta \epsilon \sigma^{-3})$. So it suffices to pick $\epsilon_1' := \mathcal{O}(\eta \epsilon \sigma^{-3})$. On the other hand, it is easy to show that $\|\hat{g}_{\sigma, T}\|_1 = \Theta(\sigma^{-2})$. This implies that
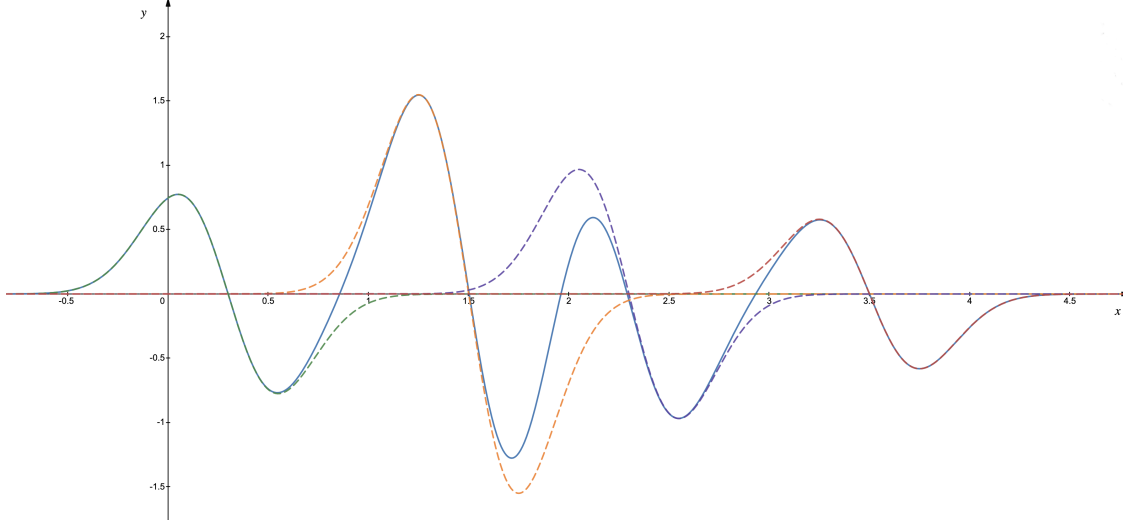
Figure 5: This figure illustrates the convolution $(g_\sigma * p)(x) = \sum_{j=0}^{3} p_j g_\sigma(x - E_j)$ for $\sigma = 0.25$, $(p_0, p_1, p_2, p_3) = (0.2, 0.4, 0.25, 0.15)$ and $(E_0, E_1, E_2, E_3) = (0.3, 1.5, 2.3, 3.5)$. The solid curve is $(g_\sigma * p)(x)$, while the dashed curves are $p_j g_\sigma(x - E_j)$ for $j = 0, 1, 2, 3$ respectively. Note that $(g_\sigma * p)(x)$ resembles $p_0 g_\sigma(x - E_0)$ in a neighborhood of $E_0 = 0.3$.

the required normalized convolution evaluation accuracy for g_σ is $\frac{\epsilon_1'}{\|\hat{g}_{\sigma,T}\|_1} = \mathcal{O}(\eta\epsilon\sigma^{-1})$. Moreover, our GSEE and convolution evaluation approaches are general so that they can be easily adapted to the Gaussian derivative filter function with almost the same parameters (i.e., $\sigma = \widetilde{\mathcal{O}}(\Delta)$ and $T = \widetilde{\mathcal{O}}(1/\sigma)$). Therefore, using g_σ in our algorithm, the maximal evolution time remains to be $\mathcal{T}_{\max} = \widetilde{\mathcal{O}}(\Delta^{-1})$ and the total evolution time is reduced to $\mathcal{T}_{\mathsf{tot}} = \widetilde{\mathcal{O}}(\eta^{-2}\epsilon^{-2}\Delta)$.

3 Estimating ground state energy via Gaussian derivative filtering

In this section, we propose a strategy for GSEE based on Gaussian derivative filtering. In Section 3.1, we define the Gaussian derivative function and prove a nice property of the convolution between this filter and the spectral measure p. In Section 3.2, we show how this property leads to a strategy for GSEE. In Section 3.3, we prove that the Gaussian derivative function can be approximated by a band-limited function, which is crucial for efficient evaluation of the convolution.

3.1 Convolving the spectral measure with a Gaussian derivative filter

Let us start by defining the Gaussian derivative function and demonstrating its properties. Specifically, let $\sigma > 0$ be arbitrary, and let $f_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}$ be a Gaussian function. The Fourier transform of f_σ is

$$\hat{f}_\sigma(\xi) = e^{-\frac{1}{2}(\sigma\pi\xi)^2}. \tag{11}$$

Now consider the derivative of f_σ, i.e.,

$$g_\sigma(x) := f_\sigma'(x) = -\frac{1}{\sqrt{2\pi}\sigma^3}xe^{-\frac{x^2}{2\sigma^2}}. \tag{12}$$

Then the Fourier transform of g_σ is

$$\hat{g}_\sigma(\xi) = 2\pi i \xi \hat{f}_\sigma(\xi) = 2\pi i \xi e^{-\frac{1}{2}(\sigma\pi\xi)^2}. \tag{13}$$

The following properties of g_σ and $\hat{g}_\sigma$ will be useful:

Fact 3.1 (Properties of the Gaussian derivative function).

1. $g_\sigma(0) = 0$.

2. $|g_\sigma(x)|$ is even, increases monotonically in $(-\infty, -\sigma] \cup [0, \sigma]$, and decreases monotonically in $[-\sigma, 0] \cup [\sigma, \infty)$.

3. $g_\sigma(x)$ decays exponentially to 0 as $x \to \pm\infty$.

4. $\hat{g}_\sigma(\xi)$ decays exponentially to 0 as $\xi \to \pm\infty$.

Now let us consider the convolution between the filter g_σ and the spectral measure p:

$$(g_\sigma * p)(x) = \sum_{j=0}^{N-1} p_j g_\sigma(x - E_j) = -\frac{1}{\sqrt{2\pi}\sigma^3} \sum_{j=0}^{N-1} p_j (x - E_j) x e^{-\frac{(x-E_j)^2}{2\sigma^2}}. \tag{14}$$

It turns out that if σ is appropriately chosen, then $|(g_\sigma * p)(x)|$ is small only if x is close to E_0, assuming x is at most $O(\sigma)$-away from E_0:

Lemma 3.2. Let $c = \sqrt{2\ln(10/9)} \approx 0.45904$, and let Δ and η be as in the problem formulation in Section 2. Suppose $\epsilon > 0$ is small enough such that $\epsilon \le c \cdot \min\left(\frac{0.9\Delta}{\sqrt{2\ln(9\Delta\epsilon^{-1}\eta^{-1})}}, 0.2\Delta\right)$. Then for

$$\sigma := \min\left(\frac{0.9\Delta}{\sqrt{2\ln(9\Delta\epsilon^{-1}\eta^{-1})}}, 0.2\Delta\right), \tag{15}$$

we have

- $|(g_\sigma * p)(x)| < \frac{0.6\epsilon p_0}{\sqrt{2\pi}\sigma^3}$, $\forall x \in [E_0 - 0.5\epsilon, E_0 + 0.5\epsilon]$.

- $|(g_\sigma * p)(x)| > \frac{0.8\epsilon p_0}{\sqrt{2\pi}\sigma^3}$, $\forall x \in [E_0 - 0.5\sigma, E_0 - \epsilon) \cup (E_0 + \epsilon, E_0 + 0.5\sigma]$.

Proof. Note that our choice of σ and the condition on ϵ imply that $\epsilon \le c\sigma < 0.5\sigma$. As a consequence, we do have $E_0 - 0.5\sigma < E_0 - \epsilon$ and $E_0 + \epsilon < E_0 + 0.5\sigma$. Thus, the interval in the second bullet is well-defined. Moreover, we have

$$|g_\sigma(0.9\Delta)| = \frac{1}{\sqrt{2\pi}\sigma^3} 0.9\Delta e^{-\frac{0.81\Delta^2}{2\sigma^2}}$$

$$\le \frac{1}{\sqrt{2\pi}\sigma^3} 0.1\epsilon\eta \qquad \text{(by the property } \sigma \le \frac{0.9\Delta}{\sqrt{2\ln(9\Delta\epsilon^{-1}\eta^{-1})}} \text{ in Eq. (15))}$$

$$\le \frac{1}{\sqrt{2\pi}\sigma^3} 0.1\epsilon p_0, \tag{16}$$

where the last step follows from $p_0 \ge \eta$.

We prove the first and the second parts of the lemma below.

13

Part I. For any $x \in [E_0 - 0.5\epsilon, E_0 + 0.5\epsilon]$, we have

$$|(g_\sigma * p)(x)| = \left| p_0 g_\sigma(x - E_0) + \sum_{j=1}^{N-1} p_j g_\sigma(x - E_j) \right| \qquad \text{(by Eq. (14))}$$

$$\leq p_0 |g_\sigma(x - E_0)| + \sum_{j=1}^{N-1} p_j |g_\sigma(x - E_j)|$$

$$\leq p_0 |g_\sigma(x - E_0)| + \max_{1 \leq j \leq N-1} |g_\sigma(x - E_j)|. \qquad (17)$$

The first term in Eq. (17) can be bounded as follows:

$$|g_\sigma(x - E_0)| \leq |g_\sigma(0.5\epsilon)| \qquad \text{(by } |x - E_0| \leq 0.5\epsilon < \sigma \text{ and Property 2 in Fact 3.1)}$$

$$= \frac{1}{\sqrt{2\pi}\sigma^3} 0.5\epsilon e^{-\frac{0.25\epsilon^2}{2\sigma^2}} \qquad \text{(by Eq. (12))}$$

$$\leq \frac{1}{\sqrt{2\pi}\sigma^3} 0.5\epsilon. \qquad (18)$$

To upper bound the second term in Eq. (17), first note that for each $j \geq 1$,

$$|x - E_j| \geq E_j - E_0 - 0.5\epsilon \qquad \text{(since } x \in [E_0 - 0.5\epsilon, E_0 + 0.5\epsilon])$$

$$\geq \Delta - 0.5\epsilon \qquad \text{(since } E_j - E_0 \geq E_1 - E_0 \geq \Delta)$$

$$> 0.9\Delta \qquad \text{(by the assumption } \epsilon \leq 0.2c\Delta < 0.1\Delta)$$

$$> \sigma, \qquad (19)$$

where the last step follows from the property $\sigma \leq 0.2\Delta$ in Eq. (15). Then we obtain

$$|g_\sigma(x - E_j)| < |g_\sigma(0.9\Delta)| \qquad \text{(by Eq. (19) and Property 2 in Fact 3.1)}$$

$$\leq \frac{1}{\sqrt{2\pi}\sigma^3} 0.1\epsilon p_0, \qquad (20)$$

where the second step follows from Eq. (16).

