Faster ground state energy estimation on early fault-tolerant quantum computers via rejection sampling

Guoming Wang, Daniel Stilck França, Gumaro Rendon, Peter Johnson

Faster ground state energy estimation on early fault-tolerant quantum computers via rejection sampling

Guoming Wang[1], Daniel Stilck França[2], Gumaro Rendon[3], and Peter D. Johnson[3]

[1]Zapata Computing Canada Inc., Toronto, ON M5C 3A1, Canada
[2]Univ Lyon, ENS Lyon, UCBL, CNRS, Inria, LIP, F-69342, Lyon Cedex 07, France
[3]Zapata Computing Inc., Boston, MA 02110 USA

Abstract

A major thrust in quantum algorithm development over the past decade has been the search for the quantum algorithms that will deliver practical quantum advantage first. Today's quantum computers and even early fault-tolerant quantum computers will be limited in the number of operations they can implement per circuit. We introduce quantum algorithms for ground state energy estimation (GSEE) that accommodate this design constraint. The first estimates ground state energies and has a quadratic improvement on the ground state overlap parameter compared to other methods in this regime. The second certifies that the estimated ground state energy is within a specified error tolerance of the true ground state energy, addressing the issue of gap estimation that beleaguers several ground state preparation and energy estimation algorithms. We note, however, that the scaling of this certification technique is, unfortunately, worse than that of the GSEE algorithm. These algorithms are based on a novel use of the quantum computer to facilitate rejection sampling. After a classical computer is used to draw samples, the quantum computer is used to accept or reject the samples. The set of accepted samples correspond to draws from a target distribution. While we use this technique for ground state energy estimation, it may find broader application. Our work pushes the boundaries of what operation-limited quantum computers are capable of and thus brings the target of quantum advantage closer to the present.

1 Introduction

Estimating the ground state energy of quantum many-body systems is a fundamental problem in condensed matter physics, quantum chemistry, material science and quantum information. It serves as a basic subroutine in drug discovery and materials design [1, 2, 3, 4]. In this problem, the system of interest is characterized as an n-qubit Hamiltonian H with unknown spectral decomposition $H = \sum_j E_j |E_j\rangle \langle E_j|$, where $E_0 < E_1 \leq E_2 \leq ...$ are the eigenvalues of H and the $|E_j\rangle$'s are the orthonormal eigenstates of H, and we need to estimate the lowest eigenvalue E_0 of H within additive error ϵ with high probability. Due to its significance, this problem has been extensively studied in the past decades, and various methods have been proposed to tackle it. Among these methods, two main approaches to ground state energy estimation (GSEE) have been explored.

GSEE assuming a fault-tolerance cost model The first approach to GSEE is based on the quantum phase estimation (QPE) algorithm [5, 6] and often has rigorous performance guarantees. In this approach, we assume that there exists an efficient procedure to prepare a state $|\psi\rangle$ that has

non-trivial overlap with the ground state of H (i.e. $|\langle E_0|\psi\rangle| = \Omega(1/\operatorname{poly}(n))$). This assumption is reasonable in many practical settings. For example, Refs. [7, 8, 9, 10] provide several methods for creating such states for quantum chemistry Hamiltonians. Given the "ansatz state" $|\psi\rangle$, we can perform standard QPE on the time evolution of H using $|\psi\rangle$ as the initial state. Assuming $p_0 \coloneqq |\langle E_0|\psi\rangle|^2$, we can obtain an ϵ-accurate estimate of E_0 by using $\mathcal{O}(p_0^{-1})$ QPE circuits, each of which uses $\mathcal{O}\left(\log\left(\epsilon^{-1}p_0^{-1}\right)\right)$ ancilla qubits and implements a time evolution of H for a duration at most $\mathcal{O}(\epsilon^{-1}p_0^{-1})$. Subsequently, the polynomial scaling of the algorithm runtime is $\mathcal{O}(\epsilon^{-1}p_0^{-2})$. Several variants of this basic algorithm have been proposed in which the overall evolution time is reduced to $\tilde{\mathcal{O}}(\epsilon^{-1}p_0^{-1})$ [11, 12, 13, 14], or the maximal evolution time is reduced to $\mathcal{O}\left(\epsilon^{-1}\log\left(p_0^{-1}\right)\right)$ [11, 12, 13, 15, 16, 14], or the number of ancilla qubits is reduced to $\mathcal{O}(1)$ [17, 18, 15, 14].

Despite these improvements, all of these algorithms require implementing a time evolution of H up to $\mathcal{O}(\epsilon^{-1})$ time. This implies that the number of operations per circuit on the quantum computer is quite large when the tolerable error ϵ is small. As an example, [19] estimated that the number of T gates required to implement ground state energy estimation to sufficient accuracy for a class of small molecules is greater than 10^{10}; the error rate of each gate must be significantly less than 10^{-10}. Today's quantum hardware and that of the near future are far more limited in the number of gates that can be implemented per circuit. As an example, IBM recently announced their 100×100 challenge [20], in which, by 2024, they aim to realize accurate expectation value calculations using circuits of $100 \times 100 = 10^4$ quantum gates, each of error rate 10^{-3}. There is large consensus in the community that the only feasible approach to increasing the number of operations from 10^4 to greater than 10^{10} is by building large-scale quantum architectures that can implement low-error-rate fault-tolerant operations.

The GSEE algorithms above were developed with a fault-tolerant quantum computer in mind. It is by now standard to assume a *cost model* where the overhead of reducing error rates (through fault-tolerant protocols) does not change the polynomial scaling of a quantum algorithm; that is, the cost is logarithmic in the number of operations. Under this model, it is usually favorable to reduce algorithm runtime at the cost of an increased number of operations per circuit. For this reason, methods for ground state energy estimation have been developed [21] that increase the number of ancilla qubits and operations to reduce the overall runtime.

GSEE assuming a pre-fault-tolerance cost model The second approach to GSEE is the variational quantum eigensolver (VQE) [22, 1, 23]. This approach was initially conceived assuming a cost model where fault-tolerant protocols are unavailable and thus gate error rates are fixed. For a given circuit error rate, this limits the maximum number of quantum operations per circuit. The method uses an ansatz circuit of limited number of operations to generate an approximation of the ground state of H. While this algorithm is amenable to near-term implementation, it lacks performance guarantees for two reasons. First, the accuracy of its output depends on the representation power of the chosen variational ansatz which is often hard to understand. Second, in order to find the optimal ansatz parameters, one needs to solve the associated non-convex optimization problem which could be challenging [24]. Furthermore, since VQE needs to repeatedly evaluate the energy of the ansatz state (via direct sampling or more sophisticated methods [25, 26]), this algorithm becomes too time-consuming to be practical for chemical system sizes of industrial relevance, as shown in a recent work [27] and are more sensitive to noise than previously expected [28, 29, 30, 31].

GSEE assuming a limited-number-of-operations cost model The two modes of ground state energy estimation algorithms introduced so far were developed assuming two very different cost models: logarithmic-overhead error reduction through fault-tolerance and fixed error rate.

These models correspond to future fault-tolerant architectures and today's near-term intermediate scale devices, respectively. These cost models may not be sufficient to describe the capabilities and limitations of early fault-tolerant quantum computers [32]. This motivates the development of cost models that interpolate between these two regimes. While rigorously developing such cost models is beyond the scope of our work, we will assume a proxy for such a model: *any early fault-tolerant quantum computer will be limited in the number of logical operations it can implement per circuit while keeping the total error rate below some amount.* This motivates the development of quantum algorithms in which the number of operations can be reduced below that allotted by the device, while paying a cost of an increase in runtime.

Several recent works have been developed under similar premises. The first was the so-called α-VQE method [33] in which a variable-depth amplitude estimation is introduced. Later, Wang et al. introduced a variable-depth amplitude estimation algorithm that is robust to substantial amounts of error [25]. Similar methods were explored in the context of quantum algorithms for finance [34, 35]. Some of these methods have been implemented on quantum hardware [36, 37]. Other algorithms have been introduced for early fault-tolerant quantum computers with the aim of reducing the number of ancilla qubits compared to their traditional counterparts [38, 15, 39].

Recently, Wang et al.[40] applied this methodology to the task of ground state energy estimation. They asked whether there exists a method for ground state energy estimation that uses fewer operations per circuit than previous approaches (i.e. the merit of VQE), while admitting a rigorous performance guarantee (i.e. the merit of traditional GSEE methods). Inspired by the recent work of Lin and Tong [15], they showed that, given a lower bound Δ on the spectral gap of H around E_0, one can estimate E_0 within additive error ϵ with high probability by using $\tilde{\mathcal{O}}\left(\Delta^2\epsilon^{-2}p_0^{-2}\right)$ quantum circuits, where each circuit evolves H up to $\mathcal{O}\left(\Delta^{-1}\operatorname{polylog}\left(\Delta\epsilon^{-1}p_0^{-1}\right)\right)$ time and requires a single ancilla qubit. Note that the maximal evolution time is only poly-logarithmic in $1/\epsilon$, which is exponentially better than previous methods. Furthermore, the parameter Δ can be chosen anywhere between $\Theta(\epsilon)$ and $\Theta(\Delta_{\text{true}})$, where $\Delta_{\text{true}} \coloneqq E_1 - E_0$, and tuning it leads to a class of GSEE algorithms that smoothly interpolate between an algorithm with circuit depth $\mathcal{O}\left(\Delta_{\text{true}}^{-1}\operatorname{polylog}\left(\Delta_{\text{true}}\epsilon^{-1}p_0^{-1}\right)\right)$ and runtime $\tilde{\mathcal{O}}\left(\Delta_{\text{true}}\epsilon^{-2}p_0^{-2}\right)$ and an algorithm with circuit depth $\mathcal{O}\left(\epsilon^{-1}\operatorname{polylog}\left(p_0^{-1}\right)\right)$ and runtime $\tilde{\mathcal{O}}(\epsilon^{-1}p_0^{-2})$. In other words, these algorithms smoothly transition from having low circuit depth to achieving Heisenberg-limit scaling in the runtime with respect to ϵ. Later, Ding and Lin [41] recovered this result using a different approach.

Although Refs. [40] and [41] significantly reduce the circuit depth for GSEE, one aspect of their algorithms is somewhat unsatisfactory — their runtimes are quadratic in $1/p_0$. This means that if the overlap between the input state $|\psi\rangle$ and the ground state $|E_0\rangle$ is small, these algorithm could take a long time to finish. Is it possible to improve the dependence on $1/p_0$ to linear while preserving the nice features of these algorithms? In this work, we answer this question in affirmative. Precisely, we prove that, given a lower bound Δ on the spectral gap of H around E_0, one can estimate E_0 within additive error ϵ with high probability by using $\tilde{\mathcal{O}}(\Delta^2\epsilon^{-2}p_0^{-1})$ quantum circuits, where each circuit has $\mathcal{O}\left(\Delta^{-1}\operatorname{polylog}\left(\Delta\epsilon^{-1}p_0^{-1}\right)\right)$ depth and requires $\mathcal{O}(1)$ ancilla qubits. This algorithm works with two common input models of H: 1. The block-encoding model, in which H is embedded as a sub-matrix of a larger unitary operator; 2. The Hamiltonian-evolution model, in which a controlled time evolution of H can be efficiently implemented. Moreover, as in Ref. [40], we can tune the parameter Δ anywhere between $\Theta(\epsilon)$ and $\Theta(\Delta_{\text{true}})$, and doing so yields a class of GSEE algorithms that smoothly transition from having circuit depth $\mathcal{O}\left(\Delta_{\text{true}}^{-1}\operatorname{polylog}\left(\Delta_{\text{true}}\epsilon^{-1}p_0^{-1}\right)\right)$ and runtime $\tilde{\mathcal{O}}\left(\Delta_{\text{true}}\epsilon^{-2}p_0^{-1}\right)$ to having circuit depth $\mathcal{O}(\epsilon^{-1}\operatorname{polylog}\left(p_0^{-1}\right))$ and runtime $\tilde{\mathcal{O}}\left(\epsilon^{-1}p_0^{-1}\right)$. Therefore, our algorithm is also able to trade circuit depth for runtime like the one in Ref. [40] but has a shorter runtime than the latter.

The core component of our algorithm is a novel procedure for performing rejection sampling on a quantum computer. We apply it to draw random values from a neighborhood of E_0. Specifically, suppose we want to draw a sample from a target Gaussian-like distribution $\nu(x)$ such that $\nu(x)$ peaks around E_0. It might be difficult to do so directly. Instead, we first draw a random value x from another distribution $\mu(x)$, and construct a unitary operator $U_x(H)$ such that $|\langle 0^k| U_x(H) |0^k\rangle |\psi\rangle|^2 \propto \frac{\nu(x)}{\mu(x)}$ for some integer k. Then we accept x if a measurement on the first k qubits of $U_x(H)|\psi\rangle$ yields outcome 0^k, and discard x otherwise. This procedure is akin to the standard rejection sampling algorithm.

We show that one can utilize quantum singular value transformation (QSVT) [42] or similar techniques [43, 14] to implement $U_x(H)$ with a low-depth circuit. Furthermore, the circuit only involves a few ancilla qubits and the rejection sampling procedure is efficient (i.e. not many raw samples are discarded). Once we obtain sufficiently many samples from the desired distribution $\nu(x)$, we compute their mean which is close to E_0 with high probability.

Although good estimates of the spectral gap are available for many physically relevant models [44], this is not always the case. Furthermore, our algorithm will yield good estimates of the ground state energy for evolution times $t \ll \Delta^{-1}$ if the initial state has more structure. For instance, if it has large overlap with the ground state and little overlap with other low-energy states. Thus, it is highly desirable to ensure that our algorithm can be proven to give reliable estimates when we do not know Δ or have reasons to believe we have a structured initial state. To address this, we devise a test that does not rely on knowledge of the spectral gap and only accepts if the ground state energy estimate produced by the algorithm is within some tolerated error of the true ground state energy [1]. Furthermore, the test will always accept if we evolve for times proportional to the inverse of the spectral gap. Although our algorithm, together with the certification result, is the first to deliver certifiably correct energy estimates in the regime of depths $\ll \epsilon^{-1}$, its sample complexity of $\tilde{\mathcal{O}}(\epsilon^{-4} p_0^{-3})$ might make its implementation prohibitive in some contexts.

The remainder of this paper is organized as follows. In Section 2, we formally define our problems and summarize our main results. Then in Section 3, we describe our method for performing rejection sampling on a quantum computer and utilize it to design a quantum algorithm for GSEE. The implementation details of certain parts of the algorithm depend on the input model of H and are deferred to Sections 4 and 5, where we develop the full-fledged versions of this algorithm in the block-encoding model and the Hamiltonian-evolution model, respectively, and analyze the costs of the resulting algorithms. After that, in Section 6, we present a method for certifying the correctness of the outputs of our algorithms. Finally, Section 7 concludes this paper.

2 Overview of the main results

2.1 Problem formulation

Let us start with a formal definition of the GSEE problem:

Definition 2.1 (GSEE)**.** *Suppose H is an n-qubit Hamiltonian with unknown spectral decomposition $H = \sum_{j=0}^{N-1} E_j |E_j\rangle\langle E_j|$, where $N = 2^n$, $-1 \le E_0 < E_1 \le E_2 \le ... \le E_{N-1} \le 1$, $E_1 - E_0 \ge \Delta$ for known $\Delta > 0$, and the $|E_j\rangle$'s are orthonormal eigenstates of H. Moreover, suppose we can prepare an n-qubit state $|\psi\rangle = \sum_{j=0}^{N-1} \gamma_j |E_j\rangle$ such that $|\gamma_0|^2 \ge \eta$ for known $\eta \in (0, 1)$. Our goal is to estimate E_0 within accuracy (i.e. additive error) ϵ with probability at least $1 - \delta$, for given $\epsilon > 0$*

[1] Note that our test certifies that a ground state energy estimate is ϵ-close to the correct one, *not* that the spectral gap of the Hamiltonian was at least some Δ_0.

and $\delta \in (0,1)$. Furthermore, we want to achieve this by using quantum circuits that depend on H and $|\psi\rangle$ and classical post-processing of their outcomes.

The above definition of the GSEE problem has omitted one important detail — how H is accessed. In this paper, we will consider two common input models of H: the block-encoding model [42] and the Hamiltonian-evolution model.

Definition 2.2 (block-encoding)**.** *Let $n, m \in \mathbb{N}$ and $\alpha, \gamma \in \mathbb{R}^+$ be arbitrary. We say that an $(n+m)$-qubit unitary operator U is an (α, m, γ)-block-encoding of an n-qubit linear operator A if*

$$\|\alpha\left(\langle 0^m| \otimes I\right) U \left(|0^m\rangle \otimes I\right) - A\| \leq \gamma. \tag{1}$$

For simplicity, we will only consider $(1, m, 0)$-block-encodings of H in this work. In fact, our results in Sections 4 can be readily extended to general block-encodings of H, but doing so would complicate the notation and obscure our main message.

Definition 2.3 (GSEE in the block-encoding model)**.** *For GSEE in the block-encoding model, we assume that there exists a method to implement a $(1, m, 0)$-block-encoding U of H for some integer $m \geq 0$. Our quantum circuits will consist of multiple queries to U and $U^\dagger$ and other primitive quantum gates, and act on $\left|0^k\right\rangle |\psi\rangle$ for some $k \geq m$. The other assumptions are the same as in Definition 2.1.*

Definition 2.4 (GSEE in the Hamiltonian-evolution model)**.** *For GSEE in the Hamiltonian-evolution model, we assume that there exists a method to implement the controlled time evolution of H, i.e. the unitary operation controlled-$e^{\mathbf{i}Ht} \coloneqq |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{\mathbf{i}Ht}$. Each of our quantum circuits will consist of multiple queries to controlled-$e^{\mathbf{i}Ht}$ and controlled-$e^{-\mathbf{i}Ht}$ for some $t = \Theta(1)$ and other primitive quantum gates, and act on $\left|0^k\right\rangle |\psi\rangle$ for some $k \geq 1$. The other assumptions are the same as in Definition 2.1.*

2.2 Main results

Our main results are threefold: a new way of performing rejection sampling on a quantum device, how to use it to obtain state-of-the-art early fault-tolerant algorithms for GSEE, and techniques for certifying ground state energy estimates. We will now informally summarize these technical contributions.

First, we start with a novel way to perform rejection sampling on a quantum computer:

Theorem 2.5 (Rejection sampling from the spectral measure)**.** *Let H be an n-qubit Hamiltonian with spectral decomposition $H = \sum_i E_i |E_i\rangle\langle E_i|$, and let $|\psi\rangle$ be an n-qubit state. Define the probability measure p on $\mathbb{R}$ as*

$$p(x) = \sum_i p_i \delta(x - E_i), \quad p_i = |\langle E_i|\psi\rangle|^2, \tag{2}$$

Furthermore, let ν, μ be two probability density functions on $\mathbb{R}$ such that $supp(p * \nu) \subset supp(\mu)$, where $*$ denotes the convolution. Finally, assume that there exists a constant $M \in \mathbb{C}$ such that for every $x \in supp(\mu)$ we have access to a $(1, m, 0)$-block-encoding U_x of an operator $h_x(H)$ which satisfies

$$|h_x(H)|^2 = \frac{\nu(H - xI)}{|M|^2 \mu(x)}, \quad \|h_x(H)\| \leq 1, \tag{3}$$

and that we can generate samples distributed according to μ. Then we can generate a sample distributed according to $p * \nu$ from an expected number $|M|^2$ uses of such U_x's.