Combining Eqs. (17), (18), and (20), we get that for $x \in [E_0 - 0.5\epsilon, E_0 + 0.5\epsilon]$,

$$|(g_\sigma * p)(x)| < p_0 \cdot \frac{1}{\sqrt{2\pi}\sigma^3} 0.5\epsilon + \frac{1}{\sqrt{2\pi}\sigma^3} 0.1\epsilon p_0 = \frac{0.6\epsilon p_0}{\sqrt{2\pi}\sigma^3}. \qquad (21)$$

Part II. For any $x \in [E_0 - 0.5\sigma, E_0 - \epsilon) \cup (E_0 + \epsilon, E_0 + 0.5\sigma]$, we have

$$|(g_\sigma * p)(x)| = \left| p_0 g_\sigma(x - E_0) + \sum_{j=1}^{N-1} p_j g_\sigma(x - E_j) \right| \qquad \text{(by Eq. (14))}$$

$$\geq p_0 |g_\sigma(x - E_0)| - \sum_{j=1}^{N-1} p_j |g_\sigma(x - E_j)|$$

$$\geq p_0 |g_\sigma(x - E_0)| - \max_{1 \leq j \leq N-1} |g_\sigma(x - E_j)|. \qquad (22)$$

14

The first term in Eq. (22) can be lower bounded as follows:

$$|g_\sigma(x - E_0)| > |g_\sigma(\epsilon)| \qquad\qquad \text{(by } \epsilon < |x - E_0| \leq 0.5\sigma \text{ and Property 2 in Fact 3.1)}$$

$$= \frac{1}{\sqrt{2\pi\sigma^3}} \epsilon e^{-\frac{\epsilon^2}{2\sigma^2}} \qquad\qquad \text{(by Eq. (12))}$$

$$\geq \frac{1}{\sqrt{2\pi\sigma^3}} \epsilon e^{-\frac{c^2}{2}} \qquad\qquad \text{(by the assumption } \epsilon \leq c\sigma)$$

$$\geq \frac{1}{\sqrt{2\pi\sigma^3}} 0.9\epsilon, \qquad\qquad\qquad (23)$$

where the last step follows from $c = \sqrt{2\ln(10/9)}$.

To upper bound the second term in Eq. (22), note that for each $j \geq 1$,

$$|x - E_j| \geq E_j - E_0 - 0.5\sigma \qquad \text{(since } x \in [E_0 - 0.5\sigma, E_0 - \epsilon) \cup (E_0 + \epsilon, E_0 + 0.5\sigma])$$

$$\geq \Delta - 0.5\sigma \qquad\qquad\qquad \text{(since } E_j - E_0 \geq E_1 - E_0 \geq \Delta)$$

$$\geq 0.9\Delta > \sigma, \qquad\qquad\qquad\qquad (24)$$

where the last two inequalities follow from the property $\sigma \leq 0.2\Delta$ in Eq. (15). Then we obtain

$$|g_\sigma(x - E_j)| \leq |g_\sigma(0.9\Delta)| \qquad\qquad \text{(by Eq. (24) and Property 2 in Fact 3.1)}$$

$$\leq \frac{1}{\sqrt{2\pi\sigma^3}} 0.1\epsilon p_0, \qquad\qquad\qquad (25)$$

where the last step follows from Eq. (16).

Combining Eqs. (22), (23), and (25), we get that for $x \in [E_0 - 0.5\sigma, E_0 - \epsilon) \cup (E_0 + \epsilon, E_0 + 0.5\sigma]$,

$$|(g_\sigma * p)(x)| > p_0 \cdot \frac{1}{\sqrt{2\pi\sigma^3}} 0.9\epsilon - \frac{1}{\sqrt{2\pi\sigma^3}} 0.1\epsilon p_0 = \frac{0.8\epsilon p_0}{\sqrt{2\pi\sigma^3}}. \qquad (26)$$

The lemma is thus proved. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

3.2 Basic strategy for ground state energy estimation

Lemma 3.2 prompts us to develop the following strategy for estimating ground state energy. We first obtain an estimate $\tilde{E}_0$ of E_0 such that $\tilde{E}_0$ is $O(\sigma)$-close to E_0 with high probability. Then we find a point at which $|(g_\sigma * p)|$ has small value in a region $[\tilde{E}_0 - O(\sigma), \tilde{E}_0 + O(\sigma)]$. Using Lemma 3.2 we can prove that this point is ϵ-close to E_0 with high probability. Formally, we have

Lemma 3.3. *Let Δ, η, ϵ and δ be as in the problem formulation in Section 2. Suppose ϵ satisfies the condition in Lemma 3.2. Let σ be defined as Eq. (15). Suppose $\tilde{E}_0$ is a random variable such that*

$$\mathbb{P}\left[|\tilde{E}_0 - E_0| > \frac{\sigma}{4}\right] < \frac{\delta}{2}. \qquad\qquad\qquad (27)$$

Let $M := \lceil \sigma/\epsilon \rceil + 1$, and let $x_j := \tilde{E}_0 - 0.25\sigma + (0.5\sigma/M) \cdot (j-1)$ for $j \in [M]$. Suppose $h_1, h_2, \ldots, h_M$ are random variables such that

$$\mathbb{P}\left[\forall j \in [M] : |(g_\sigma * p)(x_j) - h_j| \leq \frac{0.1\epsilon\eta}{\sqrt{2\pi\sigma^3}}\right] \geq 1 - \frac{\delta}{2}. \qquad (28)$$

Let $j^* = \arg\min_{1 \leq j \leq M} |h_j|$. Then we have

$$\mathbb{P}\left[|x_{j^*} - E_0| > \epsilon\right] < \delta. \qquad\qquad\qquad (29)$$

Proof. By our assumptions about $\tilde{E}_0$ and $h_1, h_2, \ldots, h_M$ and the union bound, we get that the following events happen simultaneously with probability at least $1 - \delta$:

- $|\tilde{E}_0 - E_0| \leq 0.25\sigma$.

- $|(g_\sigma * p)(x_j) - h_j| \leq \frac{0.1\epsilon\eta}{\sqrt{2\pi}\sigma^3}, \forall j \in [M]$.

In this case, we have $x_0, x_1, \ldots, x_M \in [\tilde{E}_0 - 0.25\sigma, \tilde{E}_0 + 0.25\sigma] \subseteq [E_0 - 0.5\sigma, E_0 + 0.5\sigma]$. Then by Lemma 3.2, we have that

- If $|x_j - E_0| \leq 0.5\epsilon$, then

$$|h_j| \leq |(g_\sigma * p)(x_j)| + |(g_\sigma * p)(x_j) - h_j| < \frac{0.6\epsilon p_0}{\sqrt{2\pi}\sigma^3} + \frac{0.1\epsilon\eta}{\sqrt{2\pi}\sigma^3} \leq \frac{0.7\epsilon p_0}{\sqrt{2\pi}\sigma^3}. \tag{30}$$

- If $|x_j - E_0| > \epsilon$, then

$$|h_j| \geq |(g_\sigma * p)(x_j)| - |(g_\sigma * p)(x_j) - h_j| > \frac{0.8\epsilon p_0}{\sqrt{2\pi}\sigma^3} - \frac{0.1\epsilon\eta}{\sqrt{2\pi}\sigma^3} \geq \frac{0.7\epsilon p_0}{\sqrt{2\pi}\sigma^3}. \tag{31}$$

Meanwhile, note that $x_1 \leq E_0 \leq x_M$, and $|x_{j+1} - x_j| \leq 0.5\epsilon, \forall j \in [M-1]$. So there exists some $j_0 \in [M]$ such that $|x_{j_0} - E_0| \leq 0.5\epsilon$. This implies that $|h_{j^*}| \leq |h_{j_0}| < \frac{0.7\epsilon p_0}{\sqrt{2\pi}\sigma^3}$, which in turn implies that $|x_{j^*} - E_0| \leq \epsilon$. This lemma is thus proved. $\square$

It remains to show how to generate the random variables $\tilde{E}_0$ and $h_1, h_2, \ldots, h_M$ that satisfy the conditions Eqs. (27) and (28) respectively. To obtain $\tilde{E}_0$, we use the GSEE algorithm in [17] which takes $\tilde{O}(\epsilon^{-1})$ maximal Hamiltonian evolution time to achieve ϵ-accuracy. Since $\tilde{E}_0$ only needs $\frac{\sigma}{4}$-accuracy, this step has $\tilde{O}(\sigma^{-1})$ maximal evolution time. To obtain $h_1, h_2, \ldots, h_M$, we first introduce the band-limited version of g_σ, denoted as $g_{\sigma,T}$, in Section 3.3, and prove that $(g_\sigma * p)(x) \approx (g_{\sigma,T} * p)(x)$ for a small T. Then we design a data structure CONVEVAL in Section 4 such that this data structure can evaluate $g_{\sigma,T} * p$ at the points $x_1, x_2, \ldots, x_M$ with high accuracy and confidence after appropriate initialization.

3.3 Gaussian derivative filters with bounded band-limits

In order to efficiently evaluate $g_\sigma * p$ at any given point, we truncate the spectrum of g_σ and construct a T-bandlimit version $g_{\sigma,T}$ such that

$$(g_\sigma * p)(x) \approx (g_{\sigma,T} * p)(x), \quad \forall x \in \mathbb{R}. \tag{32}$$

Specifically, we define $g_{\sigma,T}$ by restricting $\hat{g}_\sigma$ to $[-T, T]$ and performing the inverse Fourier transform:

$$g_{\sigma,T}(x) := \int_{-T}^{T} \hat{g}_\sigma(\xi)e^{2\pi \mathbf{i}x\xi}d\xi. \tag{33}$$

Clearly, $g_{\sigma,T} \to g_\sigma$ as $T \to \infty$. The following lemma shows how to choose T such that $g_{\sigma,T}$ can approximate g_σ in L_∞-norm:

Lemma 3.4. *Let $\epsilon_1 > 0$ be arbitrary. Then for*

$$T := \pi^{-1}\sigma^{-1}\sqrt{2\ln\left(8\pi^{-1}\epsilon_1^{-1}\sigma^{-2}\right)}, \tag{34}$$

we have

$$|g_\sigma(x) - g_{\sigma,T}(x)| \leq \frac{\epsilon_1}{2}, \forall x \in \mathbb{R}. \tag{35}$$

16

Proof. By the Fourier inversion theorem, we have

$$
\begin{aligned}
|g_\sigma(x) - g_{\sigma,T}(x)| &= \left| \int_{-\infty}^{-T} \hat{g}_\sigma(\xi) e^{2\pi \mathbf{i} \xi x} d\xi + \int_{T}^{+\infty} \hat{g}_\sigma(\xi) e^{2\pi \mathbf{i} \xi x} d\xi \right| \\
&\leq \int_{-\infty}^{-T} |\hat{g}_\sigma(\xi)| \, d\xi + \int_{T}^{+\infty} |\hat{g}_\sigma(\xi)| \, d\xi \\
&= 2 \int_{T}^{+\infty} 2\pi \xi e^{-\frac{1}{2}(\sigma \pi \xi)^2} d\xi \\
&= \frac{4}{\sigma^2 \pi} e^{-\frac{1}{2}\sigma^2 \pi^2 T^2}.
\end{aligned}
\tag{36}
$$

By solving the inequality

$$
\frac{4}{\sigma^2 \pi} e^{-\frac{1}{2}\sigma^2 \pi^2 T^2} \leq \frac{\epsilon_1}{2},
\tag{37}
$$

we get that it suffices to take

$$
T \geq \pi^{-1} \sigma^{-1} \sqrt{2 \ln \left(8 \pi^{-1} \epsilon_1^{-1} \sigma^{-2} \right)}.
\tag{38}
$$

$\square$

The following claim shows that the L_∞-approximation for g_σ implies the L_∞-approximation for $g_\sigma * p$.

Claim 3.5. *Let T be defined as in Lemma 3.4. Then we have*

$$
|(g_\sigma * p)(x) - (g_{\sigma,T} * p)(x)| \leq \frac{\epsilon_1}{2}, \quad \forall x \in \mathbb{R}.
$$

Proof. For any $x \in \mathbb{R}$, we have

$$
\begin{aligned}
|(g_\sigma * p)(x) - (g_{\sigma,T} * p)(x)| &= \left| \int_{-\infty}^{\infty} (g_\sigma(z) - g_{\sigma,T}(z)) p(x - z) dz \right| \\
&\leq \int_{-\infty}^{\infty} |g_\sigma(z) - g_{\sigma,T}(z)| \, p(x - z) dz \\
&\leq \frac{\epsilon_1}{2} \int_{-\infty}^{\infty} p(x - z) dz \\
&= \frac{\epsilon_1}{2},
\end{aligned}
\tag{39}
$$

where the first step follows from the definition of convolution, the section step follows from the triangle inequality, the third step follows from Lemma 3.4, and the last step follows from the property of Dirac delta function. $\square$

Claim 3.5 implies that in order to estimate $(g_\sigma * p)(x)$ within ϵ_1-accuracy, it suffices to evaluate $(g_{\sigma,T} * p)(x)$ within $0.5\epsilon_1$-accuracy, which can be achieved by the method in Section 4.

4 Complexity of evaluating the convolution

In this section, we focus on evaluating the convolution between a filter function f and the spectral measure p to within ϵ-additive error. In Section 4.1, we develop an evaluation method for general filter functions with bounded band-limits. Then in Section 4.2, we apply the method to the Gaussian derivative filter used in our GSEE algorithm.

4.1 Evaluating the convolution via Hadamard tests

For the sake of generality, we will not restrict to a specific filter f but consider arbitrary filters with bounded band-limits. Specifically, for a parameter $T > 0$, let f_T be a function with band-limit T, i.e.,

$$f_T(x) = \int_{-T}^{T} \hat{f}_T(t) e^{2\pi \mathbf{i} t x} dt, \tag{40}$$

where $\hat{f}_T$ is the Fourier transform of f_T and satisfies $\hat{f}_T(t) = 0$ for all $t \in (-\infty, -T) \cup (T, +\infty)$. Furthermore, we require that $\hat{f}_T$ is either continuous in $[-T, T]$ or a weighted sum of Dirac delta functions (i.e., f_T has a discrete spectrum). Here we will state the results for the former case, and the reader can easily generalize them to the latter case.

Given such a function f_T, we can define a probability density ν in terms of its Fourier weights:

$$\nu(t) = \frac{|\hat{f}_T(t)|}{\|\hat{f}_T\|_1}, \quad \forall t \in [-T, T]. \tag{41}$$

Moreover, let $\phi(t)$ be the phase of $\hat{f}_T(t)$, i.e., $\hat{f}_T(t) = |\hat{f}_T(t)| e^{\mathbf{i} 2\pi \phi(t)}$. Then we have that

$$f_T(x) = \int_{-T}^{T} \|\hat{f}_T\|_1 e^{2\pi i (tx + \phi(t))} \nu(t) dt. \tag{42}$$

Now given a quantum state ρ, a Hamiltonian H and a parameter $t \in [-T, T]$, we define two random variables $\mathbf{X}_t$ and $\mathbf{Y}_t$ as follows. Let $\mathbf{b}_I$ and $\mathbf{b}_{S^\dagger}$ be the measurement outcome of the circuit in Figure 3 with $\tau = 2\pi t$ and $W = I$ or $S^\dagger$ (where S is the phase gate), respectively. Then we define $\mathbf{X}_t = (-1)^{\mathbf{b}_I}$ and $\mathbf{Y}_t = (-1)^{\mathbf{b}_{S^\dagger}}$. As mentioned in Section 2, we have that

$$\mathbb{E}\left[\mathbf{X}_t\right] = \mathrm{Re}\left(\mathrm{tr}\left[\rho e^{-2\pi \mathbf{i} H t}\right]\right), \quad \mathbb{E}\left[\mathbf{Y}_t\right] = \mathrm{Im}\left(\mathrm{tr}\left[\rho e^{-2\pi \mathbf{i} H t}\right]\right). \tag{43}$$

Now given a point $x \in \mathbb{R}$, we define the random variable $\mathbf{Z}(x)$ as follows. Let $\mathbf{t}$ be a random variable with probability density function ν. Then we define

$$\mathbf{Z}(x) := \|\hat{f}_T\|_1 \cdot e^{2\pi \mathbf{i}(\mathbf{t}x + \phi(\mathbf{t}))} \cdot (\mathbf{X}_\mathbf{t} + \mathbf{i}\mathbf{Y}_\mathbf{t}). \tag{44}$$

It turns out that $\mathbf{Z}(x)$ is an unbiased estimator of the convolution $f_T * p$ at point x:

Lemma 4.1. *For the random variable $\mathbf{Z}(x)$ defined as Eq. (44), we have that*

$$\mathbb{E}[\mathbf{Z}(x)] = (f_T * p)(x), \quad \forall x \in \mathbb{R}. \tag{45}$$

Proof. Let us first consider the conditional expectation $\mathbb{E}[\mathbf{Z}(x)|\mathbf{t} = t]$ for some $t \in [-T, T]$. By Eq. (43) and the definition of $\mathbf{Z}(x)$ in Eq. (44), we get

$$\begin{aligned}
\mathbb{E}[\mathbf{Z}(x)|\mathbf{t} = t] &= \mathbb{E}\left[\|\hat{f}_T\|_1 e^{2\pi \mathbf{i}(\mathbf{t}x + \phi(\mathbf{t}))}(\mathbf{X}_\mathbf{t} + \mathbf{i}\mathbf{Y}_\mathbf{t})\big|\mathbf{t} = t\right] \\
&= \|\hat{f}_T\|_1 e^{2\pi \mathbf{i}(t_0 x + \phi(t))} \mathrm{tr}\left[\rho e^{-2\pi \mathbf{i} H t}\right].
\end{aligned} \tag{46}$$

By the law of total expectation, we have

$$\mathbb{E}[\mathbf{Z}(x)] = \int_{-T}^{T} \mathbb{E}[\mathbf{Z}(x)|\mathbf{t} = t] \cdot \mathbb{P}[\mathbf{t} = t] \, dt$$

$$= \int_{-T}^{T} \|\hat{f}_T\|_1 e^{2\pi \mathbf{i}(tx + \phi(t))} \operatorname{tr}\left[\rho e^{-2\pi \mathbf{i}Ht}\right] \nu(t) dt$$

$$= \int_{-T}^{T} \hat{f}_T(t) e^{2\pi \mathbf{i}tx} \operatorname{tr}\left[\rho e^{-2\pi \mathbf{i}Ht}\right] dt, \tag{47}$$

where the last step follows from the definition of ν in Eq. (41) and the definition of $\phi(t)$.