Using this rejection sampling routine, we are then able to obtain the following GSEE algorithm:

Theorem 2.6 (GSEE algorithm, informal)**.** *Suppose H is an n-qubit Hamiltonian with unknown spectral decomposition $H = \sum_{j=0}^{N-1} E_j |E_j\rangle \langle E_j|$, where $N = 2^n$, $-1 \leq E_0 < E_1 \leq \cdots \leq E_{N-1} \leq 1$, $E_1 - E_0 \geq \Delta$ for known $\Delta > 0$. Moreover, suppose we can prepare an n-qubit state $|\psi\rangle$ such that $|\langle\psi|E_0\rangle|^2 \geq \eta$ for known $\eta \in (0, 1)$. Then in both the block-encoding model and the Hamiltonian-evolution model, there exists an algorithm that, for any given $\epsilon > 0$ and $\delta \in (0, 1)$, estimates E_0 within additive error ϵ with probability at least $1 - \delta$ by making use of*

$$\tilde{\mathcal{O}}\left(\eta^{-1}(\Delta^2\epsilon^{-2} + \log\left(\Delta^{-1}\right))\log\left(\delta^{-1}\right)\right)$$

quantum circuits, where each circuit has depth proportional to

$$\tilde{\mathcal{O}}\left(\Delta^{-1}\log(\eta^{-1}\Delta\epsilon^{-1})\right).$$

We emphasize that in Theorem 2.6, Δ is merely a lower bound on the true spectral gap $\Delta_{\text{true}} \coloneqq E_1 - E_0$ of H, not necessarily Δ_{true} itself. In fact, we can tune it between $\Theta(\epsilon)$ and $\Theta(\Delta_{\text{true}})$, and doing so leads to a class of GSEE algorithms:

Corollary 2.7 (GSEE algorithms, informal)**.** *Suppose H is an n-qubit Hamiltonian with unknown spectral decomposition $H = \sum_{j=0}^{N-1} E_j |E_j\rangle \langle E_j|$, where $N = 2^n$, $-1 \leq E_0 < E_1 \leq \cdots \leq E_{N-1} \leq 1$, $\Delta_{\text{true}} = E_1 - E_0$ is the spectral gap of H. Moreover, suppose we can prepare an n-qubit state $|\psi\rangle$ such that $|\langle\psi|E_0\rangle|^2 \geq \eta$ for known $\eta \in (0, 1)$. Then in both the block-encoding model and the Hamiltonian-evolution model, for every $\beta \in [0, 1]$, there exists an algorithm that, for any given $\epsilon > 0$ and $\delta \in (0, 1)$, estimates E_0 within additive error ϵ with probability at least $1 - \delta$ by making use of*

$$\tilde{\mathcal{O}}\left(\eta^{-1}\left(\epsilon^{-2+2\beta}\Delta_{\text{true}}^{2-2\beta} + \log\left(\epsilon^{-\beta}\Delta_{\text{true}}^{\beta-1}\right)\right)\log\left(\delta^{-1}\right)\right)$$

quantum circuits, where each circuit has depth proportional to

$$\tilde{\mathcal{O}}(\epsilon^{-\beta}\Delta_{\text{true}}^{\beta-1}\log(\eta^{-1}\Delta_{\text{true}}^{1-\beta}\epsilon^{\beta-1})).$$

Note that this result interpolates between a standard limit ($\beta = 0$) and a Heisenberg limit ($\beta = 1$) by exchanging the circuit depth for the number of samples. Importantly, rather than scaling as η^{-2} like previous methods [40, 41], the number of repetitions here is quadratically improved to an η^{-1} scaling. This is achieved without increasing (i.e. worsening the scaling of) the number of quantum operations per circuit, which is what standard approaches to quadratic speedups (e.g. amplitude amplification) require.

Figure 1 compares the performance of our GSEE algorithm with others in the literature. In particular, one can see that our algorithm has the best performance in terms of η in the regime of circuit depths between $\tilde{\mathcal{O}}(\Delta_{\text{true}}^{-1})$ and $\tilde{\mathcal{O}}(\epsilon^{-1})$.

So far, we have only considered algorithms for GSEE with circuit depths proportional to Δ^{-1}, where $\Delta > \epsilon$ is a lower bound on the spectral gap $\Delta_{\text{true}} = E_1 - E_0$ of H. However, it is desirable to develop algorithms that do not require this knowledge to accurately estimate the ground state energy. Thus, we develop a method for certifying the ground state energy estimate, albeit with a higher sample complexity:

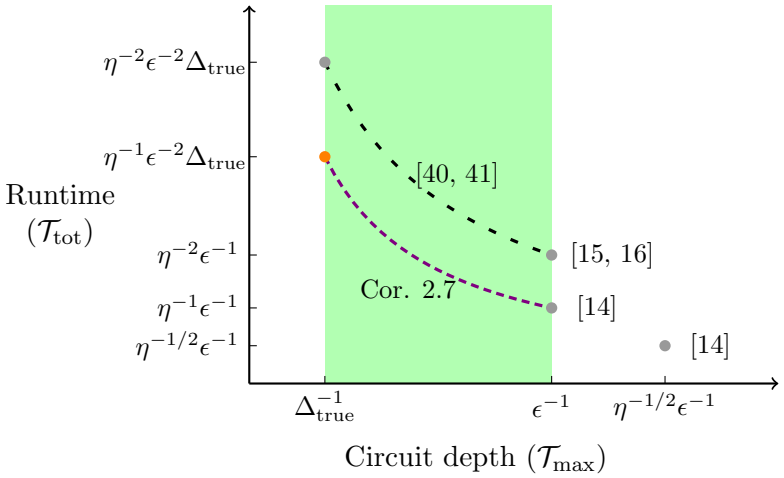

Figure 1: This figure shows the landscape of early fault-tolerant GSEE algorithms plotted according to their runtimes and circuit depths. The loosely dashed green curve represents our previous interpolation result [40] (which is also implied by Ref. [41]), while the lower densely dashed violet curve depicts the result in Corollary 2.7 (which encompasses the results in Corollary 4.8 and Corollary 5.8). For the purposes of this plot, we have assumed that $\epsilon < \eta\Delta_{\text{true}}$, leading us to place $\eta^{-1}\epsilon^{-2}\Delta_{\text{true}}$ above $\eta^{-2}\epsilon^{-1}$, though this assumption is not necessary for the main results of our paper.

Theorem 2.8 (Ground state energy certification, informal)**.** *Under the same conditions as Corollary 2.7, there exists an algorithm that is given, as an input, arbitrary $\sigma > 0$ and $\hat{E}_0 \in [E_0-\sigma, E_0+\sigma]$ (but no lower bound on Δ_{true}). It makes*

$$\tilde{\mathcal{O}}(\eta^{-3}\epsilon^{-4}\log\left(\delta^{-1}\right)) \tag{4}$$

uses of quantum circuits of depth $\tilde{\mathcal{O}}(\sigma^{-1})$ and produces a new estimate $\hat{E}_0'$ of E_0 that it either accepts or rejects. If $|\hat{E}_0'-E_0| > 4\epsilon$, it rejects $\hat{E}_0'$ with probability at least $1-\delta$. If it accepts, then $|\hat{E}_0'-E_0| \leq \epsilon$ with probability of incorrect acceptance at most δ. Furthermore, the maximal σ for which the algorithm outputs an estimate $\hat{E}_0'$ it accepts satisfies $|\hat{E}_0' - E_0| \leq \epsilon$ is $\sigma = \Omega\left(\Delta_{\text{true}}/\log\left(\epsilon^{-1}\eta^{-1}\right)\right)$.

Thus, with a higher sample complexity, we can also produce a certificate that the estimate was correct, even if we do not know the spectral gap Δ_{true} of H. Furthermore, the algorithm described above will always accept with high confidence if we go to depths proportional to the true spectral gap, the minimal depth required in Cor. 2.7. Note, however, that in some situations the algorithm can accept and, thus, give a correct estimate, even if $\sigma \gg \Delta_{\text{true}}$. As we discuss in more detail in Section 6, this can happen if, e.g. the initial state has little support on low-energy excitations. We see that our certification algorithm *does not* certify the spectral gap itself, only our ground state energy estimate. To the best of our knowledge, this is the first method to achieve certifiable GSEE in the regime where the circuit depth is $\mathcal{O}(\Delta_{\text{true}}^{-1})$. Finally, note that the assumption that we are given an estimate that satisfies $|E_0 - \hat{E}_0| \leq \sigma$ does not increase the overall depth of circuits required for the certification method. This is because we will show how to obtain such an estimate with circuits of depth $\mathcal{O}(\sigma^{-1})$ later in Section 3.3. Finally, let us briefly comment on the role of δ. Our algorithm will be based on empirically estimating the variance of a random variable up to a certain precision. As usual with such protocols, the empirical estimate will be within a tolerated error up to a given failure probability. It is the case that our protocol always rejects correctly if it

is given estimates of the variance within the tolerated error. Thus, the δ quantifies the probability of the estimates deviating more than required for the verification and, then, we do not posess a guarantee on its performance.

3 GSEE via rejection sampling

In the section, we will discuss our GSEE algorithm which will be based on a novel way of performing rejection sampling on a quantum computer. But since we believe that this rejection sampling routine will find applications elsewhere, we will first describe it in a higher level of generality.

3.1 Rejection sampling for the spectral measure

Given the spectral measure p associated with an initial state $|\psi\rangle\langle\psi|$ and a Hamiltonian $H = \sum_i |E_i\rangle\langle E_i|$, i.e.

$$p(x) = \sum_i p_i \delta(x - E_i), \quad p_i = |\langle E_i|\psi\rangle|^2, \tag{5}$$

we will show how to sample from a random variable with distribution $p * \nu$ on a quantum computer, where ν is another probability measure and $*$ denotes the convolution. It is then not difficult to see that the density of $p * \nu$ is given by:

$$(p * \nu)(x) = \sum_i p_i \nu(x - E_i). \tag{6}$$

Sampling from such distributions for various choices of ν gives us access to information about the spectrum of the Hamiltonian. For instance, if ν is itself close to a delta distribution at 0, then sampling from this distribution is close to sampling from p itself. This will, roughly speaking, correspond to our algorithm later, where we will pick ν to be a Gaussian with small variance.

Let us now discuss the sampling routine. We will assume that we can *classically* sample from a random variable X distributed according to a density μ such that $\mathrm{supp}(p * \nu) \subset \mathrm{supp}(\mu)$. Furthermore, for every $x \in \mathrm{supp}(\mu)$, we can implement a $(1, m, 0)$-block-encoding U_x of the operator $\frac{\tilde{\nu}(H - xI)}{M\tilde{\mu}(x)}$, where $\tilde{\nu}$ and $\tilde{\mu}$ satisfy:

$$\left|\frac{\tilde{\nu}(H - xI)}{M\tilde{\mu}(x)}\right|^2 = \frac{\nu(H - xI)}{|M|^2\mu(x)}, \tag{7}$$

and $M \in \mathbb{C}$ is a constant such that for all $x \in \mathrm{supp}(\mu)$:

$$\sum_i p_i \frac{\nu(E_i - x)}{|M|^2\mu(x)} \leq 1. \tag{8}$$

There are many cases where explicit block encoding circuits have been developed [45, 46].

Let us now explain the quantum algorithm that allows us to sample from $p * \nu$ in this setting with an expected number $\mathcal{O}(|M|^2)$ runs of such U_x's. First, we generate a sample x from the density μ. After that, we implement the block-encoding U_x of $\frac{\tilde{\nu}(H - xI)}{M\tilde{\mu}(x)}$ and apply it to $|0^m\rangle|\psi\rangle$. Then we measure the m ancilla qubits of the resulting state. If the outcome is 0^m, we accept the sample. Let us see why, conditioned on the acceptance, x will be distributed according to $p * \nu$.

Note that:

$$U_x \left|0^m\right\rangle \left|\psi\right\rangle = \left|0^m\right\rangle \frac{\tilde{\nu}(H - xI)}{M\tilde{\mu}(x)} \left|\psi\right\rangle + \sum_{j \in \{0,1\}^m,\ j \neq 0^m} \left|j\right\rangle \left|\psi'_j\right\rangle, \tag{9}$$

where the $\left|\psi'_j\right\rangle$'s are unnormalized states. By expanding $\frac{\tilde{\nu}(H-xI)}{M\tilde{\mu}(x)} \left|\psi\right\rangle$ in the eigenbasis of H, we see that the probability of measuring 0^m on the first register is expressed by:

$$\sum_i \left| \frac{\tilde{\nu}(E_i - x)}{M\tilde{\mu}(x)} \right|^2 |\langle E_i|\psi\rangle|^2 = \sum_i p_i \frac{\nu(E_i - x)}{|M|^2 \mu(x)} = \frac{(p * \nu)(x)}{|M|^2 \mu(x)}. \tag{10}$$

Thus the probability we accept is exactly the same as if we ran the rejection sampling algorithm for the densities $p * \nu$ and μ. Furthermore, by standard properties of rejection sampling, we will, in expectation, need $\mathcal{O}(|M|^2)$ trials to generate a single sample from the target distribution.

From the discussion above, we obtain our new rejection sampling routine:

Theorem 3.1. *Let $H = \sum_i E_i |E_i\rangle\langle E_i|$ be a Hamiltonian on n qubits and $|\psi\rangle$ an n-qubit state. Definite the probability measure p on $\mathbb{R}$ as*

$$p(x) = \sum_i p_i \delta(x - E_i), \quad p_i = |\langle E_i|\psi\rangle|^2, \tag{11}$$

Furthermore, let ν, μ be two probability density functions on $\mathbb{R}$ such that $supp(p * \nu) \subset supp(\mu)$, where $*$ denotes the convolution. Finally, assume that there exists a constant $M \in \mathbb{C}$ such that for every $x \in supp(\mu)$ we have access to a $(1, m, 0)$-block-encoding of a unitary U_x of an operator h_x which satisfies

$$|h_x(H)|^2 = \frac{\nu(H - xI)}{|M|^2 \mu(x)}, \quad \|h_x(H)\| \leq 1, \tag{12}$$

and that we can generate samples distributed according to μ. Then we can generate a sample distributed according to $p * \nu$ from an expected number $|M|^2$ uses of such U_x's.

Let us illustrate the rejection sampling routine with one of the leading examples of this work: conditioned normal random variables. Suppose that we wish to sample from $p * n_{\sigma,[a,b]}$, where $n_{\sigma,[a,b]}$ is the density of a normal random variable with mean 0 and variance σ^2 conditioned on an interval $[a, b]$. That is $n_{\sigma,[a,b]}$ corresponds to the density:

$$n_{\sigma,[a,b]}(x) = \frac{e^{-\frac{x^2}{2\sigma}}}{\int_a^b e^{-\frac{y^2}{2\sigma}} dy} \chi_{[a,b]}(x). \tag{13}$$

Here $\chi_{[a,b]}(x)$ is the indicator function. One possible strategy we will follow to sample from this distribution is to perform rejection sampling with the density μ given by the uniform distribution on $[a, b]$. In this case, if we implement a block-encoding of the operator

$$h_\sigma(H - xI) = e^{-\frac{(H-xI)^2}{4\sigma^2}}, \tag{14}$$

we see that it satisfies the requirement in Eq. (7) with $|M|^2$ given by

$$|M|^2 = \frac{\sigma\sqrt{2\pi}}{(b-a)\int_a^b (p * n_{\sigma,[-\infty,\infty]})(x) dx}, \tag{15}$$

as with this choice of $|M|^2$, the ratio of the density of $p * n_{\sigma,[a,b]}$ and the uniform distribution on $[a, b]$ coincide with the probability of accepting. Furthermore, the convolution $p*h_\sigma$ is clearly upper bounded by 1. Note that we do not need to know the constant in Eq. (15) to run the algorithm, only h_σ. But $|M|^2$ will determine the number of repetitions.

Building upon recent work on GSEE, we will show how to estimate the ground state energy by first sampling from the convolution of the spectral measure with a uniform distribution on $[0, 1]$ (the approach of Ref. [15]) and then refining the estimate by sampling with the convolution of the same measure with a normal random variable (the approach of Ref. [40]). However, the previous algorithms worked by *estimating the density* $\nu * p$ through generalized Hadamard tests. As here we generate samples instead of estimating densities, we will obtain a better scaling in the ground-state overlap.

We believe that the rejection sampling subroutine developed in this work is of independent interest and will find application elsewhere. Indeed, to the best of our knowledge, it is the first quantum algorithm for sampling from distributions associated with the spectral measure with a constant number of ancilla qubits.