It remains to prove that the above expression indeed coincides with $f_T * p(x)$. Indeed, we have that:

$$(f_T * p)(x) = \int_{-\infty}^{\infty} p(x - y) f_T(y) \mathrm{d}y$$

$$= \int_{-\infty}^{\infty} \int_{-T}^{T} p(x - y) \hat{f}_T(t) e^{2\pi \mathbf{i}ty} \mathrm{d}t \mathrm{d}y$$

$$= \int_{-T}^{T} \hat{f}_T(t) \mathrm{d}t \int_{-\infty}^{\infty} p(x - y) e^{2\pi \mathbf{i}ty} \mathrm{d}y. \tag{48}$$

By the definition of $p(x)$ in Eq. (7), we have that

$$\int_{-\infty}^{\infty} p(x - y) e^{2\pi \mathbf{i}ty} \mathrm{d}y = \int_{-\infty}^{\infty} \sum_{k \geq 0} p_k \delta(x - y - E_k) e^{2\pi \mathbf{i}ty} \mathrm{d}y = \sum_{k \geq 0} p_k e^{2\pi \mathbf{i}t(x - E_k)}, \tag{49}$$

where the last step follows from the integration of Dirac delta function. Then, it implies that

$$(f_T * p)(x) = \int_{-T}^{T} \hat{f}_T(t) \mathrm{d}t \cdot \sum_{k \geq 0} p_k e^{2\pi \mathbf{i}t(x - E_k)} = \int_{-T}^{T} \hat{f}_T(t) e^{2\pi \mathbf{i}tx} \operatorname{tr}\left[\rho e^{-2\pi \mathbf{i}Ht}\right] \mathrm{d}t, \tag{50}$$

where the last step follows from $\operatorname{tr}\left[\rho e^{-2\pi \mathbf{i}Ht}\right] = \sum_{k \geq 0} p_k e^{-2\pi \mathbf{i}tE_k}$.

Comparing Eqs. (47) and (50), we conclude that $\mathbb{E}[\mathbf{Z}(x)] = (f_T * p)(x)$ for all $x \in \mathbb{R}$. The lemma is thus proved.

$\square$

With Lemma 4.1 established, it is now straightforward to analyze how many samples we need to estimate the function $f_T * p$ at various points within a target accuracy.

Lemma 4.2 (Sample complexity of the convolution evaluation). *Let $\{(t^{(i)}, X^{(i)}, Y^{(i)})\}_{i=1}^{S}$ be S i.i.d. samples such that $t^{(i)} \sim \nu$, $X^{(i)} \sim \mathbf{X}_{t_i}$ and $Y^{(i)} \sim \mathbf{Y}_{t_i}$, where ν is defined as Eq. (41), and $\mathbf{X}_t$ and $\mathbf{Y}_t$ are the measurement outcome of the circuit in Figure 3 with $\tau = 2\pi t$ and $W = I$ or S, respectively. Let $x_1, x_2, \ldots, x_M \in \mathbb{R}$ be arbitrary. For each $j \in [M]$, let $\overline{Z}_j$ be defined as follows:*

$$\overline{Z}_j := \frac{\|\hat{f}_T\|_1}{S} \sum_{i=1}^{S} e^{2\pi \mathbf{i}(t^{(i)} x_j + \phi(t^{(i)}))} \cdot (X^{(i)} + \mathbf{i}Y^{(i)}). \tag{51}$$

Then for any $\epsilon_1 > 0$ and $\delta_1 \in (0, 1)$, letting

$$S := \left\lceil \frac{\|\hat{f}_T\|_1^2 \ln(4M/\delta_1)}{\epsilon_1^2} \right\rceil, \tag{52}$$

19

we have

$$\mathbb{P}\left[\forall j \in [M] : |\overline{Z}_j - (f_T * p)(x_j)| \leq \epsilon_1\right] \geq 1 - \delta_1. \tag{53}$$

Proof. Recall that $\mathbf{Z}(x) = \|\hat{f}_T\|_1 \cdot e^{2\pi \mathbf{i}(\mathbf{t}x + \phi(\mathbf{t}))} \cdot (\mathbf{X_t} + \mathbf{i}\mathbf{Y_t})$ for any $x \in \mathbb{R}$. Then $\bar{Z}_j$ is the empirical mean of S i.i.d. samples of $\mathbf{Z}(x_j)$ that correspond to $\{(t^{(i)}, X^{(i)}, Y^{(i)})\}_{i=1}^S$, for each $j \in [M]$. Since $\mathbf{X_t}$ and $\mathbf{Y_t}$ take values in $\{1, -1\}$, we know that $\mathrm{Re}\,(\mathbf{Z}(x))$ and $\mathrm{Im}\,(\mathbf{Z}(x))$ take values in $[-\|\hat{f}_T\|_1, \|\hat{f}_T\|_1]$. It then follows from Hoeffding's inequality [29] that for our choice of S in Eq. (52), for any $j \in [M]$, it holds that

$$\mathbb{P}\left[|\mathrm{Re}\,(\overline{Z}_j) - \mathbb{E}[\mathrm{Re}\,(\mathbf{Z}(x_j))]| > \frac{\epsilon_1}{\sqrt{2}}\right] < \frac{\delta_1}{2M}, \tag{54}$$

$$\mathbb{P}\left[|\mathrm{Im}\,(\overline{Z}_j) - \mathbb{E}[\mathrm{Im}\,(\mathbf{Z}(x_j))]| > \frac{\epsilon_1}{\sqrt{2}}\right] < \frac{\delta_1}{2M}. \tag{55}$$

Then by the triangle inequality and union bound, we get

$$\mathbb{P}\left[|\overline{Z}_j - \mathbb{E}[\mathbf{Z}(x_j)]| > \epsilon_1\right] < \frac{\delta_1}{M}. \tag{56}$$

Meanwhile, by Lemma 4.1, we know that

$$\mathbb{E}[\mathbf{Z}(x_j)] = (f_T * p)(x_j). \tag{57}$$

Thus, we have

$$\mathbb{P}\left[|\overline{Z}_j - (f_T * p)(x_j)| > \epsilon_1\right] < \frac{\delta_1}{M}. \tag{58}$$

By a union bound over all $j \in [M]$, we get that

$$\mathbb{P}\left[\exists j \in [M] : |\overline{Z}_j - (f_T * p)(x_j)| > \epsilon_1\right] < \delta_1. \tag{59}$$

$\square$

Remark 4.3. *Note that $\overline{Z}_j$ is a complex number in general, but $(f_T * p)(x_j)$ is real provided that f_T is real. In this case, we can re-define $\overline{Z}_j$ as the real part of the right-hand side of Eq. (51) and Lemma 4.2 will still hold. We envision that in some scenarios, it is useful to have a complex filter f_T, and hence define $\overline{Z}_j$ as Eq. (51) for sake of generality.*

Now we give a data structure CONVEVAL in Algorithm 1 for evaluating the convolution $f_T * p$ at multiple points.

Lemma 4.2 immediately implies that:

Corollary 4.4. *Let $x_1, x_2, \ldots, x_M \in \mathbb{R}$ be arbitrary. Suppose the data structure CONVEVAL is initialized with parameters $(f_T, \epsilon, \delta, M)$. Let h_j be the output of the procedure CONVEVAL.EVAL(x_j) for $j \in [M]$. Then we have*

$$\mathbb{P}\left[\forall j \in [M] : |(f_T * p)(x_j) - h_j| \leq \epsilon\right] \geq 1 - \delta. \tag{60}$$

Algorithm 1 Convolution evaluation data structure.

1: **data structure** FILTERSAMPLER
2:     INIT(f_T)   ▷ Initialize for the filter f_T
3:     SAMPLE()   ▷ Sample $\xi \in \mathbb{R}$ with probability $\propto |\hat{f}_T(\xi)|$
4:     NORM()   ▷ Return $\|\hat{f}_T\|_1$
5: **end data structure**
6:
7: **data structure** CONVEVAL
8: **members**
9:     $\mathcal{C}(t, W)$ ▷ Run the circuit in Figure 3 with $\tau = 2\pi t$ and $W = I$ or $S^\dagger$
10:    $\{(t^{(i)}, z^{(i)})\}_{i \in [S]} \subset \mathbb{R} \times \mathbb{C}$ ▷ Fourier samples
11:    FILTERSAMPLER FS   ▷ Filter function's sampler
12: **end members**
13:
14: **procedure** INIT(f_T, ϵ, δ, M)   ▷ ϵ is the target accuracy, δ is the tolerable failure probability,
    M is the maximal number of points at which the convolution is evaluated
15:    FS.INIT(f_T)
16:    $L \leftarrow$ FS.NORM()
17:    $S \leftarrow \left\lceil \frac{L^2 \ln(4M/\delta)}{\epsilon^2} \right\rceil$ ▷ Lemma 4.2
18:    **for** $i \leftarrow 1, 2, \ldots, S$ **do**
19:        $t^{(i)} \leftarrow$ FS.SAMPLE()
20:        $x^{(i)} \leftarrow \mathcal{C}(t^{(i)}, I)$ ▷ Hadamard test
21:        $y^{(i)} \leftarrow \mathcal{C}(t^{(i)}, S^\dagger)$ ▷ Hadamard test
22:        $z^{(i)} \leftarrow L \cdot e^{2\pi\mathbf{i}\phi(t^{(i)})}(x^{(i)} + \mathbf{i}y^{(i)})$
23:    **end for**
24: **end procedure**
25:
26: **procedure** EVAL(x)   ▷ Approximate $(f_T * p)(x)$ within accuracy ϵ
27:    $\overline{Z} \leftarrow \frac{1}{S} \sum_{i \in [S]} e^{2\pi\mathbf{i}t^{(i)}x} \cdot z^{(i)}$
28:    **return** $\overline{Z}$
29: **end procedure**
30: **end data structure**

Lemma 4.5 (Running time of the convolution evaluation data structure). *Suppose the data structure* FILTERSAMPLER *runs in* $\mathcal{O}(1)$-*time. Then, the data structure* CONVEVAL *in Algorithm 1 has the following running times:*

- *Procedure* INIT(f_T, ϵ, δ, M) *has* $\mathcal{O}(T)$ *maximal evolution time and* $\mathcal{O}(ST)$ *total evolution time, where* $S = \mathcal{O}(\epsilon^{-2}\|\hat{f}_T\|_1^2 \log(\delta^{-1}M))$.

- *Procedure* EVAL(x) *has* $\mathcal{O}(S)$ *classical post-processing time.*

Proof. The CONVEVAL.INIT procedure runs the Hadamard test circuit $2S$ times to get the samples $\{(x^{(i)}, y^{(i)})\}_{i=1}^S$. Since the filter function f_T has spectrum bounded in $[-T, T]$, the maximal evolution time is $2\pi T$ and the total evolution time is at most $4\pi ST$.

The CONVEVAL.EVAL procedure then uses the S samples to compute the estimate of $(f_T * p)(x)$. Moreover, the computation is classical and elementary.   □

4.2 Application to Gaussian derivative filters

In this section, we apply the data structure CONVEVAL to the band-limited Gaussain derivative filter $g_{\sigma,T}$:

$$g_{\sigma,T}(x) = \int_{-T}^{T} \hat{g}_\sigma(\xi) e^{2\pi \mathbf{i} x \xi} d\xi = 2\pi \mathbf{i} \int_{-T}^{T} \xi e^{-\frac{1}{2}(\sigma \pi \xi)^2 + 2\pi \mathbf{i} x \xi} d\xi. \tag{61}$$

To apply Lemma 4.2, we first bound the L_1-norm of its spectrum.

Claim 4.6. Let $g_{\sigma,T}$ be defined as Eq. (61). Then we have $\|\hat{g}_{\sigma,T}\|_1 \leq \frac{4}{\pi \sigma^2}$.

Proof. By the fact that $\hat{g}_{\sigma,T}(\xi) = \hat{g}_\sigma(\xi) \mathbf{1}_{|\xi| \leq T}$ and direct calculation, we obtain

$$\|\hat{g}_{\sigma,T}\|_1 \leq \|\hat{g}_\sigma\|_1 = \int_{-\infty}^{+\infty} \left| 2\pi \mathbf{i} \xi e^{-\frac{1}{2}(\sigma \pi \xi)^2} \right| d\xi = \int_0^{+\infty} 4\pi \xi e^{-\frac{1}{2}(\sigma \pi \xi)^2} d\xi = \frac{4}{\pi \sigma^2}. \tag{62}$$

$\square$

Then we get the following corollary on the sample complexity of evaluating $g_{\sigma,T} * p$ on M points.

Corollary 4.7. Let $\epsilon_1 > 0$, $\delta_1 \in (0,1)$ and $x_1, x_2, \ldots, x_M \in \mathbb{R}$ be arbitrary. Suppose the data structure CONVEVAL is initialized with parameters $(g_{\sigma,T}, \epsilon_1, \delta_1, M)$. Let h_j be the output of the procedure CONVEVAL.EVAL(x_j) for $j \in [M]$. Then we have

$$\mathbb{P}\left[\forall j \in [M] : |(g_{\sigma,T} * p)(x_j) - h_j| \leq \epsilon_1\right] \geq 1 - \delta_1. \tag{63}$$

Furthermore, it take $S = \mathcal{O}(\epsilon_1^{-2} \sigma^{-4} \log(M/\delta_1))$ samples from Hadamard tests to obtain $h_1, h_2, \ldots, h_M$.

Proof. Claim 4.6 implies $\|\hat{g}_{\sigma,T}\|_1 = \mathcal{O}(\sigma^{-2})$. Thus the procedure CONVEVAL.INIT$(g_{\sigma,T}, \epsilon_1, \delta_1, M)$ draws $S = \mathcal{O}(\epsilon_1^{-2} \|\hat{g}_{\sigma,T}\|_1^2 \log(M/\delta_1)) = \mathcal{O}(\epsilon_1^{-2} \sigma^{-4} \log(M/\delta_1))$ samples from Hadamard tests. Then Eq. (63) follows immediately from Corollary 4.4. $\square$

5 Main Theorem

In this section, we describe our main results about ground state energy estimation. We first present a $\widetilde{\mathcal{O}}(\Delta^{-1})$-depth algorithm for GSEE in Algorithm 2.

Then we prove the following theorem:

Theorem 5.1 (Ground state energy estimation). Let $H = \sum_{j=0}^{N-1} E_j |E_j\rangle\langle E_j|$ be a Hamiltonian such that $E_0 < E_1 \leq E_2 \leq \cdots \leq E_{N-1}$ are the eigenvalues of H, and the $|E_j\rangle$'s are orthonormal eigenstates of H. Suppose we are given access to the Hamiltonian evolution $e^{\mathbf{i}Ht}$ for any $t \in \mathbb{R}$. Let $\Delta > 0$ be given such that $\Delta \leq E_1 - E_0$. Moreover, suppose we can prepare a state ρ such that $\langle E_0| \rho |E_0\rangle \geq \eta$ for known $\eta > 0$.

Then, for sufficiently small $\epsilon > 0$ and any $\delta \in (0,1)$, there exists an algorithm that estimates E_0 within accuracy ϵ with probability at least $1 - \delta$ such that:

- The maximal Hamiltonian evolution time is $\widetilde{\mathcal{O}}(\Delta^{-1})$;

- The total Hamiltonian evolution time is $\widetilde{\mathcal{O}}(\eta^{-2}\epsilon^{-2}\Delta)$;

- The classical running time is $\widetilde{\mathcal{O}}(\eta^{-2}\epsilon^{-3}\Delta^3)$.

Algorithm 2 Low-depth ground state energy estimation algorithm.

1: **procedure** GSEE(ϵ, δ, $\widetilde{E}_0$, Δ, η)

2: $\quad \sigma \leftarrow \min\left(\dfrac{0.9\Delta}{\sqrt{2\ln(9\Delta\epsilon^{-1}\eta^{-1})}}, 0.2\Delta \right)$

3: $\quad M \leftarrow \lceil \sigma/\epsilon \rceil + 1, \ \tilde{\epsilon} \leftarrow \dfrac{0.1\epsilon\eta}{\sqrt{2\pi}\sigma^3}$

4: $\quad T \leftarrow \pi^{-1}\sigma^{-1}\sqrt{2\ln\left(8\pi^{-1}\tilde{\epsilon}^{-1}\sigma^{-2}\right)}$ $\qquad\qquad\qquad$ ▷ Filter band-limit (Lemma 3.4)

5: $\quad$ CONVEVAL.INIT($g_{\sigma,T}$, $\tilde{\epsilon}/2$, $\delta/2$, M) $\qquad\qquad\qquad\qquad\qquad\qquad$ ▷ Algorithm 1

6: $\quad$ **for** $j = 1, 2, \ldots, M$ **do**

7: $\quad\quad x_j \leftarrow \widetilde{E}_0 - 0.25\sigma + (0.5\sigma/M) \cdot (j-1)$

8: $\quad\quad h_j \leftarrow$ CONVEVAL.EVAL(x_j) $\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ ▷ Algorithm 1

9: $\quad$ **end for**

10: $\quad j^* \leftarrow \arg\min_{1 \leq j \leq M} |h_j|$.

11: $\quad$ **return** x_{j^*}

12: **end procedure**

Proof. **Algorithm:** Suppose $\epsilon > 0$ is small enough such that it satisfies the condition in Lemma 3.2. We first run the algorithm in [17] to obtain a coarse estimate $\widetilde{E}_0$ of E_0 such that $\widetilde{E}_0$ is $\sigma/4$-close to E_0 with probability at least $1 - \delta/2$, where σ is defined as Eq. (15). Then we run Algorithm 2 with parameters $(\epsilon, \delta, \widetilde{E}_0, \eta)$ and return its output x_{j^*} as the final estimate of E_0.

Correctness: By construction, $\widetilde{E}_0$ satisfies Eq. (27) in Lemma 3.3. Meanwhile, by Claim 3.5 and the choice of T in Algorithm 2, we have

$$|(g_\sigma * p)(x) - (g_{\sigma,T} * p)(x)| \leq \frac{\tilde{\epsilon}}{2}, \quad \forall x \in \mathbb{R}, \tag{64}$$

Meanwhile, since CONVEVAL is initialized with parameters $(g_{\sigma,T}, \tilde{\epsilon}/2, \delta/2, M)$ in Algorithm 2, by Corollary 4.7, we get

$$\mathbb{P}\left[\forall j \in [M] : |(g_{\sigma,T} * p)(x_j) - h_j| \leq \frac{\tilde{\epsilon}}{2} \right] \geq 1 - \frac{\delta}{2}. \tag{65}$$

Then it follows from Eqs. (64) and (65) and the triangle inequality that

$$\mathbb{P}\left[\forall j \in [M] : |(g_\sigma * p)(x_j) - h_j| \leq \tilde{\epsilon} \right] \geq 1 - \frac{\delta}{2}, \tag{66}$$

which coincides with Eq. (28) in Lemma 3.3, given the choice of $\tilde{\epsilon}$ in Algorithm 2. Now with both of its conditions met, Lemma 3.3 implies that the output of Algorithm 2, i.e., x_{j^*}, is ϵ-close to E_0 with probability at least $1 - \delta$, as desired.