3.2 Basic idea for GSEE

We will now discuss our algorithm for GSEE based on rejection sampling. This algorithm consists of two stages. The first stage generates a coarse estimate $\tilde{E}_0$ of E_0 by using the strategy of Ref. [14]. The second stage refines this estimate and improves its accuracy to arbitrarily small ϵ. To achieve this goal, we make the following observation about Gaussian functions. Let $f_\sigma(x) \coloneqq e^{-x^2/(2\sigma^2)}$ be (an unnormalized) Gaussian. If we truncate $f_\sigma^2(x) = e^{-x^2/\sigma^2}$ to a sufficiently large interval around 0, then the mean of this truncated function is extremely close to 0, as long as the middle point of the interval is not far from 0:

Lemma 3.2. *Let $\sigma > 0$ and $\epsilon \in (0, \sigma]$ be arbitrary, and let $w = 2\sigma\sqrt{\ln(e\sigma/\epsilon)}$. Then for any $s \in [-w/2, w/2]$, we have $\int_{s-w}^{s+w} e^{-x^2/\sigma^2} dx \geq 1.12\sigma$ and $\left|\int_{s-w}^{s+w} x e^{-x^2/\sigma^2} dx\right| \leq \frac{\sigma\epsilon}{2e}$, and hence*

$$\frac{\left|\int_{s-w}^{s+w} x e^{-x^2/\sigma^2} dx\right|}{\int_{s-w}^{s+w} e^{-x^2/\sigma^2} dx} \leq \frac{\epsilon}{2.24e}. \tag{16}$$

If we perturb $f_\sigma^2(x)$ slightly, then the resulting function still possesses a similar property:

Lemma 3.3. *Let $\sigma > 0$ and $\epsilon \in (0, \sigma]$ be arbitrary, and let $w = 2\sigma\sqrt{\ln(e\sigma/\epsilon)}$. Then for any $s \in [-w/2, w/2]$, if $g : [s - w, s + w] \to \mathbb{R}^+$ satisfies that $|g(x) - e^{-x^2/\sigma^2}| \leq \frac{0.03\epsilon}{\sigma \ln(e\sigma/\epsilon)}$ for all $x \in [s - w, s + w]$, then we have $\int_{s-w}^{s+w} g(x)dx \geq \sigma$ and $\left|\int_{s-w}^{s+w} x g(x)dx\right| \leq 0.55\sigma\epsilon$, and hence*

$$\frac{\left|\int_{s-w}^{s+w} x g(x)dx\right|}{\int_{s-w}^{s+w} g(x)dx} \leq 0.55\epsilon. \tag{17}$$

The proofs of Lemma 3.2 and Lemma 3.3 are given in Appendix A. As a corollary of Lemma 3.3, we obtain the following strategy for estimating an unknown quantity E with high accuracy:

Corollary 3.4. *Under the same assumptions as in Lemma 3.3, if X is a random variable such that it takes values from $[E+s-w, E+s+w]$ for some unknown $E \in \mathbb{R}$ and its probability density function is proportional to $g(x-E)$, then its expected value $\mathbb{E}[X]$ is $\mathcal{O}(\epsilon)$-close to E.*

Now we apply this strategy to GSEE in which the target quantity is E_0. We first choose appropriate $\sigma = \mathcal{O}(\Delta/\sqrt{\log(\eta^{-1}\Delta\epsilon^{-1})})$ and $w = \mathcal{O}(\sigma\sqrt{\log(\sigma/\epsilon)})$. Then we find a coarse estimate $\tilde{E}_0$ of E_0 such that $s \coloneqq \tilde{E}_0 - E_0 \subseteq [-w/2, w/2]$. Next, we pick x uniformly at random from $[\tilde{E}_0 - w, \tilde{E}_0 + w]$, and construct a quantum circuit that implements $f_\sigma(H - xI)$ approximately and probabilistically, and run this circuit on the initial state $|0^k\rangle|\psi\rangle$ for appropriate $k \geq 1$. Then the probability of this circuit approximately performing $f_\sigma(H-xI)$ on $|\psi\rangle$ is roughly proportional to $f_\sigma^2(x - E_0)$. Informally, this is because this probability is close to

$$\|f_\sigma(H - xI)|\psi\rangle\|^2 = \sum_{j=0}^{N-1} p_j f_\sigma^2(E_j - x) \approx p_0 f_\sigma^2(E_0 - x), \tag{18}$$

since Gaussian functions decrease exponentially from their peaks and $x \in [\tilde{E}_0 - w, \tilde{E}_0 + w]$ is quite far from the E_j's for all $j \geq 1$. We can determine whether this event happens by measuring the k ancilla qubits and checking if particular outcomes are obtained in the end. If this event happens, then x is accepted and appended to an array S; otherwise, x is rejected and discarded. Clearly, the entries of S are i.i.d. random variables. Let Y be such a random variable. By Corollary 3.4, its expected value $\mathbb{E}[Y]$ is $\mathcal{O}(\epsilon)$-close to E_0. Furthermore, we can prove that the probability of a random $x \in [\tilde{E}_0 - w, \tilde{E}_0 + w]$ getting accepted by this procedure is $\Omega(\eta/\sqrt{\log(\Delta\epsilon^{-1})})$, i.e. this rejection sampling method is efficient. Finally, we repeat the above procedure $\tilde{\mathcal{O}}(\eta^{-1}\Delta^2\epsilon^{-2})$ times and compute the mean of the accepted samples. Then by Chernoff's bound, this quantity will be $\mathcal{O}(\epsilon)$-close to $\mathbb{E}[Y]$ with high probability, which implies that it is also an $\mathcal{O}(\epsilon)$-accurate estimate of E_0 with high probability, as desired.

3.3 Formal description of the GSEE algorithm

We will now describe our GSEE algorithm in a way that is agnostic to the input model of H. This algorithm can be executed in both the block-encoding and Hamiltonian-evolution models. It requires the block-encodings of certain functions of H whose implementation details depend on the specific input model and will be explained in Sections 4 and 5. The other components of the algorithm remains identical in these two models.

We first propose a basic GSEE algorithm (i.e. Algorithm 1) which generalizes the one from Ref. [14]. This algorithm will be used to generate a coarse estimate of E_0.

Before we give a formal description of the algorithm, let us consider a simplified and idealized version to build up some intuition. Suppose that for $\lambda \in (-1, 1)$, we can implement a $(c, m, 0)$-block-encoding V_λ of the operator $f_\lambda(H)$, where $f_\lambda : [-1, 1] \to [0, 1]$ satisfies $f(x) = 1$ for $x \in [-1, \lambda]$ and $f(x) = 0$ for $x \in (\lambda, 1]$. Let us now suppose that $|\psi\rangle$ is an initial state such that $|\langle\psi|E_0\rangle|^2 \geq \eta$ for some known $\eta \in (0, 1)$. If we pick $\lambda > E_0$, then $\|f_\lambda(H)|\psi\rangle\|^2 \geq \eta$, as the operator $f_\lambda(H)$ will act as the identity on the ground space of H. On the other hand, if we pick $\lambda < E_0$ we have $\|f_\lambda(H)|\psi\rangle\|^2 = 0$, as it will send all the eigenspaces of H to 0.

Furthermore, by the definition of block-encoding, if we measure the m ancilla qubits of the state $V_\lambda|0^m\rangle|\psi\rangle$, the probability of getting outcome 0^m is given by $c^2\|f_\lambda(H)|\psi\rangle\|^2$. Thus, if we could implement such a block-enconding, we could also infer what the ground state energy is by

performing a binary search. Indeed, we only need to find the smallest λ such that $\|f_\lambda(H)\,|\psi\rangle\,\|^2 \geq \eta$. As for a given λ, whether $\|f_\lambda(H)\,|\psi\rangle\,\|^2 \geq \eta$ or $\|f_\lambda(H)\,|\psi\rangle\,\|^2 = 0$ can be determined with success probability of $2/3$ by repeating the aforementioned experiment $\mathcal{O}(\eta^{-1})$ times and checking how often we obtain measurement outcome 0^m. So the block-encoding V_λ would give rise to an algorithm with $\mathcal{O}(\eta^{-1}\log\left(\epsilon^{-1}\right)))$ runs to estimate the ground state energy with accuracy ϵ. However, implementing block-encodings of non-smooth functions like f_λ cannot be accomplished directly and it is necessary to perform suitable approximations of f_λ by smooth functions. This motivates the introduction of the function $f_{a,b;\epsilon}$ below. The fact that we only approximate f_λ also introduces subtleties on how we post-process the samples and perform the binary search, but the strategy is comparable to the one described above.

Algorithm 1 Basic GSEE algorithm

procedure BASIC_GSEE(ϵ, δ, η, c, m)
  $l \leftarrow -1$;
  $r \leftarrow 1$;
  $L \leftarrow \lceil \log_{3/2}\left(\epsilon^{-1}\right)\rceil$;
  $M \leftarrow \lceil 12c\eta^{-1}\ln\left(L\delta^{-1}\right)\rceil$;
  $\epsilon' \leftarrow \min(\sqrt{0.1\eta}, 0.05)$;
  while $r - l > 2\epsilon$ **do**
    $K \leftarrow 0$;
    $a \leftarrow (2l + r)/3$;
    $b \leftarrow (l + 2r)/3$;
    Let $V_{a,b;\epsilon'}$ be a $(c, m, 0)$-block-encoding of $f_{a,b;\epsilon'}(H)$, where $f_{a,b;\epsilon'} : [-1, 1] \to \mathbb{R}$ satisfies that $1 - \epsilon' \leq f_{a,b;\epsilon'}(x) \leq 1$ for all $x \in [-1, a]$, $|f_{a,b;\epsilon'}(x)| \leq 1$ for all $x \in [a, b]$, and $|f_{a,b;\epsilon'}(x)| \leq \epsilon'$ for all $x \in [b, 1]$;
    for $j = 1, 2, \ldots, M$ **do**
      Measure the first m qubits of $V_{a,b;\epsilon'}\,|0^m\rangle\,|\psi\rangle$ in the standard basis;
      if the measurement outcome is 0^m **then**
        $K \leftarrow K + 1$;
      end if
    end for
    if $K < 0.5c^{-1}\eta M$ **then**
      $l \leftarrow a$;
    else
      $r \leftarrow b$;
    end if
  end while
  return $(l + r)/2$.
end procedure

Theorem 3.5. *Suppose H is an n-qubit Hamiltonian with unknown spectral decomposition $H = \sum_{j=0}^{N-1} E_j\,|E_j\rangle\,\langle E_j|$, where $N = 2^n$, $-1 \leq E_0 < E_1 \leq E_2 \leq \ldots \leq E_{N-1} \leq 1$, and the $|E_j\rangle$'s are orthonormal eigenstates of H. Moreover, suppose we can prepare an n-qubit state $|\psi\rangle = \sum_{j=0}^{N-1}\gamma_j\,|E_j\rangle$ such that $|\gamma_0|^2 \geq \eta$ for known $\eta \in (0, 1)$. Finally, suppose for arbitrary $a, b \in (-1, 1)$ and $\epsilon' \geq 0$, we can implement a $(c, m, 0)$-block-encoding of $f_{a,b;\epsilon'}(H)$, where $f_{a,b;\epsilon'} : [-1, 1] \to \mathbb{R}$ satisfies that $1 - \epsilon' \leq f_{a,b;\epsilon'}(x) \leq 1$ for all $x \in [-1, a]$, $|f_{a,b;\epsilon'}(x)| \leq 1$ for all $x \in [a, b]$, and $|f_{a,b;\epsilon'}(x)| \leq \epsilon'$ for all $x \in [b, 1]$. Then for arbitrary $\epsilon > 0$ and $\delta \in (0, 1)$, the output of Algorithm 1 (i.e.*

BASIC_GSEE$(\epsilon, \delta, \eta, c, m))$ *is* ϵ*-close to* E_0 *with probability at least* $1-\delta$.

Proof. We claim that in each iteration of the **while**-loop, if $E_0 \leq a = (2l+r)/3$, then $K > 0.5c^{-1}\eta M$ with probability at least $1-\delta/L$; if $E_0 \geq b = (l+2r)/3$, then $K < 0.5c^{-1}\eta M$ with probability at least $1-\delta/L$. If this is true, then after each iteration, E_0 remains between l and r with probability at least $1-\delta/L$. Since there are at most L iterations in the **while**-loop, E_0 is between the final l and r with probability at least $1-\delta$. Then since $l-r \leq 2\epsilon$ in the end, we know that $(l+r)/2$ is ϵ-close to E_0 with probability at least $1-\delta$, as desired.

To prove the above claims, we need to bound the probability of obtaining measurement outcome 0^m in each iteration of the **for**-loop in the cases $E_0 \leq a$ and $E_0 \geq b$ respectively. If $E_0 \leq a$, then this probability is

$$c^{-1}\sum_{j=0}^{N-1} p_j f_{a,b;\epsilon'}^2(E_j) \geq c^{-1} p_0 f_{a,b;\epsilon'}^2(E_0) \geq c^{-1}\eta(1-\epsilon')^2 \geq 0.9c^{-1}\eta. \tag{19}$$

On the other hand, if $E_0 \geq b$, then this probability is

$$c^{-1}\sum_{j=0}^{N-1} p_j f_{a,b;\epsilon'}^2(E_j) \leq c^{-1} \max_{0\leq j\leq N-1} f_{a,b;\epsilon'}^2(E_j) \leq c^{-1}(\epsilon')^2 \leq 0.1c^{-1}\eta. \tag{20}$$

Next, we invoke the Chernoff-Hoeffding Theorem which states that if $X_1, \ldots, X_n$ are i.i.d. random variables which take values in $\{0,1\}$. Let $p = \mathbb{E}[X_1]$. Then for any $q \in (p,1]$ and $r \in [0,p)$, we have

$$\mathbb{P}\left[\frac{1}{n}\sum_{i=1}^n X_i \geq q\right] \leq e^{-nD(q\|p)}, \tag{21}$$

$$\mathbb{P}\left[\frac{1}{n}\sum_{i=1}^n X_i \leq r\right] \leq e^{-nD(r\|p)}, \tag{22}$$

where $D(x\|y) \coloneqq x\ln(x/y) + (1-x)\ln((1-x)/(1-y))$ is the Kullback–Leibler divergence between Bernoulli random variables with parameters x and y respectively. We also use the fact that

$$D(x\|y) \geq \frac{(x-y)^2}{2\max(x,y)}, \quad \forall x, y \in (0,1). \tag{23}$$

If $E_0 \leq a$, then we have $p \geq 0.9c^{-1}\eta$, and hence

$$\mathbb{P}\left[K/M \leq 0.5c^{-1}\eta\right] \leq e^{-MD(0.5c^{-1}\eta\|p)} \tag{24}$$

in which

$$D(0.5c^{-1}\eta\|p) \geq \frac{(p-0.5c^{-1}\eta)^2}{2p} \geq 0.088c^{-1}\eta, \tag{25}$$

as $\frac{(p-0.5c^{-1}\eta)^2}{2p}$ increases monotonically for $p \geq 0.5c^{-1}\eta$. Then by our choice of M, we know that $\mathbb{P}\left[K \leq 0.5c^{-1}\eta M\right] \leq \delta/L$.

On the other hand, if $E_0 \geq b$, we have $p \leq 0.1c^{-1}\eta$, and hence

$$\mathbb{P}\left[K/M \geq 0.5c^{-1}\eta\right] \leq e^{-MD(0.5c^{-1}\eta\|p)} \tag{26}$$

in which

$$D(0.5c^{-1}\eta\|p) \geq \frac{(0.5c^{-1}\eta - p)^2}{2 \cdot 0.5c^{-1}\eta} \geq 0.16c^{-1}\eta. \tag{27}$$

Then by our choice of M, we obtain that $\mathbb{P}\left[K \geq 0.5c^{-1}\eta M\right] \leq \delta/L$.

The theorem is thus proved. □

Next, we present our main GSEE algorithm (i.e. Algorithm 2) which utilizes Algorithm 1 to obtain a coarse estimate of E_0 and then refines this estimate by the strategy mentioned in Section 3.2:

Algorithm 2 Advanced GSEE algorithm

procedure ADV_GSEE($\epsilon, \delta, \Delta, \eta, c_1, m_1, c_2, m_2$)
  if $\epsilon \geq \Delta/8$ **then**
    return BASIC_GSEE $(\epsilon, \delta, \eta, c_1, m_1)$;
  end if
  $\sigma \leftarrow \frac{\Delta}{5\sqrt{\ln(ec_2)}\sqrt{\ln(e\eta^{-1}\Delta\epsilon^{-1})}}$;
  $\epsilon_1 \leftarrow \min(\frac{\epsilon}{1.1}, \sigma)$;
  $w \leftarrow 2\sigma\sqrt{\ln(e\sigma/\epsilon_1)}$;
  $\tilde{E}_0 \leftarrow$ BASIC_GSEE $(w/2, \delta/3, \eta, c_1, m_1)$;
  $\epsilon'' \leftarrow \frac{0.0075\epsilon_1\eta}{\sigma\ln(e\sigma/\epsilon_1)}$;
  $M \leftarrow \lceil 32c_2^2\eta^{-1}\sqrt{\ln(e\sigma/\epsilon_1)} \cdot \max(2w^2\epsilon^{-2}\ln\left(6\delta^{-1}\right), \ln\left(3\delta^{-1}\right))\rceil$;
  $S \leftarrow$ empty array; ▷ S may contain duplicate values
  for $i = 1, 2, \ldots, M$ **do**
    Choose ξ_i uniformly at random from $[\tilde{E}_0 - w, \tilde{E}_0 + w]$;
    Let $W_{\sigma,\xi_i;\epsilon''}$ be a $(c_2, m_2, 0)$-block-encoding of $g_{\sigma,\xi_i;\epsilon''}(H)$, where $g_{\sigma,\xi_i;\epsilon''} : [-1,1] \to \mathbb{R}$ satisfies that $|g_{\sigma,\xi_i;\epsilon''}(x) - e^{-(x-\xi_i)^2/(2\sigma^2)}| \leq \epsilon''$ and $|g_{\sigma,\xi_i;\epsilon''}(x)| \leq 1$ for all $x \in [-1,1]$.
    Measure the first m_2 qubits of $W_{\sigma,\xi_i;\epsilon''}\left|0^{m_2}\right\rangle|\psi\rangle$ in the standard basis;
    if the measurement outcome is 0^{m_2} **then**
      Append ξ_i to S;
    end if
  end for
  return $\frac{1}{|S|}\sum_{y\in S} y$.
end procedure

Theorem 3.6. *Suppose H is an n-qubit Hamiltonian with unknown spectral decomposition $H = \sum_{j=0}^{N-1} E_j |E_j\rangle\langle E_j|$, where $N = 2^n$, $-1 \leq E_0 < E_1 \leq E_2 \leq \ldots \leq E_{N-1} \leq 1$, $E_1 - E_0 \geq \Delta$ for known $\Delta > 0$, and the $|E_j\rangle$'s are orthonormal eigenstates of H. Moreover, suppose we can prepare an n-qubit state $|\psi\rangle = \sum_{j=0}^{N-1}\gamma_j|E_j\rangle$ such that $|\gamma_0|^2 \geq \eta$ for known $\eta \in (0,1)$. In addition, suppose for arbitrary $a, b \in (-1,1)$ and $\epsilon' \geq 0$, we can implement a $(c_1, m_1, 0)$-block-encoding of $f_{a,b;\epsilon'}(H)$, where $f_{a,b;\epsilon'} : [-1,1] \to \mathbb{R}$ satisfies that $1 - \epsilon' \leq f_{a,b;\epsilon'}(x) \leq 1$ for all $x \in [-1,a]$, $|f_{a,b;\epsilon'}(x)| \leq 1$ for all $x \in [a,b]$, and $|f_{a,b;\epsilon'}(x)| \leq \epsilon'$ for all $x \in [b,1]$. Finally, suppose for arbitrary $\sigma, \epsilon'' \in (0,1)$ and $\xi \in [-2,2]$, we can implement a $(c_2, m_2, 0)$-block-encoding of $g_{\sigma,\xi;\epsilon''}(H)$, where $c_2 \geq 1$ and $g_{\sigma,\xi;\epsilon''} : [-1,1] \to \mathbb{R}$ satisfies that $|g_{\sigma,\xi;\epsilon''}(x) - e^{-(x-\xi)^2/(2\sigma^2)}| \leq \epsilon''$ and $|g_{\sigma,\xi;\epsilon''}(x)| \leq 1$ for all $x \in [-1,1]$. Then for arbitrary $\epsilon > 0$ and $\delta \in (0,1)$, the output of Algorithm 2 (i.e. ADV_GSEE$(\epsilon, \delta, \Delta, \eta, c_1, m_1, c_2, m_2)$) is ϵ-close to E_0 with probability at least $1 - \delta$.*

Proof. Let $\hat{E}_0$ denote the output of Algorithm 2. If $\epsilon \geq \Delta/8$, then $\hat{E}_0 = \text{BASIC\_GSEE}\,(\epsilon, \delta, \eta, c_1, m_1)$ is ϵ-close to E_0 with probability at least $1-\delta$ by Theorem 3.5. So we only need to focus on the case $\epsilon < \Delta/8$.

For convenience, we introduce the following notation. Let $A_i = 1$ if ξ_i is added to S, and $A_i = 0$ otherwise, for $i = 1, 2, \ldots, M$. Then since $\xi_1, \xi_2, \ldots, \xi_M$ are i.i.d. random variables, $A_1, A_2, \ldots, A_M$ are i.i.d. Bernoulli variables. Next, let $K = \sum_{i=1}^{M} A_i$ be the size of S at the end of the algorithm, and let X_k be the k-th number added to S, for $k = 1, 2, \ldots, K$. Then $X_1, X_2, \ldots, X_K$ are i.i.d. random variables, and $\hat{E}_0 = \frac{1}{K}\sum_{k=1}^{K} X_k$ has the same expected value as X_1.

We claim that:

$$\mathbb{P}\left[|\mathbb{E}\left[X_1\right] - E_0| \geq \epsilon/2\right] \leq \delta/3, \tag{28}$$

$$\mathbb{P}\left[K \geq 8w^2\epsilon^{-2}\ln\left(6\delta^{-1}\right)\right] \geq 1 - \delta/3, \tag{29}$$

and

$$\mathbb{P}\left[|\hat{E}_0 - \mathbb{E}\left[X_1\right]| \geq \epsilon/2 \mid K \geq 8w^2\epsilon^{-2}\ln\left(6\delta^{-1}\right)\right] \leq \delta/3. \tag{30}$$

Let G denote the event $K \geq 8w^2\epsilon^{-2}\ln\left(6\delta^{-1}\right)$. If the above claims are true, then by Eqs. (29), (30) and union bound, we get

$$\mathbb{P}\left[|\hat{E}_0 - \mathbb{E}\left[X_1\right]| \geq \epsilon/2\right] \leq \mathbb{P}\left[|\hat{E}_0 - \mathbb{E}\left[X_1\right]| \geq \epsilon/2 \mid G\right] + \mathbb{P}\left[\neg G\right] \leq \delta/3 + \delta/3 = 2\delta/3. \tag{31}$$

Then by Eqs. (28), (31), triangle inequality and union bound, we know that

$$\mathbb{P}\left[|\hat{E}_0 - E_0| \geq \epsilon\right] \leq \mathbb{P}\left[|\hat{E}_0 - \mathbb{E}\left[X_1\right]| \geq \epsilon/2\right] + \mathbb{P}\left[|\mathbb{E}\left[X_1\right] - E_0| \geq \epsilon/2\right] \leq 2\delta/3 + \delta/3 = \delta, \tag{32}$$

as desired.

It remains to prove Eqs. (28), (29) and (30). Consider each iteration in the **for**-loop. Suppose $\xi \in [\tilde{E}_0 - w, \tilde{E}_0 + w]$ is chosen [2]. Let $W_{\sigma,\xi;\epsilon''}$ be the $(c_2, m_2, 0)$-block-encoding of $g_{\sigma,\xi;\epsilon'}(H)$. That is,

$$W_{\sigma,\xi;\epsilon''}\left|0^{m_2}\right\rangle|\psi\rangle = c_2^{-1}\left|0^{m_2}\right\rangle g_{\sigma,\xi;\epsilon''}(H)\,|\psi\rangle + \sum_{s\in\{0,1\}^{m_2},\ s\neq 0^{m_2}} |s\rangle\left|\phi_{s,\xi,\epsilon''}\right\rangle, \tag{33}$$

where the $\left|\phi_{s,\xi,\epsilon''}\right\rangle$'s are unnormalized states. Consequently, the probability of obtaining measurement outcome 0^{m_2} when measuring the first m_2 qubits of this state (in the standard basis) is

$$q(\xi) \coloneqq \mathbb{P}\left[A_i = 1 | \xi_i = \xi\right] = c_2^{-2}\left\langle\psi\right| g^2_{\sigma,\xi;\epsilon''}(H)\,|\psi\rangle = c_2^{-2}\sum_{j=0}^{N-1} p_j g^2_{\sigma,\xi;\epsilon''}(E_j). \tag{34}$$

Meanwhile, let $f_\sigma(x) \coloneqq e^{-x^2/(2\sigma^2)}$ be Gaussian, and let

$$r(\xi) \coloneqq c_2^{-2}\left\langle\psi\right| f^2_\sigma(H - \xi I)\,|\psi\rangle = c_2^{-2}\sum_{j=0}^{N-1} p_j f^2_\sigma(E_j - \xi). \tag{35}$$

[2] One can prove $w \leq 0.8$ by the problem assumptions and our choice of parameters. Then since $\tilde{E}_0 \in [-1, 1]$, we have $-1.8 \leq \tilde{E}_0 - w < \tilde{E}_0 + w \leq 1.8$. This means that we might choose $\xi_i > 1$ or < -1 occasionally. This decision is made to ease the proof.

Then by the assumptions on $g_{\sigma,\xi;\epsilon''}$, we have

$$|g^2_{\sigma,\xi;\epsilon''}(x) - f^2_\sigma(x-\xi)| = |g_{\sigma,\xi;\epsilon''}(x) - f_\sigma(x-\xi)| \cdot |g_{\sigma,\xi;\epsilon''}(x) + f_\sigma(x-\xi)| \tag{36}$$
$$\leq 2|g_{\sigma,\xi;\epsilon''}(x) - f_\sigma(x-\xi)| \tag{37}$$
$$\leq 2\epsilon'', \quad \forall x \in [-1,1], \tag{38}$$

and hence $|q(\xi) - r(\xi)| \leq 2c_2^{-2}\epsilon''$.

Now since ξ_i is uniformly distributed in $[\tilde{E}_0 - w, \tilde{E}_0 + w]$, we get that

$$\mathbb{P}[A_i = 1] = \int_{\tilde{E}_0-w}^{\tilde{E}_0+w} p(\xi_i = \xi) \cdot \mathbb{P}[A_i = 1|\xi_i = \xi]\, d\xi = \frac{1}{2w}\int_{\tilde{E}_0-w}^{\tilde{E}_0+w} q(\xi)d\xi. \tag{39}$$

This implies each X_k has probability density function

$$p(X_k = \xi) = p(\xi_i = \xi | A_i = 1) \tag{40}$$
$$= \frac{p(\xi_i = \xi) \cdot \mathbb{P}[A_i = 1|\xi_i = \xi]}{\mathbb{P}[A_i = 1]} \tag{41}$$
$$= \frac{q(\xi)}{\int_{\tilde{E}_0-w}^{\tilde{E}_0+w} q(\xi)d\xi}, \quad \forall \xi \in [\tilde{E}_0 - w, \tilde{E}_0 + w]. \tag{42}$$

Furthermore, the expected value of X_k is

$$\mathbb{E}[X_k] = \int_{\tilde{E}_0-w}^{\tilde{E}_0+w} \xi \cdot p(X_k = \xi)d\xi = \frac{\int_{\tilde{E}_0-w}^{\tilde{E}_0+w} \xi q(\xi)d\xi}{\int_{\tilde{E}_0-w}^{\tilde{E}_0+w} q(\xi)d\xi}. \tag{43}$$

Next, it is convenient to make a change of variables. Let $s \coloneqq \tilde{E}_0 - E_0$. Then by Theorem 3.5, we know that $s \in [-w/2, w/2]$ with probability at least $1 - \delta/3$. From now on, we will focus on the case when this event happens. Moreover, let $x \coloneqq \xi - E_0$ and $h(x) \coloneqq q(E_0 + x) = q(\xi)$. Then we have $x \in [s - w, s + w] \subseteq [-3w/2, 3w/2]$. Moreover, we have

$$\mathbb{P}[A_i = 1] = \frac{1}{2w}\int_{s-w}^{s+w} h(x)dx, \tag{44}$$

and

$$\mathbb{E}[X_k] = E_0 + \frac{\int_{s-w}^{s+w} xh(x)dx}{\int_{s-w}^{s+w} h(x)dx}. \tag{45}$$

In Appendix B, we prove that

$$E_j - \xi \geq \sigma\sqrt{\ln\left(0.5c_2^2/\epsilon''\right)} \tag{46}$$

for all $j \geq 1$. Consequently, $f^2_\sigma(E_j - \xi) \leq 2\epsilon'' c_2^{-2}$ for all $j \geq 1$. Then by Eq. (35), we obtain

$$\left|c_2^2 p_0^{-1} r(\xi) - f^2_\sigma(E_0 - \xi)\right| = c_2^2 p_0^{-1} \sum_{j=1}^{N-1} p_j f^2_\sigma(E_j - \xi) \tag{47}$$
$$\leq c_2^2 p_0^{-1} \max_{1\leq j\leq N-1} f^2_\sigma(E_j - \xi) \tag{48}$$
$$\leq c_2^2 p_0^{-1} \cdot 2\epsilon'' c_2^{-2} \tag{49}$$
$$= 2p_0^{-1}\epsilon'', \tag{50}$$

Meanwhile, recall that $|q(\xi) - r(\xi)| \le 2c_2^{-2}\epsilon''$. Therefore, we have

$$\begin{aligned}
\left|c_2^2 p_0^{-1} q(\xi) - f_\sigma^2(E_0 - \xi)\right| &\le \left|c_2^2 p_0^{-1}\left(q(\xi) - r(\xi)\right)\right| + \left|c_2^2 p_0^{-1} r(\xi) - f_\sigma^2(E_0 - \xi)\right| && (51)\\
&\le 2p_0^{-1}\epsilon'' + 2p_0^{-1}\epsilon'' && (52)\\
&= 4p_0^{-1}\epsilon'' && (53)\\
&\le 4\eta^{-1}\epsilon'' && (54)\\
&= \frac{0.03\epsilon_1}{\sigma\sqrt{\ln\left(e\sigma/\epsilon_1\right)}} && (55)
\end{aligned}$$

where the fourth step follows from $p_0 \ge \eta$ and the last step follows from the definition of ϵ'. Now since $h(x) = q(\xi)$ and $f_\sigma^2(E_0 - \xi) = f_\sigma^2(x)$, we get

$$\left|c_2^2 p_0^{-1} h(x) - f_\sigma^2(x)\right| \le \frac{0.03\epsilon_1}{\sigma\sqrt{\ln\left(e\sigma/\epsilon_1\right)}}, \quad \forall x \in [s-w, s+w]. \tag{56}$$

Thus, by Lemma 3.3 we know that

$$\frac{\left|\int_{s-w}^{s+w} x h(x) dx\right|}{\int_{s-w}^{s+w} h(x) dx} \le 0.55\epsilon_1 \le 0.5\epsilon. \tag{57}$$

Combining Eqs. (45) and (57) yields that $|\mathbb{E}[X_k] - E_0| \le \epsilon/2$ whenever $|s| \le w/2$, which happens with probability at least $1 - \delta/3$. So Eq. (28) holds.

Furthermore, Lemma 3.3 also implies that $\int_{s-w}^{s+w} c_2^2 p_0^{-1} h(x) dx \ge \sigma$. Combining this fact and Eq. (44) yields

$$\mathbb{P}[A_i = 1] \ge \frac{p_0\sigma}{2c_2^2 w} \ge \frac{\eta\sigma}{2c_2^2 w} = \frac{\eta}{4c_2^2\sqrt{\ln\left(e\sigma/\epsilon_1\right)}}, \tag{58}$$

where the second step follows from $p_0 \ge \eta$ and the last steps follows from the definition of w. Then we have

$$\mathbb{E}[K] = M\mathbb{P}[A_1 = 1] \ge \frac{M\eta}{4c_2^2\sqrt{\ln\left(e\sigma/\epsilon_1\right)}} = \max(16w^2\epsilon^{-2}\ln\left(6\delta^{-1}\right), 8\ln\left(3\delta^{-1}\right)), \tag{59}$$

where the last step follows from the definition of M. Now we invoke the multiplicative Chernoff bound:

$$\mathbb{P}\left[K \ge \frac{\mathbb{E}[K]}{2}\right] \ge 1 - e^{-\mathbb{E}[K]/8}, \tag{60}$$

and get

$$\mathbb{P}\left[K \ge 8w^2\epsilon^{-2}\ln\left(6\delta^{-1}\right)\right] \ge 1 - \delta/3. \tag{61}$$

So Eq. (29) is proved.

Finally, since $\hat{E}_0 = \frac{1}{K}\sum_{k=1}^K X_k$ where $X_1, X_2, \ldots, X_k$ are i.i.d. and each X_k takes values in $[\tilde{E}_0 - w, \tilde{E}_0 + w]$, by Hoeffding's inequality, we get

$$\mathbb{P}\left[\left|\hat{E}_0 - \mathbb{E}[X_1]\right| \ge \epsilon/2\right] \le 2e^{-\epsilon^2 K/(8w^2)}. \tag{62}$$

If $K \ge 8w^2\epsilon^{-2}\ln\left(6\delta^{-1}\right)$, then the RHS of Eq. (62) is at most $\delta/3$. So Eq. (30) is proved as well.

The theorem is thus proved.

□

Note that Algorithm 2 relies on block-encodings of $f_{a,b;\epsilon'}(H)$ and $g_{\sigma,\xi;\epsilon''}(H)$ (which are approximate threshold and Gaussian functions of H respectively), and their implementation depends on how H is accessed. In the next two sections, we will give concrete implementation of these unitary operations in the block-encoding and Hamiltonian-evolution models respectively, and analyze the costs of the resulting GSEE algorithms in these models.

4 GSEE in the block-encoding model

In this section, we focus on GSEE in the block-encoding model. That is, we assume that there exists a method to implement an $(1, m, 0)$-block-encoding U of H for some integer $m \geq 0$. Our quantum circuits will consist of multiple queries to U and $U^\dagger$ as well as other primitive quantum gates, and act on $\left|0^k\right\rangle |\psi\rangle$ for some integer $k \geq m$.

We will develop a full-fledged version of Algorithm 2 in the block-encoding model. This means that we need to give concrete implementation of the block-encodings of $f_{a,b;\epsilon'}(H)$ and $g_{\sigma,\xi;\epsilon''}(H)$ in Theorem 3.6 in this input model. To accomplish this, we utilize the technique of quantum singular value transformation (QSVT) [42]. This technique allows us to implement any bounded low-degree polynomial of a linear operator A by using a block-encoding of A and its inverse a small number of times. Specifically, we will invoke Theorem 31 of Ref. [42] which states that:

Lemma 4.1. *Suppose U is a $(1, m, 0)$-block-encoding of a Hermitian matrix A. Let $P \in \mathbb{R}[x]$ be a degree-d real polynomial such that $|P(x)| \leq 1$ for all $x \in [-1, 1]$. Then we can implement a $(2, m + 2, 0)$-block-encoding $\tilde{U}$ of $P(A)$ with $2d$ queries of U and $U^\dagger$, and $\mathcal{O}((m + 1)d)$ other primitive quantum gates.*

We will use QSVT to approximately implement the threshold and Gaussian functions of the Hamiltonian H. To this end, we need low-degree polynomial approximations of such functions in the relevant regions. Precisely, for the threshold functions, we have the following result which follows directly from Corollary 6 of Ref. [43] or Lemma 14 of Ref. [42]:

Lemma 4.2. *For every $a, b \in [-1, 1]$, $a < b$ and $\epsilon' \in (0, 1)$, there exists an efficiently-computable real polynomial $f_{a,b;\epsilon'}(x)$ of degree $\mathcal{O}\left(\frac{\log(1/\epsilon')}{b-a}\right)$ such that $1 - \epsilon' \leq f_{a,b;\epsilon'}(x) \leq 1$ for all $x \in [-1, a]$, $|f_{a,b;\epsilon'}(x)| \leq 1$ for all $x \in [a, b]$, and $|f_{a,b;\epsilon'}(x)| \leq \epsilon'$ for all $x \in [b, 1]$.*

Combining Lemma 4.1 and Lemma 4.2 yields:

Corollary 4.3. *Suppose U is a $(1, m, 0)$-block-encoding of a Hermitian matrix A. For every $a, b \in [-1, 1]$, $a < b$ and $\epsilon' \in (0, 1)$, let $f_{a,b;\epsilon'}(x)$ be the polynomial in Lemma 4.2. Then we can implement a $(2, m+2, 0)$-block-encoding V of $f_{a,b;\epsilon'}(A)$ with $\mathcal{O}\left(\frac{\log(1/\epsilon')}{b-a}\right)$ queries of U, $U^\dagger$, and $\mathcal{O}(\frac{(m+1)\log(1/\epsilon')}{b-a})$ other primitive quantum gates.*

Meanwhile, for the Gaussian functions, we have:

Lemma 4.4. *For every $\sigma, \epsilon'' \in (0, 1)$, there exists an efficiently-computable real polynomial $P_{\sigma;\epsilon''}(x)$ of degree $\mathcal{O}(\max(\sigma^{-1}, \sqrt{\log(1/\epsilon'')}) \cdot \sqrt{\log(1/\epsilon'')})$ such that $|P_{\sigma;\epsilon''}(x)| \leq 1$ and $|P_{\sigma;\epsilon''}(x) - e^{-x^2/(2\sigma^2)}| \leq \epsilon''$ for all $x \in [-\pi, \pi]$.*

The proof of Lemma 4.4 is given in Appendix C. This lemma implies:

Corollary 4.5. *For every $\sigma, \epsilon'' \in (0, 1)$ and $\xi \in [-2, 2]$, there exists an efficiently-computable real polynomial $g_{\sigma,\xi;\epsilon''}(x)$ of degree $\mathcal{O}(\max(\sigma^{-1}, \sqrt{\log(1/\epsilon'')}) \cdot \sqrt{\log(1/\epsilon'')})$ such that $|g_{\sigma,\xi;\epsilon''}(x)| \leq 1$ and $|g_{\sigma,\xi;\epsilon''}(x) - e^{-(x-\xi)^2/(2\sigma^2)}| \leq \epsilon''$ for all $x \in [-1, 1]$.*

Combining Lemma 4.1 and Corollary 4.5 yields:

Corollary 4.6. *Suppose U is a $(1, m, 0)$-block-encoding of a Hermitian matrix A. For every $\sigma, \epsilon'' \in (0, 1)$ and $\xi \in [-2, 2]$, let $g_{\sigma,\xi;\epsilon''}(x)$ be the polynomial in Corollary 4.5. Then we can implement a $(2, m+2, 0)$-block-encoding W of $g_{\sigma,\xi;\epsilon''}(A)$ with $\mathcal{O}(\max(\sigma^{-1}, \sqrt{\log(1/\epsilon'')}) \cdot \sqrt{\log(1/\epsilon'')})$ queries of U and $U^\dagger$, and $\mathcal{O}((m+1) \cdot \max(\sigma^{-1}, \sqrt{\log(1/\epsilon'')}) \cdot \sqrt{\log(1/\epsilon'')})$ other primitive quantum gates.*

Equipped with concrete implementation of $f_{a,b;\epsilon'}(H)$ and $g_{\sigma,\xi;\epsilon''}(H)$, we are now ready to state the main result about GSEE in the block-encoding model:

Theorem 4.7. *Suppose H is an n-qubit Hamiltonian with unknown spectral decomposition $H = \sum_{j=0}^{N-1} E_j |E_j\rangle \langle E_j|$, where $N = 2^n$, $-1 \leq E_0 < E_1 \leq E_2 \leq ... \leq E_{N-1} \leq 1$, $E_1 - E_0 \geq \Delta$ for known $\Delta > 0$, and the $|E_j\rangle$'s are orthonormal eigenstates of H. Moreover, suppose we can prepare an n-qubit state $|\psi\rangle = \sum_{j=0}^{N-1} \gamma_j |E_j\rangle$ such that $|\gamma_0|^2 \geq \eta$ for known $\eta \in (0, 1)$. Finally, suppose U is a $(1, m, 0)$-block-encoding of H. Then there exists a quantum algorithm that, for any given $\epsilon > 0$ and $\delta \in (0, 1)$, estimates E_0 within additive error ϵ with probability at least $1 - \delta$ by making use of*

$$\tilde{\mathcal{O}}(\eta^{-1}(\Delta^2 \epsilon^{-2} + \log(\Delta^{-1})) \log(\delta^{-1}))$$

quantum circuits, where each circuit uses

$$\tilde{\mathcal{O}}(\Delta^{-1} \log(\eta^{-1} \Delta \epsilon^{-1}))$$

queries of U and $U^\dagger$, and

$$\tilde{\mathcal{O}}((m+1)\Delta^{-1} \log(\eta^{-1} \Delta \epsilon^{-1}))$$

other primitive quantum gates.