Cost analysis: First, we run the algorithm in [17] to obtain $\widetilde{E}_0$. Since $\sigma = \widetilde{\Omega}(\Delta)$, this step has maximal evolution time $\widetilde{\mathcal{O}}(\Delta^{-1})$, total evolution time $\widetilde{\mathcal{O}}(\Delta^{-1}\eta^{-2})$, and classical post-processing time $\widetilde{\mathcal{O}}(\Delta^{-1}\eta^{-2})$.

Then, in Line 5 of Algorithm 2, we run CONVEVAL.INIT($g_{\sigma,T}$, $\tilde{\epsilon}/2$, $\delta/2$, M) to initialize the data structure CONVEVAL in Algorithm 1. We choose the parameters as follows:

- $\tilde{\epsilon} = \Omega(\epsilon\eta\sigma^{-3}) = \widetilde{\Omega}(\epsilon\eta\Delta^{-3})$,

- $T = \widetilde{\mathcal{O}}(\sigma^{-1}) = \widetilde{\mathcal{O}}(\Delta^{-1})$,

- $M = \Theta(\sigma\epsilon^{-1}) = \widetilde{\Theta}(\Delta\epsilon^{-1})$.

23

Thus, by Corollary 4.7, we have

$$S = \widetilde{\Theta}(\epsilon_1^{-2}\sigma^{-4}) = \widetilde{\Theta}(\epsilon^{-2}\eta^{-2}\Delta^6 \cdot \Delta^{-4}) = \widetilde{\Theta}(\epsilon^{-2}\eta^{-2}\Delta^2). \tag{67}$$

The for-loop in Procedure CONVEVAL.INIT of Algorithm 1 draws S samples from the Hadamard test circuit. The sampling process has the maximal evolution time $2\pi T = \widetilde{\mathcal{O}}(\Delta^{-1})$ and total evolution time at most $\mathcal{O}(TS) = \widetilde{\mathcal{O}}(\epsilon^{-2}\eta^{-2}\Delta)$.

Next, in Line 6 of Algorithm 2, we call the procedure CONVEVAL.EVAL M times to evaluate the convolutions at $x_1, x_2, \ldots, x_M$. Each evaluation takes $\mathcal{O}(S) = \widetilde{\mathcal{O}}(\epsilon^{-2}\eta^{-2}\Delta^2)$ classical time. Hence, this step takes $\mathcal{O}(MS) = \widetilde{\mathcal{O}}(\Delta\epsilon^{-1} \cdot \epsilon^{-2}\eta^{-2}\Delta^2) = \widetilde{\mathcal{O}}(\epsilon^{-3}\eta^{-2}\Delta^3)$ classical time.

Combining these steps together, we get that the whole GSEE algorithm takes:

- maximal evolution time $\widetilde{\mathcal{O}}(\Delta^{-1})$,

- total evolution time $\widetilde{\mathcal{O}}(\Delta^{-1}\eta^{-2} + \epsilon^{-2}\eta^{-2}\Delta) = \widetilde{\mathcal{O}}(\epsilon^{-2}\eta^{-2}\Delta)$, and

- classical post-processing time $\widetilde{\mathcal{O}}(\epsilon^{-3}\eta^{-2}\Delta^3)$,

as claimed. □

As described in the introduction, it is favorable to be able to reduce the total evolution time at the cost of increased maximal evolution time (or circuit depth). This allows one to make the most use of the available circuit depth afforded by the quantum architecture. Such a feature is desirable in the era of early fault-tolerant quantum computing where there is likely to be a limit to the available coherence of the device [30]. Fortunately, this feature follows directly from the above theorem and we present it as a corollary below. Note that in Theorem 5.1, Δ is merely a lower bound on the true spectral gap $\Delta_{\text{true}} := E_1 - E_0$ of Hamiltonian H, not necessarily Δ_{true} itself. In fact, Δ can range from $\widetilde{\mathcal{O}}(\epsilon)$ (in order to satisfy the condition in Lemma 3.2) to Δ_{true}. By setting $\Delta = \widetilde{\mathcal{O}}(\epsilon^\alpha\Delta_{\text{true}}^{1-\alpha})$ with $\alpha \in [0, 1]$, we obtain:

Corollary 5.2. *Let* $H = \sum_{j=0}^{N-1} E_j|E_j\rangle\langle E_j|$ *be a Hamiltonian such that* $E_0 < E_1 \leq E_2 \leq \cdots \leq E_{N-1}$ *are the eigenvalues of* H, *and the* $|E_j\rangle$*'s are orthonormal eigenstates of* H. *Let* $\Delta_{\text{true}} = E_1 - E_0$ *be the spectral gap of* H. *Suppose we are given access to the Hamiltonian evolution* $e^{\mathbf{i}Ht}$ *for any* $t \in \mathbb{R}$. *Moreover, suppose we can prepare a state* ρ *such that* $\langle E_0|\rho|E_0\rangle \geq \eta$ *for known* $\eta > 0$.

Then for any $\alpha \in [0, 1]$, *for sufficiently small* $\epsilon > 0$ *and any* $\delta \in (0, 1)$, *there exists an algorithm that estimates* E_0 *within accuracy* ϵ *with probability at least* $1 - \delta$ *such that:*

- *The maximal Hamiltonian evolution time is* $\widetilde{\mathcal{O}}(\epsilon^{-\alpha}\Delta_{\text{true}}^{-1+\alpha})$;

- *The total Hamiltonian evolution time is* $\widetilde{\mathcal{O}}(\eta^{-2}\epsilon^{-2+\alpha}\Delta_{\text{true}}^{1-\alpha})$;

- *The classical running time is* $\widetilde{\mathcal{O}}(\eta^{-2}\epsilon^{-3+\alpha}\Delta_{\text{true}}^{3-\alpha})$.

In particular, setting $\alpha = 0$ *or* 1 *leads to:*

- $\Delta = \Delta_{\text{true}}$, *for which Theorem 5.1 yields an algorithm with maximal evolution time* $\widetilde{\mathcal{O}}(\Delta_{\text{true}}^{-1})$ *and total evolution time* $\widetilde{\mathcal{O}}(\eta^{-2}\epsilon^{-2}\Delta_{\text{true}})$; *or*

- $\Delta = \widetilde{\mathcal{O}}(\epsilon)$, *for which Theorem 5.1 yields an algorithm with maximal evolution time* $\widetilde{\mathcal{O}}(\epsilon^{-1})$ *and total evolution time* $\widetilde{\mathcal{O}}(\eta^{-2}\epsilon^{-1})$ *(i.e., the Heisenberg limit).*

For general $\Delta = \widetilde{\mathcal{O}}(\epsilon^\alpha\Delta_{\text{true}}^{1-\alpha})$ with $\alpha \in [0, 1]$, Theorem 5.1 yields an algorithm with maximal evolution time $\widetilde{\mathcal{O}}(\epsilon^{-\alpha}\Delta_{\text{true}}^{-1+\alpha})$ and total evolution time $\widetilde{\mathcal{O}}(\eta^{-2}\epsilon^{-2+\alpha}\Delta_{\text{true}}^{1-\alpha})$. In other words, tuning Δ between the two extremes gives a trade-off between the circuit depth and total runtime of the algorithm.

24

6 Discussion and Outlook

In this work we have introduced a new quantum algorithm (Algorithm 2) for estimating the ground state energy of a Hamiltonian. The main innovation is that the maximum circuit depth required by this algorithm is governed by the gap of the Hamiltonian rather than the target accuracy. Hamiltonians describing chemical systems typically have gaps that are several orders of magnitude larger than chemical accuracy (the standard target accuracy in quantum chemistry). Accordingly, the algorithm developed in this work can reduce the required circuit depths of ground state energy estimation by several orders of magnitude. This lowers the quantum resources required for solving industrially-relevant quantum chemistry problems, which may enable quantum advantage using early fault-tolerant quantum computers. The cost of reducing the circuit depth is an increase in the overall runtime of the algorithm. As with other recent ground state energy estimation algorithms [17], the algorithm is "embarrassingly parallel"; the runtime can be reduced by parallelizing the sample gathering accross multiple quantum computers (trading time resources for space resources). As shown in Corollary 5.2, as the quantum computer is able to implement deeper quantum circuits (say, through improvements in scaling the fault-tolerant quantum architecture), we are able to convert this available coherence into a runtime reduction. This ability to trade circuit depth for runtime is an important feature of algorithms designed for early fault-tolerant quantum computing [31, 32, 33, 18].

Beyond direct application of the algorithm for achieving quantum advantage, our work helps to establish the paradigm of developing quantum algorithms using the tools of classical signal processing [17, 34, 33]. The quantum computer generates a stochastic signal that encodes properties of a matrix of interest. This stochastic signal can be processed to learn the matrix properties of interest. This paradigm will be essential in the development of algorithms for early fault-tolerant quantum computers. Quantum computations with such architectures will be error prone, generating noisy signals. The tools of classical signal processing have been designed to handle such noisy signals and can aid in the design and analysis of robust quantum algorithms [32, 35].

We now discuss several directions for future work. One requirement of the algorithm is that a lower bound on the energy gap must be specified. There exist quantum chemistry methods for estimating the gap (e.g. using the ORCA software [25, 26] as we did for our numerical comparisons). However, such estimates can become inaccurate for large systems. It may be helpful to incorporate a step into the quantum algorithm that estimates this gap. Although this estimation is computationally hard in general [36], many physical systems of interest have structure that make the estimation feasible.

In addition to estimating the gap, it may be possible to increase the size of an "effective gap" (i.e. the energy separation between the ground state and the lowest eigenvalue for which the input state has significant overlap with its eigenstate). One approach would be to exploit known symmetries in the Hamiltonian to prepare an input state that realizes a larger effective gap. Another approach would be to apply state preparation boosters [33] that filter out low-lying excited states. Such approaches would help to reduce the circuit depth required for energy estimation (by increasing the effective gap), but may increase the circuit depth cost of state preparation. Further work is needed to investigate such approaches and understand the trade-offs.