Proof. We estimate the ground state energy of H by running Algorithm 2 with appropriate parameters. Specifically, Corollary 4.3 implies that for arbitrary $a, b \in (-1, 1)$ and $\epsilon' \geq 0$, a $(2, m+2, 0)$-block-encoding of $f_{a,b;\epsilon'}(H)$ can be implemented with $\mathcal{O}\left(\frac{\log(1/\epsilon')}{b-a}\right)$ queries of U and $U^\dagger$, and $\mathcal{O}(\frac{(m+1)\log(1/\epsilon')}{b-a})$ other primitive quantum gates, where $f_{a,b;\epsilon'} : [-1, 1] \to \mathbb{R}$ satisfies that $1 - \epsilon' \leq f_{a,b;\epsilon'}(x) \leq 1$ for all $x \in [-1, a]$, $|f_{a,b;\epsilon'}(x)| \leq 1$ for all $x \in [a, b]$, and $|f_{a,b;\epsilon'}(x)| \leq \epsilon'$ for all $x \in [b, 1]$. Moreover, Corollary 4.6 implies that for arbitrary $\sigma, \epsilon'' \in (0, 1)$ and $\xi \in [-2, 2]$, a $(2, m+2, 0)$-block-encoding of $g_{\sigma,\xi;\epsilon''}(H)$ can be implemented with $\mathcal{O}(\max(\sigma^{-1}\sqrt{\log(1/\epsilon'')}, \log(1/\epsilon'')))$ queries of U and $U^\dagger$, and $\mathcal{O}((m+1)\max(\sigma^{-1}\sqrt{\log(1/\epsilon'')}, \log(1/\epsilon'')))$ other primitive quantum gates, where $g_{\sigma,\xi;\epsilon''} : [-1, 1] \to \mathbb{R}$ satisfies that $|g_{\sigma,\xi;\epsilon''}(x)| \leq 1$ and $|g_{\sigma,\xi;\epsilon''}(x) - e^{-(x-\xi)^2/(2\sigma^2)}| \leq \epsilon''$ for all $x \in [-1, 1]$. Consequently, we can run Algorithm 2 with parameters $(\epsilon, \delta, \Delta, \eta, 2, m+2, 2, m+2)$, and by Theorem 3.6, its output is ϵ-close to E_0 with probability at least $1 - \delta$, as desired.

It remains to analyze the resource cost of ADV_GSEE$(\epsilon, \delta, \Delta, \eta, 2, m+2, 2, m+2)$. If $\epsilon \geq \Delta/8$, then we only need to run BASIC_GSEE $(\epsilon, \delta, \eta, 2, m+2)$. This procedure makes use of

$$\tilde{\mathcal{O}}(\eta^{-1} \log(\epsilon^{-1}) \log(\delta^{-1})) = \tilde{\mathcal{O}}(\eta^{-1} \log(\Delta^{-1}) \log(\delta^{-1}))$$

quantum circuits, where each circuit implements a $(2, m+2, 0)$-block-encoding of $f_{a,b;\epsilon'}(H)$ in which $b - a = \Omega(\epsilon)$ and $\epsilon' = \Omega(\sqrt{\eta})$. We can construct this circuit by using

$$\mathcal{O}\left(\frac{\log(1/\epsilon')}{b-a}\right) = \mathcal{O}(\epsilon^{-1} \log(\eta^{-1})) = \mathcal{O}(\Delta^{-1} \log(\eta^{-1}))$$

queries of U and $U^\dagger$, and

$$\mathcal{O}\left(\frac{(m+1)\log\left(1/\epsilon'\right)}{b-a}\right) = \mathcal{O}((m+1)\epsilon^{-1}\log\left(\eta^{-1}\right)) = \mathcal{O}((m+1)\Delta^{-1}\log\left(\eta^{-1}\right))$$

other primitive quantum gates.

Otherwise, $\epsilon < \Delta/8$. In this case, we need to first run BASIC_GSEE $(w/2, \delta/3, \eta, 2, m+2)$, where

$$w \geq 2\sigma = \Omega\left(\frac{\Delta}{\sqrt{\log\left(\eta^{-1}\Delta\epsilon^{-1}\right)}}\right). \tag{63}$$

As a result, BASIC_GSEE $(w/2, \delta/3, \eta, 2, m+2)$ makes use of

$$\tilde{\mathcal{O}}(\eta^{-1}\log\left(w^{-1}\right)\log\left(\delta^{-1}\right)) = \tilde{\mathcal{O}}(\eta^{-1}\log\left(\Delta^{-1}\log\left(\eta^{-1}\Delta\epsilon^{-1}\right)\right)\log\left(\delta^{-1}\right))$$

quantum circuits, where each circuit implements a $(2, m+2, 0)$-block-encoding of $f_{a,b;\epsilon'}(H)$ in which $b-a = \Omega(w)$ and $\epsilon' = \Omega(\sqrt{\eta})$. We can construct this circuit by using

$$\mathcal{O}\left(\frac{\log\left(1/\epsilon'\right)}{b-a}\right) = \mathcal{O}(w^{-1}\log\left(\eta^{-1}\right)) = \mathcal{O}(\Delta^{-1}\sqrt{\log\left(\eta^{-1}\Delta\epsilon^{-1}\right)}\log\left(\eta^{-1}\right))$$

queries of U and $U^\dagger$, and

$$\mathcal{O}\left(\frac{(m+1)\log\left(1/\epsilon'\right)}{b-a}\right) = \mathcal{O}((m+1)w^{-1}\log\left(\eta^{-1}\right)) = \mathcal{O}((m+1)\Delta^{-1}\sqrt{\log\left(\eta^{-1}\Delta\epsilon^{-1}\right)}\log\left(\eta^{-1}\right))$$

other primitive quantum gates. Moreover, we also need to run

$$M = \tilde{\mathcal{O}}(\eta^{-1}\Delta^2\epsilon^{-2}\log\left(\delta^{-1}\right)) \tag{64}$$

quantum circuits in the **for**-loop, where each circuit implements a $(2, m+2, 0)$-block-encoding of $g_{\sigma,\xi;\epsilon''}(H)$ in which $\sigma^{-1} = \mathcal{O}(\Delta^{-1}\sqrt{\log\left(\eta^{-1}\Delta\epsilon^{-1}\right)})$ and $(\epsilon'')^{-1} = \tilde{\mathcal{O}}(\eta^{-1}\Delta\epsilon^{-1})$. We can build each of these circuits by using

$$\mathcal{O}(\max(\sigma^{-1}\sqrt{\log\left(1/\epsilon''\right)}, \log\left(1/\epsilon''\right))) = \tilde{\mathcal{O}}(\Delta^{-1}\log(\eta^{-1}\Delta\epsilon^{-1}))$$

queries of U and $U^\dagger$, and

$$\mathcal{O}((m+1)\max(\sigma^{-1}\sqrt{\log\left(1/\epsilon''\right)}, \log\left(1/\epsilon''\right))) = \tilde{\mathcal{O}}((m+1)\Delta^{-1}\log(\eta^{-1}\Delta\epsilon^{-1}))$$

other primitive quantum gates.

Overall, regardless of the value of ϵ, Algorithm 2 always makes use of

$$\tilde{\mathcal{O}}(\eta^{-1}(\Delta^2\epsilon^{-2} + \log\left(\Delta^{-1}\right))\log\left(\delta^{-1}\right))$$

quantum circuits, where each circuit uses

$$\tilde{\mathcal{O}}(\Delta^{-1}\log(\eta^{-1}\Delta\epsilon^{-1}))$$

queries of U and $U^\dagger$, and

$$\tilde{\mathcal{O}}((m+1)\Delta^{-1}\log(\eta^{-1}\Delta\epsilon^{-1}))$$

other primitive quantum gates, as claimed.

□

Note that in Theorem 4.7, Δ is merely a lower bound on the true spectral gap $\Delta_{\text{true}} \coloneqq E_1 - E_0$ of the Hamiltonian H. In fact, we can tune it between $\Theta(\epsilon)$ and $\Theta(\Delta_{\text{true}})$, and doing so leads to a smooth interpolation between an algorithm with low circuit depth and an algorithm with Heisenberg-limit scaling of runtime with respect to ϵ:

Corollary 4.8. *Suppose H is an n-qubit Hamiltonian with unknown spectral decomposition $H = \sum_{j=0}^{N-1} E_j |E_j\rangle \langle E_j|$, where $N = 2^n$, $-1 \leq E_0 < E_1 \leq E_2 \leq ... \leq E_{N-1} \leq 1$, $\Delta_{\text{true}} = E_1 - E_0$ is the spectral gap of H, and the $|E_j\rangle$'s are orthonormal eigenstates of H. Moreover, suppose we can prepare an n-qubit state $|\psi\rangle = \sum_{j=0}^{N-1} \gamma_j |E_j\rangle$ such that $|\gamma_0|^2 \geq \eta$ for known $\eta \in (0, 1)$. Finally, suppose U is a $(1, m, 0)$-block-encoding of H. Then for every $\beta \in [0, 1]$, there exists an algorithm that, for any given $\epsilon > 0$ and $\delta \in (0, 1)$, estimates E_0 within additive error ϵ with probability at least $1 - \delta$ by making use of*

$$\tilde{\mathcal{O}}\left(\eta^{-1}\left(\epsilon^{-2+2\beta}\Delta_{\text{true}}^{2-2\beta} + \log\left(\epsilon^{-\beta}\Delta_{\text{true}}^{\beta-1}\right)\right)\log\left(\delta^{-1}\right)\right)$$

quantum circuits, where each circuit uses

$$\tilde{\mathcal{O}}(\epsilon^{-\beta}\Delta_{\text{true}}^{\beta-1}\log(\eta^{-1}\Delta_{\text{true}}^{1-\beta}\epsilon^{\beta-1}))$$

queries of U and $U^\dagger$, and

$$\tilde{\mathcal{O}}((m+1)\epsilon^{-\beta}\Delta_{\text{true}}^{\beta-1}\log(\eta^{-1}\Delta_{\text{true}}^{1-\beta}\epsilon^{\beta-1}))$$

other primitive quantum gates.

Proof. This corollary is the direct consequence of setting $\Delta = \Theta(\epsilon^\beta \Delta_{\text{true}}^{1-\beta})$ for $\beta \in [0, 1]$ in Theorem 4.7. □

In particular, setting $\beta = 0$ or 1 in Corollary 4.8 yields:

- $\Delta = \Theta(\Delta_{\text{true}})$, for which we get an algorithm that makes use of

  $$\tilde{\mathcal{O}}\left(\eta^{-1}\left(\epsilon^{-2}\Delta_{\text{true}}^2 + \log\left(\Delta_{\text{true}}^{-1}\right)\right)\log\left(\delta^{-1}\right)\right)$$

  quantum circuits, where each circuit uses

  $$\tilde{\mathcal{O}}(\Delta_{\text{true}}^{-1}\log(\eta^{-1}\Delta_{\text{true}}\epsilon^{-1}))$$

  queries of U and $U^\dagger$, and

  $$\tilde{\mathcal{O}}((m+1)\Delta_{\text{true}}^{-1}\log(\eta^{-1}\Delta_{\text{true}}\epsilon^{-1}))$$

  other primitive quantum gates; or

- $\Delta = \Theta(\epsilon)$, for which we get an algorithm that makes use of

  $$\tilde{\mathcal{O}}\left(\eta^{-1}\log\left(\epsilon^{-1}\right)\log\left(\delta^{-1}\right)\right)$$

  quantum circuits, where each circuit uses

  $$\tilde{\mathcal{O}}(\epsilon^{-1}\log(\eta^{-1}))$$

  queries of U and $U^\dagger$, and

  $$\tilde{\mathcal{O}}((m+1)\epsilon^{-1}\log(\eta^{-1}))$$

  other primitive quantum gates.

5 GSEE in the Hamiltonian-evolution model

In this section, we focus on GSEE in the Hamiltonian-evolution model. That is, we assume that there exists a method to implement the controlled time evolution of H, i.e. the unitary operation controlled-$e^{\mathbf{i}Ht} \coloneqq |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{\mathbf{i}Ht}$. Each of our quantum circuits will consist of multiple queries to controlled-$e^{\mathbf{i}Ht}$ and controlled-$e^{-\mathbf{i}Ht}$ for some $t = \Theta(1)$ and other primitive quantum gates, and act on $\left|0^k\right\rangle |\psi\rangle$ for some $k \geq 1$.

We will develop a full-fledged version of Algorithm 2 in the Hamiltonian-evolution model. This means that we need to give concrete implementation of the block-encodings of $f_{a,b;\epsilon'}(H)$ and $g_{\sigma,\xi;\epsilon''}(H)$ in Theorem 3.6 in this input model. To this end, we will utilize the technique of quantum eigenvalue transformation of unitary matrices with real polynomials (QET-U) [14]. This technique allows us to implement certain low-degree trigonometric polynomials of $H/2$ by using controlled-$e^{\mathbf{i}H}$ and controlled-$e^{-\mathbf{i}H}$ a small number of times. Specifically, we will invoke Theorem 1 of Ref. [14] which states that:

Lemma 5.1. *For every even real polynomial $F(x)$ of degree d satisfying $|F(x)| \leq 1$ for all $x \in [-1, 1]$, we can implement a $(1, 1, 0)$-block-encoding of $F(\cos(H/2))$ with $\mathcal{O}(d)$ queries of controlled-$e^{\mathbf{i}H}$ and controlled-$e^{-\mathbf{i}H}$, and $\mathcal{O}(d)$ other primitive quantum gates.*

It turns out that both the threshold and Gaussian functions can be approximated by such trigonometric polynomials. Specifically, for the threshold functions, Lemma 6 of Ref. [15] implies:

Lemma 5.2. *For every $a, b \in [0, 1]$, $a < b$ and $\epsilon' \in (0, 1)$, there exists an efficiently-computable even real polynomial $F_{a,b;\epsilon'}(x)$ of degree $\tilde{\mathcal{O}}\left(\frac{\log(1/\epsilon')}{b-a}\right)$ such that $1 - \epsilon' \leq F_{a,b;\epsilon'}(\cos(x/2)) \leq 1$ for all $x \in [0, a]$, $|F_{a,b;\epsilon'}(\cos(x/2))| \leq 1$ for all $x \in [a, b]$, and $|F_{a,b;\epsilon'}(\cos(x/2))| \leq \epsilon'$ for all $x \in [b, 1]$.*

Combining Lemma 5.1 and Lemma 5.2 yields:

Corollary 5.3. *For every $a, b \in [0, 1]$, $a < b$ and $\epsilon' \in (0, 1)$, let $F_{a,b;\epsilon'}(x)$ be the polynomial in Lemma 5.2. Then we can implement a $(1, 1, 0)$-block-encoding of $F_{a,b;\epsilon'}(\cos(H/2))$ with $\tilde{\mathcal{O}}\left(\frac{\log(1/\epsilon')}{b-a}\right)$ queries of controlled-$e^{\mathbf{i}H}$ and controlled-$e^{-\mathbf{i}H}$, and $\mathcal{O}\left(\frac{\log(1/\epsilon')}{b-a}\right)$ other primitive quantum gates.*

Meanwhile, for the Gaussian functions, we have the following result:

Lemma 5.4. *For every $\sigma, \epsilon'' \in (0, 1)$, there exist an efficiently-computable even real polynomial $G_{\sigma;\epsilon''}(x)$ of degree $\mathcal{O}(\sigma^{-1} T \sqrt{\log(1/\epsilon'')})$, where $T = \max(2\pi, \Theta(\sigma\sqrt{\log(1/\epsilon'')}))$, such that $|G_{\sigma;\epsilon''}(\cos(\pi x/T))| \leq 1$ and $|G_{\sigma;\epsilon''}(\cos(\pi x/T)) - e^{-x^2/(2\sigma^2)}| \leq \epsilon''$ for all $x \in [-\pi, \pi]$.*

The proof of Lemma 5.4 is given in Appendix C. Combining Lemma 5.1 and Lemma 5.4 yields:

Corollary 5.5. *For every $\sigma, \epsilon'' \in (0, 1)$, let T and $G_{\sigma,\epsilon''}(x)$ be defined as in Lemma 5.4. Then we can implement a $(1, 1, 0)$-block-encoding of $G_{\sigma;\epsilon''}(\cos(\pi H/T))$ with $\mathcal{O}(\sigma^{-1} T \sqrt{\log(1/\epsilon'')})$ queries of controlled-$e^{\mathbf{i}2\pi H/T}$ and controlled-$e^{-\mathbf{i}2\pi H/T}$, and $\mathcal{O}(\sigma^{-1} T \sqrt{\log(1/\epsilon'')})$ other primitive quantum gates.*

Note that for any $\xi, t \in \mathbb{R}$, we can implement controlled-$e^{-\mathbf{i}(H-\xi I)t}$ by using one controlled-$e^{\mathbf{i}Ht}$ and one Z-rotation gate on the control qubit. Combining this fact and Corollary 5.5, we conclude that:

Corollary 5.6. *For every $\sigma, \epsilon'' \in (0,1)$, let T and $G_{\sigma,\epsilon''}(x)$ be defined as in Lemma 5.4. Then for every $\xi \in [-2,2]$, we can implement a $(1,1,0)$-block-encoding of $G_{\sigma;\epsilon''}(\cos(\pi(H-\xi I)/T))$ with $\mathcal{O}(\sigma^{-1}T\sqrt{\log(1/\epsilon'')})$ queries of controlled-$e^{\mathbf{i}2\pi H/T}$ and controlled-$e^{-\mathbf{i}2\pi H/T}$, and $\mathcal{O}(\sigma^{-1}T\sqrt{\log(1/\epsilon'')})$ other primitive quantum gates.*

Corollaries 5.3 and 5.6 provide the methods to approximately implement the threshold functions and Gaussian functions of H. This enables us to obtain the following result about GSEE in the Hamiltonian-evolution model:

Theorem 5.7. *Suppose H is an n-qubit Hamiltonian with unknown spectral decomposition $H = \sum_{j=0}^{N-1} E_j |E_j\rangle \langle E_j|$, where $N = 2^n$, $0 \leq E_0 < E_1 \leq E_2 \leq ... \leq E_{N-1} \leq 1$, $E_1 - E_0 \geq \Delta$ for known $\Delta > 0$, and the $|E_j\rangle$'s are orthonormal eigenstates of H. Moreover, suppose we can prepare an n-qubit state $|\psi\rangle = \sum_{j=0}^{N-1} \gamma_j |E_j\rangle$ such that $|\gamma_0|^2 \geq \eta$ for known $\eta \in (0,1)$. Then there exists a quantum algorithm that, for any given $\epsilon > 0$ and $\delta \in (0,1)$, estimates E_0 within additive error ϵ with probability at least $1-\delta$ by making use of*

$$\tilde{\mathcal{O}}(\eta^{-1}(\Delta^2\epsilon^{-2} + \log(\Delta^{-1}))\log(\delta^{-1}))$$

quantum circuits, where each circuit uses

$$\tilde{\mathcal{O}}(\Delta^{-1}\log(\eta^{-1}\Delta\epsilon^{-1}))$$

queries of controlled-$e^{\mathbf{i}Ht}$ and controlled-$e^{-\mathbf{i}Ht}$ for some $t = \Theta(1)$, and

$$\tilde{\mathcal{O}}(\Delta^{-1}\log(\eta^{-1}\Delta\epsilon^{-1}))$$

other primitive quantum gates.