As with all ground state energy estimation algorithms, the performance of our algorithm depends on the overlap of the input state with the ground state (see Table I in [9]). A concern with using quantum computers to solve problems in quantum chemistry is that for large systems, we are unable to prepare sufficiently good ground state approximations [37]. This is known as the orthogonality catastrophe [38] or the Van Vleck catastrophe [39, 40]. An important direction for future work is the development of well-motivated quantum heuristics for approximating ground states.

In this work, we did not consider the impact of implementation error on the performance of the algorithm. We expect that our algorithm is able to tolerate some degree of variation between the ideal Hadamard tests and the implemented Hadamard tests. Furthermore, we believe the algorithm can be operated so as to accommodate such deviations. We leave for future work, the investigation of robust quantum algorithms for ground state energy estimation.

In this work we have introduced a framework for implementing ground state energy estimation using tunable-depth quantum circuits. As shown in Figure 1, the algorithms developed in this work are applicable to the maximum circuit depths ranging from $\widetilde{\mathcal{O}}(1/\Delta)$ to $\widetilde{\mathcal{O}}(1/\epsilon)$. We leave to future work the development of tunable-depth quantum algorithms outside of this region. While we have made progress in establishing upper bounds over a range of circuit depths, an important future direction is to establish lower bounds on depth-limited ground state energy estimation. These directions would further the research program of characterizing the performance and limitations of using depth-limited quantum computers to estimate properties of Hamiltonians.

The methods introduced here may help to bring the target of useful quantum computing closer to the present. Yet, there is still much work needed to carry out detailed resource estimations that predict the onset of quantum advantage using methods such as those we have introduced. More broadly, or hope is that this work contributes to the general understanding of how to use quantum computers given practical constraints on their capabilities and might inspire the development of quantum algorithms in other application domains.

A   Comparison to the approach of Lin et al

The main advantage of our approach compared to [17] is in the minimal evolution time required to achieve a desired precision. Indeed, in their approach the evolution time scales inverse linearly with the desired precision. For our approach, the minimal evolution time is dictated by the reciprocal of the energy gap of the Hamiltonian and any additional precision we wish to attain only causes a poly-logarithmic factor in the evolution time. Of course, this comes at the expense of a higher sample complexity at smaller evolution times. This trade-off between the evolution time and the sample complexity is discussed in Corollary 5.2.

Note that this improvement in the minimal evolution time comes from two conceptual differences in our approach compared to [17]: the choice of the filter function (Heaviside versus Gaussian derivative) and how we then infer the value of the ground state energy from the convolution (jump versus 0 of derivative).

Both our approach and that of [17] require a truncated approximation of the underlying filter function to implement the algorithm with only finite-time evolutions. However, as the Heaviside function has a jump at 0, the degree of the Fourier series necessarily has to increase the better we want the approximation outside a small neighborhood of 0 to be. For instance, in [17] they find an

approximation $F_{d,\epsilon}$ such that for $d = \mathcal{O}(\epsilon^{-1} \log (\epsilon^{-1} \delta^{-1}))$ and

$$F_{d,\epsilon}(x) = \frac{1}{\sqrt{2\pi}} \sum_{k=-d}^{d} \hat{F}_{d,\epsilon,k} e^{ikx}, \tag{68}$$

we have[4]

1. $-\frac{\delta}{2} \leq F_{d,\epsilon}(x) \leq 1 + \delta$ for all $x \in \mathbb{R}$.

2. $|F_{d,\epsilon}(x) - \Theta(x)| \leq \delta$ for all $x \in [-\pi + \epsilon, -\epsilon] \cup [\epsilon, \pi - \epsilon]$.

Note that the evolution time required to implement the representation in Eq. (68) is $\mathcal{O}(d)$. This scales logarithmically with the precision with which we approximate the Heaviside function outside of the intervals $[-\pi + \epsilon, -\epsilon] \cup [\epsilon, \pi - \epsilon]$ and inverse-polynomially in size of the interval around 0, $(-\epsilon, \epsilon)$, where we are not guaranteed that the two functions are close.

The approach of [17] consists of finding the smallest point x such that $(p * F_{d,\epsilon})(x) \geq \eta$. If we were convolving with the Heaviside function, this would correspond to the ground state energy. But we will now argue that the neighborhood around 0 for which we have the approximation can shift where the jump occurs. Indeed, note that for $x \in [E_0 - \epsilon, E_0 + \epsilon]$ and $\epsilon \leq \Delta$ we have that:

$$(p * F_{d,\epsilon})(x) = (p * \Theta)(x) + \int_{-\epsilon}^{\epsilon} p(x-y)(F_{d,\epsilon}(y) - \Theta(y))dy + \int_{[-1,1]\setminus[-\epsilon,\epsilon]} p(x-y)(F_{d,\epsilon}(y) - \Theta(y))dy$$

Note that as $|F_{d,\epsilon}(x) - \Theta(x)| \leq \delta$ for all $x \in [-\pi + \epsilon, -\epsilon] \cup [\epsilon, \pi - \epsilon]$ we have that

$$\left| \int_{[-1,1]\setminus[-\epsilon,\epsilon]} p(x-y)(F_{d,\epsilon}(y) - \Theta(y))dy \right| \leq \delta. \tag{69}$$

However, as we only have the promise that $-\frac{\delta}{2} \leq F_{d,\epsilon}(x) \leq 1 + \delta$ for points in $[-\epsilon, \epsilon]$, we will not be able to infer the precise point of the jump at a precision larger than $\mathcal{O}(\epsilon)$ with this approach. This is because the residual integral term (i.e. the one over $(-\epsilon, \epsilon)$ in Eq. (69)) will cause fluctuations in this interval and we will not be able to pin down the jump.

On the other hand, by choosing our filter to be Gaussian derivatives, we are able to obtain a good approximation everywhere on the real line. Furthermore, by choosing the zeros of the derivative as criteria, we only need to make sure that the standard deviation is small enough to separate different eigenvalues. This way we obtain a smaller maximal evolution time.

B Results about Gaussian filters

In this Appendix, we analyze the cost of evaluating the convolution of the spectral measure $p(x) = \sum_{j=0}^{N-1} p_j \delta(x - E_j)$ and a Gaussian filter. Although our GSEE algorithm is based on Gaussian derivative convolution instead of Gaussian convolution, we believe that the latter is still useful for many tasks. Hence it is worth mentioning the following results.

Recall that for a Gaussian probability distribution function (pdf) f_σ, defined as

$$f_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{x^2}{\sigma^2}}, \tag{70}$$

---

[4]note that [17] adopts different notations. There δ is the precision with which we approximate the ground state energy, and here it is ϵ.

we have that its Fourier transform $\hat{f}_\sigma$ is given by

$$\hat{f}_\sigma(t) = e^{-\frac{1}{2}(\sigma\pi t)^2}. \tag{71}$$

That is, it just corresponds to the pdf of another Gaussian random variable up to a $\sqrt{\frac{\pi}{2}}\sigma$ factor.

Our first step is to define a truncated version of the pdf f_σ stemming from the Fourier transform. Given a parameter $T > 0$, we define $f_{\sigma,T}$ to be given by:

$$f_{\sigma,T}(x) := \int_{-T}^{T} \hat{f}_\sigma(t)e^{2\pi\mathbf{i}xt}\mathrm{d}t. \tag{72}$$

Clearly, as $T \to \infty$ we have that $f_{\sigma,T}(x)$ converges to $f_\sigma(x)$. The following lemma shows how to choose T such that the $f_{\sigma,T}$ can approximate f_σ in L_∞-norm:

Lemma B.1. *Let $\epsilon > 0$ be arbitrary. Then for*

$$T := \sigma^{-1}\sqrt{\ln\left(\epsilon^{-1}\sigma^{-1}\right)}, \tag{73}$$

we have that

$$|f_\sigma(x) - f_{\sigma,T}(x)| \leq \frac{\epsilon}{2}, \quad \forall x \in \mathbb{R}. \tag{74}$$

Proof. By the Fourier inversion theorem we have

$$\begin{aligned}
|f_\sigma(x) - f_{\sigma,T}(x)| &= \left| \int_{-\infty}^{-T} \hat{f}(t)e^{2\pi\mathbf{i}xt}\mathrm{d}t + \int_{T}^{+\infty} \hat{f}(t)e^{2\pi\mathbf{i}xt}\mathrm{d}t \right| \\
&\leq \int_{-\infty}^{-T} |\hat{f}(t)|\mathrm{d}t + \int_{T}^{+\infty} |\hat{f}(t)|\mathrm{d}t \\
&= 2\int_{T}^{+\infty} e^{-\frac{1}{2}(\sigma\pi t)^2}\mathrm{d}t \\
&= \sqrt{\frac{2}{\pi}}\sigma^{-1} \cdot \mathbb{P}_{X\sim\mathcal{N}(0,(\pi\sigma)^{-2})}[|X| > T] \\
&\leq \sqrt{\frac{2}{\pi}}\sigma^{-1} \cdot 2e^{-T^2\pi^2\sigma^2/2},
\end{aligned}$$

where the third step follows from Eq. (71), and the forth step follows from the concentration of Gaussian random variables.

By solving the following inequality:

$$\sqrt{\frac{2}{\pi}}\sigma^{-1} \cdot 2e^{-T^2\pi^2\sigma^2/2} \leq \frac{\epsilon}{2},$$

we get that it suffices to take:

$$T \geq \sigma^{-1}\sqrt{\ln\left(\sigma^{-1}\epsilon^{-1}\right)}. \tag{75}$$

$\square$

The L_∞-approximation for f_σ implies the L_∞-approximation for $f_\sigma * p$. That is, letting T be defined as in Lemma B.1, we have that

$$|(f_\sigma * p)(x) - (f_{\sigma,T} * p)(x)| \leq \frac{\epsilon}{2}, \quad \forall x \in \mathbb{R}. \tag{76}$$

Therefore, in order to estimate $(f_\sigma * p)(x)$ within ϵ-accuracy, it suffices to evaluate $(f_{\sigma,T} * p)(x)$ within 0.5ϵ-accuracy, which can be achieved by applying the data structure CONVEVAL to the band-limited Gaussain filter $f_{\sigma,T}$.