Proof. We estimate the ground state energy of H by running Algorithm 2 with appropriate parameters. Specifically, Corollary 5.3 implies that for arbitrary $a, b \in [0,1]$, $a < b$, and $\epsilon' \geq 0$, a $(1,1,0)$-block-encoding of $f_{a,b;\epsilon'}(H) \coloneqq F_{a,b;\epsilon'}(\cos(H/2))$ can be implemented with $\tilde{\mathcal{O}}\left(\frac{\log(1/\epsilon')}{b-a}\right)$ queries of controlled-$e^{\mathbf{i}H}$ and controlled-$e^{-\mathbf{i}H}$, and $\tilde{\mathcal{O}}(\frac{\log(1/\epsilon')}{b-a})$ other primitive quantum gates, where $f_{a,b;\epsilon'}$ satisfies that $1-\epsilon' \leq f_{a,b;\epsilon'}(x) \leq 1$ for all $x \in [0,a]$, $|f_{a,b;\epsilon'}(x)| \leq 1$ for all $x \in [a,b]$, and $|f_{a,b;\epsilon'}(x)| \leq \epsilon'$ for all $x \in [b,1]$. Moreover, Corollary 5.6 implies that for arbitrary $\sigma, \epsilon'' \in (0,1)$ and $\xi \in [-2,2]$, there exists $T = \max(2\pi, \Theta(\sigma\sqrt{\log(1/\epsilon'')}))$ such that a $(1,1,0)$-block-encoding of $g_{\sigma,\xi;\epsilon''}(H) \coloneqq G_{\sigma,\epsilon''}(\cos(\pi(H-\xi I)/T))$ can be implemented with $\mathcal{O}(\sigma^{-1}T\sqrt{\log(1/\epsilon'')})$ queries of controlled-$e^{\mathbf{i}2\pi H/T}$, controlled-$e^{-\mathbf{i}2\pi H/T}$, and $\mathcal{O}(\sigma^{-1}T\sqrt{\log(1/\epsilon'')})$ other primitive quantum gates, where $g_{\sigma,\xi;\epsilon''}$ satisfies that $|g_{\sigma,\xi;\epsilon''}(x)| \leq 1$ and $|g_{\sigma,\xi;\epsilon''}(x) - e^{-(x-\xi)^2/(2\sigma^2)}| \leq \epsilon''$ for all $x \in [-1,1]$. Consequently, we can run Algorithm 2 with parameters $(\epsilon, \delta, \Delta, \eta, 1, 1, 1, 1)$, and by Theorem 3.6, its output is ϵ-close to E_0 with probability at least $1-\delta$, as desired.

It remains to analyze the resource cost of ADV_GSEE$(\epsilon, \delta, \Delta, \eta, 1, 1, 1, 1)$. If $\epsilon \geq \Delta/8$, then we only need to run BASIC_GSEE $(\epsilon, \delta, \eta, 1, 1)$. This procedure makes use of

$$\tilde{\mathcal{O}}(\eta^{-1}\log(\epsilon^{-1})\log(\delta^{-1})) = \tilde{\mathcal{O}}(\eta^{-1}\log(\Delta^{-1})\log(\delta^{-1}))$$

quantum circuits, where each circuit implements a $(1,1,0)$-block-encoding of $f_{a,b;\epsilon'}(H)$ in which $b-a = \Omega(\epsilon)$ and $\epsilon' = \Omega(\sqrt{\eta})$. We can build this circuit by using

$$\tilde{\mathcal{O}}\left(\frac{\log(1/\epsilon')}{b-a}\right) = \tilde{\mathcal{O}}(\epsilon^{-1}\log(\eta^{-1})) = \tilde{\mathcal{O}}(\Delta^{-1}\log(\eta^{-1}))$$

queries of controlled-$e^{\mathbf{i}H}$ and controlled-$e^{-\mathbf{i}H}$, and

$$\tilde{\mathcal{O}}\left(\frac{\log\left(1/\epsilon'\right)}{b-a}\right)=\tilde{\mathcal{O}}(\epsilon^{-1}\log\left(\eta^{-1}\right))=\tilde{\mathcal{O}}(\Delta^{-1}\log\left(\eta^{-1}\right))$$

other primitive quantum gates.

Otherwise, $\epsilon<\Delta/8$. In this case, we need to first run BASIC_GSEE $(w/2,\delta/3,\eta,1,1)$, where

$$w\geq 2\sigma=\Omega\left(\frac{\Delta}{\sqrt{\log\left(\eta^{-1}\Delta\epsilon^{-1}\right)}}\right). \tag{65}$$

As a result, BASIC_GSEE $(w/2,\delta/3,\eta,1,1)$ makes use of

$$\tilde{\mathcal{O}}(\eta^{-1}\log\left(w^{-1}\right)\log\left(\delta^{-1}\right))=\tilde{\mathcal{O}}(\eta^{-1}\log\left(\Delta^{-1}\log\left(\eta^{-1}\Delta\epsilon^{-1}\right)\right)\log\left(\delta^{-1}\right))$$

quantum circuits, where each circuit implements a $(1,1,0)$-block-encoding of $f_{a,b;\epsilon'}(H)$ in which $b-a=\Omega(w)$ and $\epsilon'=\Omega(\sqrt{\eta})$. We can build this circuit by using

$$\tilde{\mathcal{O}}\left(\frac{\log\left(1/\epsilon'\right)}{b-a}\right)=\tilde{\mathcal{O}}(w^{-1}\log\left(\eta^{-1}\right))=\tilde{\mathcal{O}}(\Delta^{-1}\sqrt{\log\left(\eta^{-1}\Delta\epsilon^{-1}\right)}\log\left(\eta^{-1}\right))$$

queries of controlled-$e^{\mathbf{i}H}$ and controlled-$e^{-\mathbf{i}H}$, and

$$\tilde{\mathcal{O}}\left(\frac{\log\left(1/\epsilon'\right)}{b-a}\right)=\tilde{\mathcal{O}}(w^{-1}\log\left(\eta^{-1}\right))=\tilde{\mathcal{O}}(\Delta^{-1}\sqrt{\log\left(\eta^{-1}\Delta\epsilon^{-1}\right)}\log\left(\eta^{-1}\right))$$

other primitive quantum gates. Moreover, we also need to run

$$M=\tilde{\mathcal{O}}(\eta^{-1}\Delta^{2}\epsilon^{-2}\log\left(\delta^{-1}\right)) \tag{66}$$

quantum circuits in the **for**-loop, where each circuit implements a $(1,1,0)$-block-encoding of $g_{\sigma,\xi;\epsilon''}(H)$ in which $\sigma^{-1}=\mathcal{O}(\Delta^{-1}\sqrt{\log\left(\eta^{-1}\Delta\epsilon^{-1}\right)})$ and $(\epsilon'')^{-1}=\tilde{\mathcal{O}}(\eta^{-1}\Delta\epsilon^{-1})$. Furthermore, one can verify that $\sigma\sqrt{\log\left(1/\epsilon''\right)}=\mathcal{O}(\Delta)=\mathcal{O}(1)$ by our choice of parameters. This implies that we can construct each circuit in the **for**-loop by using

$$\mathcal{O}(\sigma^{-1}T\sqrt{\log\left(1/\epsilon''\right)})=\tilde{\mathcal{O}}(\Delta^{-1}\log(\eta^{-1}\Delta\epsilon^{-1}))$$

queries of controlled-$e^{\mathbf{i}2\pi H/T}$ and controlled-$e^{-\mathbf{i}2\pi H/T}$ for $T=\max(2\pi,\Theta(\sigma\sqrt{\log\left(1/\epsilon''\right)}))=\Theta(1)$, and

$$\mathcal{O}(\sigma^{-1}T\sqrt{\log\left(1/\epsilon''\right)})=\tilde{\mathcal{O}}(\Delta^{-1}\log(\eta^{-1}\Delta\epsilon^{-1}))$$

other primitive quantum gates.

Overall, regardless of the value of ϵ, Algorithm 2 always makes use of

$$\tilde{\mathcal{O}}(\eta^{-1}(\Delta^{2}\epsilon^{-2}+\log\left(\Delta^{-1}\right))\log\left(\delta^{-1}\right))$$

quantum circuits, where each circuit uses

$$\tilde{\mathcal{O}}(\Delta^{-1}\log(\eta^{-1}\Delta\epsilon^{-1}))$$

queries of controlled-$e^{\mathbf{i}Ht}$ and controlled-$e^{-\mathbf{i}Ht}$ for some $t=\Theta(1)$, and

$$\tilde{\mathcal{O}}(\Delta^{-1}\log(\eta^{-1}\Delta\epsilon^{-1}))$$

other primitive quantum gates, as claimed.

□

Note again that Δ in Theorem 5.7 is just a lower bound on the true spectral gap $\Delta_{\text{true}} = E_1 - E_0$ of the Hamiltonian H. By tuning it between $\Theta(\epsilon)$ and $\Theta(\Delta_{\text{true}})$, we obtain a class of GSEE algorithms that smoothly transition from having low circuit depth to having Heisenberg-limit scaling of runtime with respect to ϵ:

Corollary 5.8. *Suppose H is an n-qubit Hamiltonian with unknown spectral decomposition $H = \sum_{j=0}^{N-1} E_j |E_j\rangle \langle E_j|$, where $N = 2^n$, $0 \leq E_0 < E_1 \leq E_2 \leq ... \leq E_{N-1} \leq 1$, $\Delta_{\text{true}} = E_1 - E_0$ is the spectral gap of H, and the $|E_j\rangle$'s are orthonormal eigenstates of H. Moreover, suppose we can prepare an n-qubit state $|\psi\rangle = \sum_{j=0}^{N-1} \gamma_j |E_j\rangle$ such that $|\gamma_0|^2 \geq \eta$ for known $\eta \in (0, 1)$. Then for every $\beta \in [0, 1]$, there exists an algorithm that, for any given $\epsilon > 0$ and $\delta \in (0, 1)$, estimates E_0 within additive error ϵ with probability at least $1 - \delta$ by making use of*

$$\tilde{\mathcal{O}}\left(\eta^{-1}\left(\epsilon^{-2+2\beta}\Delta_{\text{true}}^{2-2\beta} + \log\left(\epsilon^{-\beta}\Delta_{\text{true}}^{\beta-1}\right)\right)\log\left(\delta^{-1}\right)\right)$$

quantum circuits, where each circuit uses

$$\tilde{\mathcal{O}}(\epsilon^{-\beta}\Delta_{\text{true}}^{-1+\beta}\log(\eta^{-1}\Delta_{\text{true}}^{1-\beta}\epsilon^{\beta-1}))$$

queries of controlled-$e^{\mathbf{i}Ht}$ and controlled-$e^{-\mathbf{i}Ht}$ for some $t = \Theta(1)$, and

$$\tilde{\mathcal{O}}(\epsilon^{-\beta}\Delta_{\text{true}}^{-1+\beta}\log(\eta^{-1}\Delta_{\text{true}}^{1-\beta}\epsilon^{\beta-1}))$$

other primitive quantum gates.

Proof. This corollary is the direct consequence of setting $\Delta = \Theta(\epsilon^{\beta}\Delta_{\text{true}}^{1-\beta})$ for $\beta \in [0, 1]$ in Theorem 5.7. □

In particular, setting $\beta = 0$ or 1 yields:

- $\Delta = \Theta(\Delta_{\text{true}})$, for which we get an algorithm that makes use of $\tilde{\mathcal{O}}(\eta^{-1}\epsilon^{-2}\Delta_{\text{true}}^2\log\left(\delta^{-1}\right))$ quantum circuits, where each circuit uses $\tilde{\mathcal{O}}(\Delta_{\text{true}}^{-1}\log(\eta^{-1}\Delta_{\text{true}}\epsilon^{-1}))$ queries of controlled-$e^{\mathbf{i}Ht}$ and controlled-$e^{-\mathbf{i}Ht}$ for some $t = \Theta(1)$, and $\tilde{\mathcal{O}}(\Delta_{\text{true}}^{-1}\log(\eta^{-1}\Delta_{\text{true}}\epsilon^{-1}))$ other primitive quantum gates; or
- $\Delta = \Theta(\epsilon)$, for which we get an algorithm that makes use of $\tilde{\mathcal{O}}(\eta^{-1}\log\left(\epsilon^{-1}\right)\log\left(\delta^{-1}\right))$ quantum circuits, where each circuit uses $\tilde{\mathcal{O}}(\epsilon^{-1}\log(\eta^{-1}))$ queries of controlled-$e^{\mathbf{i}Ht}$ and controlled-$e^{-\mathbf{i}Ht}$ for some $t = \Theta(1)$, and $\tilde{\mathcal{O}}(\epsilon^{-1}\log(\eta^{-1}))$ other primitive quantum gates.

6 Certifying the correctness of the energy estimate

In this section we introduce a method for certifying that the energy estimate output by Algorithm 2 is correct. This may be favorable in cases where we are not confident in the correctness of Δ, the lower bound on the energy gap. Furthermore, in the case that the input state has additional structure (i.e. has sufficiently small weight on low energy excited states) our algorithm can provide a correct estimate for evolution times that are arbitrarily smaller than the spectral gap. To illustrate this point, consider that the input state is given by the ground state. In this case, any constant evolution time can be used to obtain a reliable estimate. In both of these cases, it may be desirable to assume that the algorithm settings are correct, run Algorithm 2, and then certify the result. This way we do not need to necessarily trust our estimate of the spectral gap and can profit from better input states.

We describe a procedure that achieves exactly that. That is, given a lower-bound on the overlap of the input state with the ground state, access to samples from the distribution $p * n_\sigma$, an error tolerance $\epsilon > 0$ and an estimate of the ground state $\hat{E}_0$, we will be able to certify that $|E_0 - \hat{E}_0| \leq \epsilon$. It is also important to stress once more what our test *will not achieve*, namely certifying a spectral gap estimate. It only certifies the ground state energy estimate. However, we will show that if the evolution time is proportional to the inverse of the spectral gap, our test will accept the estimate with high probability. It is also worth mentioning that our test comes at the cost of a higher sample complexity than the estimation procedure itself, namely $\tilde{\mathcal{O}}(\epsilon^{-4}\eta^{-3})$.

We give an intuitive explanation of how the verification works in Fig. 6 and explain the protocol in detail in Appendix D. But it essentially relies on two observations. The first, formalized in Lemma D.3, is that if the density of $p * n_\sigma$ has a well-defined peak around our ground state energy estimate, then $p * n_\sigma$ conditioned on an interval of size $\sim \sigma$ around $\hat{E}_0$ will again approximately have the moments of a convex combination of Gaussians. Furthermore, we then show that the variance of the random variable conditioned on this interval will exceed a given threshold if the estimate is incorrect. Finally, we show that if we evolve by a time proportional to the inverse of the spectral gap, the test will accept with high probability. We summarize the certification algorithm below:

Theorem 6.1. *Under the same conditions as Theorem 3.6, there exists an algorithm that, given arbitrary $\sigma > 0$ and $\hat{E}_0 \in [E_0 - \sigma, E_0 + \sigma]$ (but no bound on the spectral gap Δ_{true}), makes*

$$\tilde{\mathcal{O}}(\eta^{-3}\epsilon^{-4}\log\left(\delta^{-1}\right)) \tag{67}$$

uses of the block-encoding of $g_{\sigma,\xi;\epsilon''}(H)$ in expectation and generates a new estimate $\hat{E}_0'$ of E_0. If $|\hat{E}_0' - E_0| \geq 4\epsilon$, it and rejects $\hat{E}_0'$ with probability at least $1 - \delta$. If it accepts, then $|\hat{E}_0' - E_0| \leq \epsilon$, with probability of incorrect acceptance at most δ. Furthermore, there exists some

$$\sigma_0 = \Omega\left(\frac{\Delta_{\text{true}}}{\log\left(\epsilon^{-1}\eta^{-1}\right)}\right) \tag{68}$$

such that if $\sigma \leq \sigma_0$, then this algorithm outputs an estimate $\hat{E}_0'$ that it accepts with probability $1 - \delta$ and is such that $|\hat{E}_0' - E_0| \leq \epsilon$.

We defer all the proofs and descriptions of the algorithm to Appendix D. Note that the minimal time for which the test accepts, in Eq. (68), is of the same order as when our algorithm is guaranteed to succeed. Unfortunately, the sample complexity of the test is higher than that of the algorithm at this depth. Thus, the test is only advantageous when we do not wish to trust our estimate of the spectral gap or have reasons to believe that the input is such that we can obtain a good estimate with smaller circuit depths. Our algorithm will be based on empirically estimating the variance of a random variable up to a certain precision. As is common with such protocols, depending on the number of samples the empirical estimate will have a tolerated error up to a certain failure probability. It is noteworthy that our protocol consistently provides accurate rejections when given variance estimates within the tolerated error. Therefore, the parameter δ quantifies the probability of the estimates deviating beyond the acceptable threshold for verification, and as a result, we cannot give guarantees on its performance when given wrong estimates.

Furthermore, the test can be combined with a binary search procedure to determine σ_0, the largest σ for which the test accepts. Although this would yield reliable ground state energy estimates at depths that are potentially $\ll \Delta_{\text{true}}$, the sample complexity of such an algorithm would be $\tilde{\mathcal{O}}(\epsilon^{-4}\eta^{-3})$. However, this might still be advantageous if we wish to optimize the maximal circuit depth.

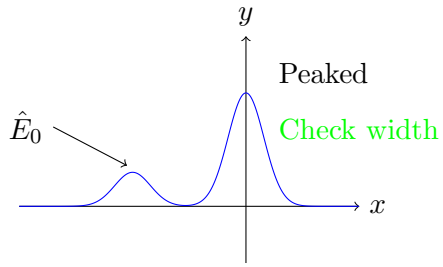

Figure 2: Example of density of mixture of Gaussians. In this case, the density quickly goes to 0 around our estimate $\hat{E}_0$. We then continue to check the variance conditioned around $\hat{E}_0$ to see if we accept the estimate.

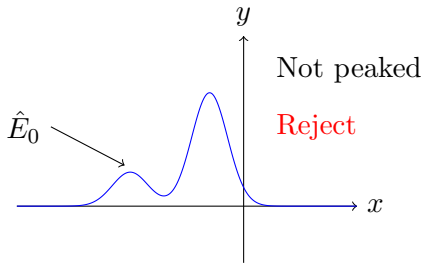

Figure 3: Example of density of mixture of Gaussians. In this case, the density quickly *does not* go to 0 quickly around $\hat{E}_0$. We reject the estimate.

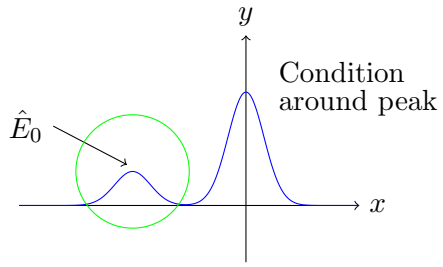

Figure 4: As the distribution is peaked around $\hat{E}_0$, we condition around it in the next step of the verification.

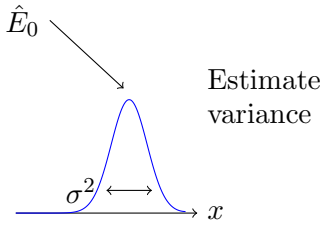

Figure 5: We estimate the variance of the random variable conditioned around $\hat{E}_0$. If it is close to σ^2, we accept it. If it differs from σ^2 by $\sim \eta\epsilon^2$, we reject it.

7 Conclusion and outlook

To summarize, we have introduced a quantum algorithm for estimating the ground state energy of a Hamiltonian using fewer operations per circuit than previous methods. This makes the algorithm more suitable for early fault-tolerant quantum computation. This algorithm relies on a new way to perform rejection sampling on quantum computers. We have shown how it works with both the block-encoding and Hamiltonian-evolution input models of the Hamiltonian. The circuit depth required by our algorithm is governed by the spectral gap of the Hamiltonian rather than the target accuracy, and the runtime of our algorithm is smaller than that of previous algorithms with similar circuit depths. Furthermore, this algorithm can be tuned to trade circuit depth for runtime, and its runtime can be also reduced by parallelizing the sample collection across multiple quantum devices (i.e. trading time resources for space resources). All of these features make our algorithm a great candidate for realizing quantum advantage for industrially-relevant problems on early fault-tolerant quantum computers [33, 25, 15, 14, 39, 47, 40, 48].

In addition, we also devised the first method for certifying ground state energy estimates in the regime of circuit depth $\tilde{\mathcal{O}}(\Delta_{\text{true}}^{-1})$. Although the sample complexity of the current method is high, it can provide certifiably correct estimates of the ground state energy from arbitrarily small circuit depth depending on the properties of the initial state.

There are several related research directions that are worth further exploration:

- In this work, we have focused on GSEE in the regime of circuit depth $\geq 1/\Delta_{\text{true}}$ (recall that Δ_{true} is the spectral gap of H around E_0). It remains unclear how hard this problem is to solve when the circuit depth (accounting for both the state preparation and estimation subroutines) is $\ll 1/\Delta_{\text{true}}$? Is it possible to devise an algorithm with reasonable runtime (e.g. $\text{poly}\,(1/\epsilon, 1/p_0)$) in this setting?

- In many applications, one needs to estimate ground state properties beyond their energies. So far, this problem was not extensively studied in the low-depth regime. One exception is Ref. [39] which develops a low-depth algorithm for estimating the expectation value of an observable O with respect to the ground state of a Hamiltonian H. This problem is referred to as ground state property estimation (GSPE). In order to estimate $\langle E_0| O |E_0\rangle$ within additive error ϵ with high probability, the algorithm of Ref. [39] requires circuit depth $\tilde{\mathcal{O}}(\Delta^{-1}\,\text{polylog}\left(\epsilon^{-1}p_0^{-1}\right))$ and runtime $\tilde{\mathcal{O}}(\Delta^{-1}\epsilon^{-2}p_0^{-2})$ (where Δ and p_0 are defined as before). Note that this runtime is worse than our runtime for ground state energy estimation. Since GSEE can be viewed as a special case of GSPE in which $O = H$, can we apply the techniques in this paper to improve the efficiency of the latter?

- Can we design more efficient methods for certifying the correctness of ground state energy estimates? Our current certifcation method has a quadratically worse sample complexity than our GSEE algorithm. If we could devise a certification method with the same sample complexity as the GSEE algorithm, we could run binary search to find the minimal evolution time required to obtain reliable estimates on the ground state energy without extra assumptions.

- Finally, we expect that our rejection sampling method could find more applications beyond GSEE. In many applications, one needs to generate samples from a continuous distribution X with probability density $\nu(x)$. The traditional approach is to prepare a quantum state that approximates $\int_\Omega \sqrt{\nu(x)}\,|x\rangle\,dx$ and then to measure it. But this approach has two shortcomings. First, the quantum circuit for creating this state might be difficult to implement. Second, one needs to discretize the target distribution, and the number of grid points (= the dimension of the resulting system) depends on the desired accuracy. On the other hand, our approach does not rely on this discretization, as the sample value is encoded into the circuit parameter. Furthermore, our circuit might be easier to implement than the one in the direct approach and might require $O(1)$ qubits regardless of the target accuracy. Provided that rejection sampling is efficient (i.e. not too many raw samples are discarded), our method could accelerate this central component of many algorithms or make it more amenable to near-term implementation.

We have contributed to the growing body of work developing quantum algorithms for early fault-tolerant quantum computers. Our hope is that these results will inspire further exploration to discover the quantum algorithms that will deliver advantage first.

References

[1] 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.

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

[3] Vincent E Elfving, Benno W Broer, Mark Webber, Jacob Gavartin, Mathew D Halls, K Patrick Lorton, and A Bochevarov. How will quantum computers provide an industrially relevant computational advantage in quantum chemistry? *arXiv preprint arXiv:2009.12472*, 2020.

[4] 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. *Proceedings of the National Academy of Sciences*, 119(38):e2203533119, 2022.

[5] A Yu Kitaev. Quantum measurements and the abelian stabilizer problem. *arXiv preprint quant-ph/9511026*, 1995.

[6] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002.

[7] Ryan Babbush, Jarrod McClean, Dave Wecker, Alán Aspuru-Guzik, and Nathan Wiebe. Chemical basis of trotter-suzuki errors in quantum chemistry simulation. *Physical Review A*, 91(2):022311, 2015.

[8] 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.

[9] Kenji Sugisaki, Shigeaki Nakazawa, Kazuo Toyota, Kazunobu Sato, Daisuke Shiomi, and Takeji Takui. Quantum chemistry on quantum computers: A method for preparation of multiconfigurational wave functions on quantum computers without performing post-hartree–fock calculations. *ACS central science*, 5(1):167–175, 2018.

[10] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C Benjamin, and Xiao Yuan. Quantum computational chemistry. *Reviews of Modern Physics*, 92(1):015003, 2020.

[11] Emanuel Knill, Gerardo Ortiz, and Rolando D. Somma. Optimal quantum measurements of expectation values of observables. *Phys. Rev. A*, 75:012328, Jan 2007.

[12] Daniel Nagaj, Pawel Wocjan, and Yong Zhang. Fast amplification of qma. *arXiv preprint arXiv:0904.1549*, 2009.

[13] David Poulin and Pawel Wocjan. Sampling from the thermal quantum gibbs state and evaluating partition functions with a quantum computer. *Physical review letters*, 103(22):220502, 2009.

[14] 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. *PRX Quantum*, 3:040305, Oct 2022.

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

[16] Kianna Wan, Mario Berta, and Earl T. Campbell. Randomized quantum algorithm for statistical phase estimation. *Phys. Rev. Lett.*, 129:030503, Jul 2022.

[17] 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.

[18] Brendon L Higgins, Dominic W Berry, Stephen D Bartlett, Howard M Wiseman, and Geoff J Pryde. Entanglement-free heisenberg-limited phase estimation. *Nature*, 450(7168):393–396, 2007.

[19] 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.

[20] Jay Gambetta. Quantum-centric supercomputing: The next wave of computing, 2022.

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

[22] 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.

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

[24] Lennart Bittel and Martin Kliesch. Training variational quantum algorithms is NP-hard. *Physical Review Letters*, 127(12), September 2021.

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

[26] Dax Enshan Koh, Guoming Wang, Peter D. Johnson, and Yudong Cao. Foundations for bayesian inference with engineered likelihood functions for robust amplitude estimation. *Journal of Mathematical Physics*, 63(5):052202, 2022.

[27] 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.

[28] Daniel Stilck Franca and Raul Garcia-Patron. Limitations of optimization algorithms on noisy quantum devices. *Nature Physics*, 17(11):1221–1227, October 2021.

[29] Giacomo De Palma, Milad Marvian, Cambyse Rouzé, and Daniel Stilck França. Limitations of variational quantum algorithms: A quantum optimal transport approach. *PRX Quantum*, 4(1), January 2023.

[30] Yihui Quek, Daniel Stilck França, Sumeet Khatri, Johannes Jakob Meyer, and Jens Eisert. Exponentially tighter bounds on limitations of quantum error mitigation, 2023.

[31] Samson Wang, Enrico Fontana, M. Cerezo, Kunal Sharma, Akira Sone, Lukasz Cincio, and Patrick J. Coles. Noise-induced barren plateaus in variational quantum algorithms. *Nature Communications*, 12(1), November 2021.

[32] Marco Fellous-Asiani, Jing Hao Chai, Robert S Whitney, Alexia Auffèves, and Hui Khoon Ng. Limitations in quantum computing from resource constraints. *PRX Quantum*, 2(4):040335, 2021.

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

[34] Javier Alcazar, Andrea Cadarso, Amara Katabarwa, Marta Mauri, Borja Peropadre, Guoming Wang, and Yudong Cao. Quantum algorithm for credit valuation adjustments. *New Journal of Physics*, 24(2):023036, 2022.

[35] Tudor Giurgica-Tiron, Iordanis Kerenidis, Farrokh Labib, Anupam Prakash, and William Zeng. Low depth algorithms for quantum amplitude estimation. *Quantum*, 6:745, 2022.

[36] 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.

[37] Tudor Giurgica-Tiron, Sonika Johri, Iordanis Kerenidis, Jason Nguyen, Neal Pisenti, Anupam Prakash, Ksenia Sosnova, Ken Wright, and William Zeng. Low-depth amplitude estimation on a trapped-ion quantum computer. *Physical Review Research*, 4(3):033034, 2022.

[38] Yohichi Suzuki, Shumpei Uno, Rudy Raymond, Tomoki Tanaka, Tamiya Onodera, and Naoki Yamamoto. Amplitude estimation without phase estimation. *Quantum Information Processing*, 19:1–17, 2020.

[39] Ruizhe Zhang, Guoming Wang, and Peter Johnson. Computing Ground State Properties with Early Fault-Tolerant Quantum Computers. *Quantum*, 6:761, July 2022.

[40] Guoming Wang, Daniel Stilck-França, Ruizhe Zhang, Shuchen Zhu, and Peter D Johnson. Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision. *arXiv preprint arXiv:2209.06811*, 2022.

[41] Zhiyan Ding and Lin Lin. Even shorter quantum circuit for phase estimation on early fault-tolerant quantum computers with applications to ground-state energy estimation. *arXiv preprint arXiv:2211.11973*, 2022.

[42] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and beyond: Exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, page 193–204, New York, NY, USA, 2019. Association for Computing Machinery.

[43] Guang Hao Low and Isaac L. Chuang. Hamiltonian simulation by uniform spectral amplification, 2017.

[44] 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.

[45] Joonho Lee, Dominic W Berry, Craig Gidney, William J Huggins, Jarrod R McClean, Nathan Wiebe, and Ryan Babbush. Even more efficient quantum computations of chemistry through tensor hypercontraction. *PRX Quantum*, 2(3):030305, 2021.

[46] Ignacio Loaiza, Alireza Marefat Khah, Nathan Wiebe, and Artur F Izmaylov. Reducing molecular electronic hamiltonian simulation cost for linear combination of unitaries approaches. *arXiv preprint arXiv:2208.08272*, 2022.

[47] Guoming Wang, Sukin Sim, and Peter D. Johnson. State Preparation Boosters for Early Fault-Tolerant Quantum Computation. *Quantum*, 6:829, October 2022.

[48] Samson Wang, Sam McArdle, and Mario Berta. Qubit-efficient randomized quantum algorithms for linear algebra. *arXiv preprint arXiv:2302.01873*, 2023.

[49] Sushant Sachdeva, Nisheeth K Vishnoi, et al. Faster algorithms via approximation theory. *Foundations and Trends® in Theoretical Computer Science*, 9(2):125–210, 2014.

A Proofs of Lemma 3.2 and Lemma 3.3

Proof of Lemma 3.2. Let us focus on the case $s \geq 0$ here. The other case $s < 0$ can be handled similarly.

We will use the following bound on the complementary error function (erfc). For any $z > 0$,

$$\operatorname{erfc} z = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-t^2} dt = \frac{2}{\pi} \int_0^{\pi/2} e^{-z^2/\sin^2\theta} d\theta \leq \frac{2}{\pi} \int_0^{\pi/2} e^{-z^2} d\theta = e^{-z^2}. \tag{69}$$

It implies that

$$\int_0^z e^{-t^2} dt = \int_0^\infty e^{-t^2} dt - \int_z^\infty e^{-t^2} dt \geq \frac{\sqrt{\pi}}{2} \left(1 - e^{-z^2}\right). \tag{70}$$

Now by the definition of w and the assumption $0 \leq s \leq w/2$, we have $w + s \geq w - s \geq w/2 > 0$. Consequently, we get

$$\begin{aligned}
\int_{s-w}^{s+w} e^{-x^2/\sigma^2} dx &\geq 2 \int_0^{w-s} e^{-x^2/\sigma^2} dx && (71)\\
&= 2\sigma \int_0^{(w-s)/\sigma} e^{-t^2} dt && (72)\\
&\geq \sqrt{\pi}\sigma \left(1 - e^{-(w-s)^2/\sigma^2}\right) && (73)\\
&\geq \sqrt{\pi}\sigma \left(1 - e^{-w^2/(4\sigma^2)}\right) && (74)\\
&\geq \sqrt{\pi}\sigma \left(1 - \frac{\epsilon}{e\sigma}\right) && (75)\\
&\geq \sqrt{\pi}\sigma \left(1 - e^{-1}\right) && (76)\\
&\geq 1.12\sigma, && (77)
\end{aligned}$$

where the sixth step follows from $0 < \epsilon \leq \sigma$. Meanwhile, a direct calculation indicates

$$\begin{aligned}
\int_{s-w}^{s+w} x e^{-x^2/\sigma^2} dx &= \frac{\sigma^2}{2} \left(e^{-(w-s)^2/\sigma^2} - e^{-(w+s)^2/\sigma^2}\right) && (78)\\
&\leq \frac{\sigma^2}{2} e^{-(w-s)^2/\sigma^2} && (79)\\
&= \frac{\sigma^2}{2} e^{-w^2/(4\sigma^2)} && (80)\\
&\leq \frac{\sigma\epsilon}{2e}. && (81)
\end{aligned}$$

The lemma is thus proved. □

Proof of Lemma 3.3. As in the proof of Lemma 3.2, we will focus on the case $s \geq 0$ here, and the other case $s < 0$ can be handled similarly.

On one hand, we have

$$\int_{s-w}^{s+w} g(x)dx \geq \int_{s-w}^{s+w} e^{-x^2/\sigma^2}dx - \int_{s-w}^{s+w} |e^{-x^2/\sigma^2} - g(x)|dx \tag{82}$$

$$\geq 1.12\sigma - 2w \cdot \frac{0.03\epsilon}{\sigma \ln(e\sigma/\epsilon)} \tag{83}$$

$$= 1.12\sigma - \frac{0.12\epsilon}{\sqrt{\ln(e\sigma/\epsilon)}} \tag{84}$$

$$\geq 1.12\sigma - 0.12\epsilon \tag{85}$$

$$\geq \sigma, \tag{86}$$

where the first step follows from triangle inequality, the second step follows from Lemma 3.2 and the assumption on $g(x)$, the third step follows from the definition of w, and the fourth and fifth steps follow from $0 < \epsilon \leq \sigma$.

On the other hand, we also have

$$\left|\int_{s-w}^{s+w} xg(x)dx\right| \leq \left|\int_{s-w}^{s+w} xe^{-x^2/\sigma^2}dx\right| + \left|\int_{s-w}^{s+w} x(e^{-x^2/\sigma^2} - g(x))dx\right| \tag{87}$$

$$\leq \frac{\sigma\epsilon}{2e} + 2w \cdot \frac{3w}{2} \cdot \frac{0.03\epsilon}{\sigma \ln(e\sigma/\epsilon)} \tag{88}$$

$$\leq \frac{\sigma\epsilon}{2e} + 0.36\sigma\epsilon \tag{89}$$

$$\leq 0.55\sigma\epsilon, \tag{90}$$

where the first step follows from triangle inequality, the second step follows from Lemma 3.2, the assumption on $g(x)$ and the fact that $|x| \leq w + s \leq 3w/2$ for all $x \in [s-w, s+w]$, the fourth step follows from the definition of w.

The lemma is thus proved.