Specifically, to apply Lemma 4.2 to $f_{\sigma,T}$, we first bound the L_1-norm of its spectrum $\hat{f}_{\sigma,T}$ as follows:

$$\left\|\hat{f}_{\sigma,T}\right\|_1 = \int_{-T}^{T} e^{-\frac{1}{2}(\sigma\pi t)^2} dt \leq \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\sigma\pi t)^2} dt = \sqrt{\frac{2}{\pi}} \frac{1}{\sigma}. \tag{77}$$

Then by Corollary 4.4 and Lemma 4.5, we get the following result on the cost of evaluating $f_{\sigma,T} * p$ on M points:

Corollary B.2. *Let $\epsilon > 0$, $\delta \in (0,1)$ and $x_1, x_2, \ldots, x_M \in \mathbb{R}$ be arbitrary. Suppose the data structure CONVEVAL is initialized with parameters $(f_{\sigma,T}, \epsilon/2, \delta, M)$. Let h_j be the output of the procedure CONVEVAL.EVAL(x_j) for $j \in [M]$. Then we have*

$$\mathbb{P}\left[\forall j \in [M] : |(f_{\sigma,T} * p)(x_j) - h_j| \leq \frac{\epsilon}{2}\right] \geq 1 - \delta. \tag{78}$$

Furthermore, it take $S = \mathcal{O}(\epsilon^{-2}\sigma^{-2} \log(M/\delta))$ samples from Hadamard tests to obtain h_1, h_2, $\ldots$, h_M. Moreover, the maximal evolution time in these Hadamard tests is $\widetilde{\mathcal{O}}(\sigma^{-1})$.

Combining Eqs. (76) and (78) and using the triangle inequality, we obtain that

$$\mathbb{P}\left[\forall j \in [M] : |(f_\sigma * p)(x_j) - h_j| \leq \epsilon\right] \geq 1 - \delta. \tag{79}$$

That is, the h_j's in Corollary B.2 are accurate estimates of the $(f_\sigma * p)(x_j)$'s with high probability.

References

[1] Henrik R Larsson, Huanchen Zhai, Cyrus J Umrigar, and Garnet Kin Chan. The chromium dimer: closing a chapter of quantum chemistry. *arXiv preprint arXiv:2206.10738*, 2022.

[2] 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.

[3] 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.

[4] Isaac H Kim, Ye-Hua Liu, Sam Pallister, William Pol, Sam Roberts, and Eunseok Lee. Fault-tolerant resource estimate for quantum chemical simulations: Case study on li-ion battery electrolyte molecules. *Physical Review Research*, 4(2):023019, 2022.

[5] Alain Delgado, Pablo AM Casares, Roberto dos Reis, Modjtaba Shokrian Zini, Roberto Campos, Norge Cruz-Hernández, Arne-Christian Voigt, Angus Lowe, Soran Jahangiri, MA Martin-Delgado, et al. How to simulate key properties of lithium-ion batteries with a fault-tolerant quantum computer. *arXiv preprint arXiv:2204.11890*, 2022.

[6] Joshua J Goings, Alec White, Joonho Lee, Christofer S Tautermann, Matthias Degroote, Craig Gidney, Toru Shiozaki, Ryan Babbush, and Nicholas C Rubin. Reliably assessing the electronic structure of cytochrome p450 on today's classical computers and tomorrow's quantum computers. *arXiv preprint arXiv:2202.01244*, 2022.

[7] Jay Gambetta. Ibm quantum roadmap to build quantum-centric supercomputers, Aug 2022.

[8] Ryan Babbush, Craig Gidney, Dominic W Berry, Nathan Wiebe, Jarrod McClean, Alexandru Paler, Austin Fowler, and Hartmut Neven. Encoding electronic spectra in quantum circuits with linear t complexity. *Physical Review X*, 8(4):041015, 2018.

[9] Yulong Dong, Lin Lin, and Yu Tong. Ground state preparation and energy estimation on early fault-tolerant quantum computers via quantum eigenvalue transformation of unitary matrices. *arXiv preprint arXiv:2204.05955*, 2022.

[10] Dominic W Berry, Brendon L Higgins, Stephen D Bartlett, Morgan W Mitchell, Geoff J Pryde, and Howard M Wiseman. How to perform the most accurate possible phase measurements. *Physical Review A*, 80(5):052114, 2009.

[11] Emanuel Knill, Gerardo Ortiz, and Rolando D Somma. Optimal quantum measurements of expectation values of observables. *Physical Review A*, 75(1):012328, 2007.

[12] 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.

[13] Peter D Johnson, Alexander A Kunitsa, Jérôme F Gonthier, Maxwell D Radin, Corneliu Buda, Eric J Doskocil, Clena M Abuan, and Jhonathan Romero. Reducing the cost of energy estimation in the variational quantum eigensolver algorithm with robust amplitude estimation. *arXiv preprint arXiv:2203.07275*, 2022.

[14] Jarrod R McClean, Sergio Boixo, Vadim N Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature communications*, 9(1):1–6, 2018.

[15] Eric R Anschuetz and Bobak T Kiani. Beyond barren plateaus: Quantum variational algorithms are swamped with traps. *arXiv preprint arXiv:2205.05786*, 2022.

[16] Thomas E O'Brien, Brian Tarasinski, and Barbara M Terhal. Quantum phase estimation of multiple eigenvalues for small-scale (noisy) experiments. *New Journal of Physics*, 21(2):023022, 2019.

[17] Lin Lin and Yu Tong. Heisenberg-limited ground-state energy estimation for early fault-tolerant quantum computers. *PRX Quantum*, 3(1):010318, 2022.

[18] Kianna Wan, Mario Berta, and Earl T Campbell. A randomized quantum algorithm for statistical phase estimation. *arXiv preprint arXiv:2110.12071*, 2021.

[19] Earl Campbell. Random compiler for fast hamiltonian simulation. *Physical review letters*, 123(7):070503, 2019.

[20] Guang Hao Low and Isaac L. Chuang. Optimal hamiltonian simulation by quantum signal processing. *Phys. Rev. Lett.*, 118:010501, Jan 2017.

[21] John M. Martyn, Zane M. Rossi, Andrew K. Tan, and Isaac L. Chuang. Grand unification of quantum algorithms. *PRX Quantum*, 2:040203, Dec 2021.

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

[23] Ian D Kivlichan, Jarrod McClean, Nathan Wiebe, Craig Gidney, Alán Aspuru-Guzik, Garnet Kin-Lic Chan, and Ryan Babbush. Quantum simulation of electronic structure with linear depth and connectivity. *Physical review letters*, 120(11):110501, 2018.

[24] Pierre-François Loos, Filippo Lipparini, Martial Boggio-Pasqua, Anthony Scemama, and Denis Jacquemin. A mountaineering strategy to excited states: Highly accurate energies and benchmarks for medium sized molecules. *Journal of Chemical Theory and Computation*, 16(3):1711–1741, 2020.

[25] Frank Neese. The orca program system. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(1):73–78, 2012.

[26] Frank Neese. Software update: the orca program system, version 4.0. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 8(1):e1327, 2018.

[27] Alexandre Eremenko and Peter Yuditskii. Uniform approximation of sgn(x) by polynomials and entire functions. *Journal d Analyse Mathématique*, 101:313–324, 05 2007.

[28] Michael James David Powell et al. *Approximation theory and methods*. Cambridge university press, 1981.

[29] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, March 1963.

[30] Yu Tong. Designing algorithms for estimating ground state properties on early fault-tolerant quantum computers. *Quantum Views*, 6:65, 2022.

[31] Daochen Wang, Oscar Higgott, and Stephen Brierley. Accelerated variational quantum eigensolver. *Physical review letters*, 122(14):140504, 2019.

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

[33] Guoming Wang, Sukin Sim, and Peter D. Johnson. State preparation boosters for early fault-tolerant quantum computation, 2022.

[34] Ruizhe Zhang, Guoming Wang, and Peter Johnson. Computing ground state properties with early fault-tolerant quantum computers. *Quantum*, 6:761, 2022.

[35] Amara Katabarwa, Alex Kunitsa, Borja Peropadre, and Peter Johnson. Reducing runtime and error in vqe using deeper and noisier quantum circuits. *arXiv preprint arXiv:2110.10664*, 2021.

[36] 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.

[37] Seunghoon Lee, Joonho Lee, Huanchen Zhai, Yu Tong, Alexander M Dalzell, Ashutosh Kumar, Phillip Helms, Johnnie Gray, Zhi-Hao Cui, Wenyuan Liu, et al. Is there evidence for exponential quantum advantage in quantum chemistry? *arXiv preprint arXiv:2208.02199*, 2022.

[38] Norm M Tubman, Carlos Mejuto-Zaera, Jeffrey M Epstein, Diptarka Hait, Daniel S Levine, William Huggins, Zhang Jiang, Jarrod R McClean, Ryan Babbush, Martin Head-Gordon, et al. Postponing the orthogonality catastrophe: efficient state preparation for electronic structure simulations on quantum devices. *arXiv preprint arXiv:1809.05523*, 2018.

[39] JH Van Vleck. Nonorthogonality and ferromagnetism. *Physical Review*, 49(3):232, 1936.

[40] Jarrod R McClean, Ryan Babbush, Peter J Love, and Alán Aspuru-Guzik. Exploiting locality in quantum computation for quantum chemistry. *The journal of physical chemistry letters*, 5(24):4368–4380, 2014.