□

B Proof of Eq. (46)

We continue to use the notation in the proof of Theorem 3.6. Recall that $E_0 - \xi \in [-3w/2, 3w/2]$ and $E_j - E_0 \geq \Delta$ for all $j \geq 1$. So we have

$$E_j - \xi = (E_j - E_0) + (E_0 - \xi) \tag{91}$$

$$\geq \Delta - \frac{3w}{2} \tag{92}$$

$$= \Delta - 3\sigma\sqrt{\ln(e\sigma/\epsilon_1)} \tag{93}$$

$$= \Delta\left(1 - \frac{3\sqrt{\ln(e\sigma/\epsilon_1)}}{5\sqrt{\ln(ec_2)}\sqrt{\ln(e\eta^{-1}\Delta\epsilon^{-1})}}\right). \tag{94}$$

Meanwhile, by the definition of σ and ϵ'', we have

$$\sigma\sqrt{\ln\left(0.5c_2^2/\epsilon''\right)} = \frac{\Delta\sqrt{\ln\left(200c_2^2\sigma\ln(e\sigma/\epsilon_1)/(3\epsilon_1\eta)\right)}}{5\sqrt{\ln(ec_2)}\sqrt{\ln(e\eta^{-1}\Delta\epsilon^{-1})}}. \tag{95}$$

Therefore, in order to prove Eq. (46), it is sufficient to show that

$$5\sqrt{\ln(ec_2)}\sqrt{\ln(e\eta^{-1}\Delta\epsilon^{-1})} \geq 3\sqrt{\ln(e\sigma/\epsilon_1)} + \sqrt{\ln\left(200c_2^2\sigma\ln(e\sigma/\epsilon_1)/(3\epsilon_1\eta)\right)}. \tag{96}$$

To achieve this, we consider two cases separately:

- Case 1: $\epsilon_1 = \sigma$. In this case, we have

$$\begin{aligned} 3\sqrt{\ln(e\sigma/\epsilon_1)} + \sqrt{\ln\left(200c_2^2\sigma\ln(e\sigma/\epsilon_1)/(3\epsilon_1\eta)\right)} &= 3 + \sqrt{\ln\left(200c_2^2/(3\eta)\right)} && (97)\\ &\leq 5\sqrt{\ln\left(3ec_2^2\eta^{-1}\right)} && (98)\\ &\leq 5\sqrt{\ln(ec_2)}\sqrt{\ln(3e\eta^{-1})} && (99)\\ &\leq 5\sqrt{\ln(ec_2)}\sqrt{\ln(e\eta^{-1}\Delta\epsilon^{-1})}, && (100) \end{aligned}$$

  where the second step follows from the facts that $c_2, \eta^{-1} \geq 1$ and $3 + \sqrt{\ln(200t/3)} \leq 5\sqrt{\ln(3et)}$ for all $t \geq 1$, the third step follows from the facts that $c_2, \eta^{-1} \geq 1$ and $\ln\left(a^2t\right) \leq \ln(ea)\cdot\ln(t)$ for all $a \geq 1$ and $t \geq 3e$, and the last step follow from the assumption $\Delta\epsilon^{-1} > 8$. So Eq. (96) holds.

- Case 2: $\epsilon_1 = \frac{\epsilon}{1.1} < \sigma$. In this case, let $A \coloneqq \eta^{-1}\Delta\epsilon^{-1} \geq \Delta\epsilon^{-1} \geq 8$. Then we have

$$\ln(e\sigma/\epsilon_1) = \ln\left(e0.22\Delta\epsilon^{-1}/(\sqrt{\ln(ec_2)}\sqrt{\ln(eA)})\right) \leq \ln\left(e0.22A/\sqrt{\ln(eA)}\right), \tag{101}$$

  where the second step follows from $\Delta\epsilon^{-1} \leq A$ and $c_2 \geq 1$, and

$$\begin{aligned} \ln\left(200c_2^2\sigma\ln(e\sigma/\epsilon_1)/(3\epsilon_1\eta)\right) &\leq \ln\left(200c_2^2\sigma^2/(3\epsilon_1^2\eta)\right) && (102)\\ &\leq \ln\left(200c_2^2\sigma^2/(3\epsilon_1^2\eta^2)\right) && (103)\\ &\leq \ln\left(67c_2^2\right) + 2\ln\left(\eta^{-1}\sigma\epsilon_1^{-1}\right)) && (104)\\ &= \ln\left(67c_2^2\right) + 2\ln\left(0.22A/(\sqrt{\ln(ec_2)}\sqrt{\ln(eA)})\right) && (105)\\ &\leq \ln\left(67c_2^2\right) + 2\ln\left(0.22A/\sqrt{\ln(eA)}\right) && (106) \end{aligned}$$

  where the first step follows from $\ln(e\sigma/\epsilon_1) \leq \sigma/\epsilon_1$, the second step follows from $\eta^{-1} \geq 1$, the fourth step follows from the definition of σ, ϵ_1 and A, and the last step follows from $c_2 \geq 1$. Now we combine Eqs. (101), (106) and the fact that

$$5\sqrt{\ln(ce)}\sqrt{\ln(et)} \geq 3\sqrt{\ln\left(e0.22t/\sqrt{\ln(et)}\right)} + \sqrt{\ln(67c^2) + 2\ln\left(0.22t/\sqrt{\ln(et)}\right)} \tag{107}$$

  for all $c \geq 1$ and $t \geq 8$, and obtain Eq. (96).

C Polynomial and trigonometric approximations of Gaussian functions

In this appendix, we prove Lemma 4.4 and Lemma 5.4 which state that there exist low-degree polynomial and trigonometric approximations of Gaussian functions, respectively.

Proof of Lemma 4.4. By Theorem 4.1 of Ref. [49], for every $b > 0$ and $\gamma \in (0,1)$, there exists a real polynomial $R_{b,\gamma}(x)$ of degree $\mathcal{O}(\sqrt{\max(b, \log(1/\gamma)) \cdot \log(1/\gamma)})$ such that $|e^{-x} - R_{b,\gamma}(x)| \le \gamma$ for all $x \in [0, b]$. Furthermore, Ref. [49] also gives an explicit construction of $R_{b,\gamma}(x)$.

Now we replace x with $x^2/(2\sigma^2)$ and set $b = \pi^2/(2\sigma^2)$, $\gamma = \epsilon''/3$ in the above result, and obtain that: for every $\sigma > 0$ and $\epsilon'' \in (0,1)$, there exists an efficiently-computable real polynomial $\tilde{P}_{\sigma;\epsilon''}(x)$ of degree $\mathcal{O}(\max(\sigma^{-1}, \sqrt{\log(1/\epsilon'')}) \cdot \sqrt{\log(1/\epsilon'')})$ such that $|e^{-x^2/(2\sigma^2)} - \tilde{P}_{\sigma;\epsilon''}(x)| \le \epsilon''/3$ for all $x \in [-\pi, \pi]$. Note that this implies $|\tilde{P}_{\sigma;\epsilon''}(x)| \le 1 + \epsilon''/3$ for all $x \in [-\pi, \pi]$.

Now let $P_{\sigma;\epsilon''}(x) \coloneqq (1 - \epsilon''/3)\tilde{P}_{\sigma;\epsilon''}(x)$. Then we have that for every $x \in [-\pi, \pi]$,

$$|P_{\sigma;\epsilon''}(x)| = (1 - \epsilon''/3)|\tilde{P}_{\sigma;\epsilon''}(x)| \le (1 - \epsilon''/3)(1 + \epsilon''/3) \le 1, \tag{108}$$

and

$$\left|P(x) - e^{-x^2/(2\sigma^2)}\right| \le \left|P(x) - \tilde{P}(x)\right| + \left|\tilde{P}(x) - e^{-x^2/(2\sigma^2)}\right| \tag{109}$$

$$\le \frac{\epsilon''}{3} \cdot |\tilde{P}(x)| + \frac{\epsilon''}{3} \tag{110}$$

$$\le \frac{\epsilon''}{3}\left(1 + \frac{\epsilon''}{3}\right) + \frac{\epsilon''}{3} \tag{111}$$

$$\le \epsilon''. \tag{112}$$

Thus $P_{\sigma;\epsilon''}(x)$ satisfies all the requirements of Lemma 4.4. □

Proof of Lemma 5.4. We claim that for every $\sigma, \epsilon'' \in (0,1)$, there exist $T = \max(2\pi, \Theta(\sigma\sqrt{\log(1/\epsilon'')}))$ and $N = \mathcal{O}(\sigma^{-1}T\sqrt{\log(1/\epsilon'')})$ and $a_0, a_1, \ldots, a_N \in \mathbb{R}^+$ such that $\sum_{j=0}^N a_j \le 1$ and

$$\left|\sum_{j=0}^{N} a_j \cos(2\pi j x/T) - e^{-x^2/(2\sigma^2)}\right| \le \epsilon'', \quad \forall x \in [-\pi, \pi]. \tag{113}$$

If this claim holds, then since $\cos(2\pi j x/T) = \mathcal{T}_{2j}(\cos(\pi x/T))$, where $\mathcal{T}_{2j}(x)$ is the $2j$-th Chebyshev polynomial of the first kind, $\sum_{j=0}^N a_j \cos(2\pi j x/T)$ can be written as $G_{\sigma;\epsilon''}(\cos(\pi x/T))$ for some even real polynomial $G_{\sigma;\epsilon}(x)$ of degree $2N = \mathcal{O}(\sigma^{-1}T\sqrt{\log(1/\epsilon'')})$. So we get

$$\left|G_{\sigma;\epsilon'}(\cos(\pi x/T)) - e^{-x^2/(2\sigma^2)}\right| \le \epsilon'', \quad \forall x \in [-\pi, \pi]. \tag{114}$$

Furthermore, since $a_0, a_1, \ldots, a_N \ge 0$, we have

$$G_{\sigma;\epsilon''}(\cos(\pi x/T)) = \sum_{j=0}^{N} a_j \cos(2\pi j x/T) \le \sum_{j=0}^{N} a_j \le 1, \quad \forall x \in \mathbb{R}. \tag{115}$$

Therefore $G_{\sigma;\epsilon''}(x)$ satisfies all the requirements of Lemma 5.4.

The proof of the above claim is inspired by the Nyquist–Shannon sampling theorem. Specifically, let $f(x) \coloneqq \sum_{k=-\infty}^{\infty} \delta(x - kT)$ be the Dirac comb with period T. It has Fourier transform $\hat{f}(\xi) = \frac{1}{T}\sum_{n=-\infty}^{\infty} \delta(\xi - n/T)$. Meanwhile, let $g(x) \coloneqq e^{-x^2/(2\sigma^2)}$ be Gaussian with mean 0 and variance σ^2. It has Fourier transform $\hat{g}(\xi) = \sqrt{2\pi}\sigma e^{-2\pi^2\sigma^2\xi^2}$. Convolving f and g yields

$$(f * g)(x) = \sum_{k=-\infty}^{\infty} g(x - kT). \tag{116}$$

On the other hand, the Fourier transform of $f * g$ is

$$\widehat{f * g}(\xi) = \hat{f}(\xi)\hat{g}(\xi) = \frac{1}{T}\sum_{n=-\infty}^{\infty} \hat{g}(n/T)\delta(\xi - n/T). \tag{117}$$

Applying inverse Fourier transform to both sides of this equation leads to

$$(f * g)(x) = \frac{1}{T}\sum_{n=-\infty}^{\infty} \hat{g}(n/T)e^{\mathbf{i}2\pi xn/T}. \tag{118}$$

Comparison of Eqs. (116) and (118) indicates that

$$\sum_{k=-\infty}^{\infty} g(x - kT) = \frac{1}{T}\sum_{n=-\infty}^{\infty} \hat{g}(n/T)e^{\mathbf{i}2\pi xn/T}. \tag{119}$$

Now we claim that for some $T = \max(2\pi, \Theta(\sigma\sqrt{\log(1/\epsilon'')}))$ and $N = \mathcal{O}(\sigma^{-1}T\sqrt{\log(1/\epsilon'')})$,

$$\sum_{k=-\infty}^{-1} g(x - kT) + \sum_{k=1}^{\infty} g(x - kT) \leq \frac{\epsilon''}{6}, \quad \forall x \in [-\pi, \pi], \tag{120}$$

and

$$\left|\frac{1}{T}\sum_{n=-\infty}^{-N-1} \hat{g}(n/T)e^{\mathbf{i}2\pi xn/T} + \frac{1}{T}\sum_{n=N+1}^{\infty} \hat{g}(n/T)e^{\mathbf{i}2\pi xn/T}\right| \leq \frac{\epsilon''}{6}, \quad \forall x \in \mathbb{R}. \tag{121}$$

If these claims are true, then combining them and Eq. (119) yields

$$\left|g(x) - \frac{1}{T}\sum_{n=-N}^{N} \hat{g}(n/T)e^{\mathbf{i}2\pi xn/T}\right| = \left|g(x) - \hat{g}(0) - \frac{2}{T}\sum_{n=1}^{N} \hat{g}(n/T)\cos(2\pi xn/T)\right| \leq \frac{\epsilon''}{3}, \quad \forall x \in [-\pi, \pi]. \tag{122}$$

This also implies that

$$\hat{g}(0) + \frac{2}{T}\sum_{n=1}^{N} \hat{g}(n/T) \leq 1 + \frac{\epsilon''}{3}. \tag{123}$$

Now let $a_0 = (1 - \epsilon''/3)\hat{g}(0)$ and $a_j = 2T^{-1}(1 - \epsilon''/3)\hat{g}(j/T)$ for $j = 1, 2, \ldots, N$. Then we have $a_0, a_1, \ldots, a_N > 0$ and

$$\sum_{j=0}^{N} a_j = (1 - \epsilon''/3)\left(\hat{g}(0) + \frac{2}{T}\sum_{n=1}^{N} \hat{g}(n/T)\right) \leq (1 - \epsilon''/3)(1 + \epsilon''/3) < 1, \tag{124}$$

and for any $x \in [-\pi, \pi]$,

$$\left|\sum_{j=0}^{N} a_j \cos(2\pi jx/T) - g(x)\right| \leq \left|\sum_{j=0}^{N} a_j \cos(2\pi jx/T) - \hat{g}(0) - \frac{2}{T}\sum_{n=1}^{N} \hat{g}(n/T)\cos(2\pi xn/T)\right| + \left|\hat{g}(0) + \frac{2}{T}\sum_{n=1}^{N} \hat{g}(n/T)\cos(2\pi xn/T) - g(x)\right| \tag{125}$$

$$\leq \frac{\epsilon''}{3} \cdot \left|\hat{g}(0) + \frac{2}{T}\sum_{n=1}^{N} \hat{g}(n/T)\cos(2\pi xn/T)\right| + \frac{\epsilon''}{3} \tag{126}$$

$$\leq \frac{\epsilon''}{3}\left(\hat{g}(0) + \frac{2}{T}\sum_{n=1}^{N} \hat{g}(n/T)\right) + \frac{\epsilon''}{3} \tag{127}$$

$$\leq \frac{\epsilon''}{3}\left(1 + \frac{\epsilon''}{3}\right) + \frac{\epsilon''}{3} \tag{128}$$

$$\leq \epsilon'', \tag{129}$$

as desired.

It remains to prove Eqs. (120) and (121). To prove the former, we impose the constraint $T \geq 2\pi$. This means that $|x| \leq T/2$ for all $x \in [-\pi, \pi]$. Then it follows that

$$\sum_{k=-\infty}^{-1} g(x - kT) + \sum_{k=1}^{\infty} g(x - kT) \leq 2\sum_{l=0}^{\infty} g((l + 1/2)T) \tag{130}$$

$$= 2\sum_{l=0}^{\infty} \exp\left(-(l + 1/2)^2 T^2/(2\sigma^2)\right) \tag{131}$$

$$= 2\exp\left(-T^2/(8\sigma^2)\right)\sum_{l=0}^{\infty} \exp\left(-(l^2 + l)T^2/(2\sigma^2)\right) \tag{132}$$

$$\leq 2\exp\left(-T^2/(8\sigma^2)\right)\sum_{l=0}^{\infty} \exp\left(-lT^2/\sigma^2\right) \tag{133}$$

$$= \frac{2\exp\left(-T^2/(8\sigma^2)\right)}{1 - \exp\left(-T^2/\sigma^2\right)}. \tag{134}$$

By picking some $T = \Theta(\sigma\sqrt{\log(1/\epsilon'')})$, we can ensure that the RHS of Eq. (134) is at most $\epsilon''/6$. Overall, Eqs. (120) holds as long as we pick some $T = \max(2\pi, \Theta(\sigma\sqrt{\log(1/\epsilon'')}))$.

To prove Eq. (121), we note that

$$\left|\frac{1}{T}\sum_{n=-\infty}^{-N-1}\hat{g}(n/T)e^{\mathbf{i}2\pi xn/T}+\frac{1}{T}\sum_{n=N+1}^{\infty}\hat{g}(n/T)e^{\mathbf{i}2\pi xn/T}\right| \leq \frac{1}{T}\sum_{n=-\infty}^{-N-1}\hat{g}(n/T)+\frac{1}{T}\sum_{n=N+1}^{\infty}\hat{g}(n/T) \quad (135)$$

$$= \frac{2}{T}\sum_{n=N+1}^{\infty}\hat{g}(n/T) \quad (136)$$

$$= \frac{2\sqrt{2\pi}\sigma}{T}\sum_{n=N+1}^{\infty}\exp\left(-2\pi^2\sigma^2n^2/T^2\right) \quad (137)$$

$$\leq \frac{2\sqrt{2\pi}\sigma}{T}\int_{N}^{\infty}\exp\left(-2\pi^2\sigma^2y^2/T^2\right)dy \quad (138)$$

$$= \frac{2}{\sqrt{\pi}}\int_{\sqrt{2}\pi\sigma N/T}^{\infty}\exp\left(-z^2\right)dz \quad (139)$$

$$= \operatorname{erfc}(\sqrt{2}\pi\sigma N/T) \quad (140)$$

$$\leq \exp\left(-2\pi^2\sigma^2N^2/T^2\right). \quad (141)$$

By picking some $N = \mathcal{O}(\sigma^{-1}T\sqrt{\log(1/\epsilon'')})$, we can guarantee that $\exp\left(-2\pi^2\sigma^2N^2/T^2\right) \leq \epsilon''/6$ and hence Eq. (121) holds for all $x \in \mathbb{R}$.

The lemma is thus proved.

□

D Certification of the ground state energy estimate

Our algorithm has the desirable feature that it only requires us to consider circuit depths that are proportional to the spectral gap of the Hamiltonian. However, in many applications of practical interest we may not know the spectral gap in advance and would like to check if the output of the algorithm was correct.

Furthermore, as we will explain soon, depending on the initial state, it might be the case that the algorithm still outputs the right estimate even if we evolve for times that are smaller than the spectral gap. If we suspect we have such a structured initial state, it is desirable to be able to verify that the output was correct.

We will now outline a test such that, given $\sigma > 0$, an estimate $\hat{E}_0$ of the ground state energy E_0 up to σ, a precision tolerance ϵ, a failure probability $\delta > 0$ and a lower bound on the overlap $p_0 \geq \eta$, accepts only if the estimate satisfies $|\hat{E}_0 - E_0| \leq \epsilon$, with a failure probability of wrong certification at most δ. The quantum circuits required to run the algorithm are of depth $\mathcal{O}(\sigma^{-1})$. Furthermore, we will also show that if σ is proportional to the spectral gap Δ_{true}, it is guaranteed to accept with probability at least $1-\delta$ (although it might also accept at a larger value of σ/smaller circuit depth for more structured states).

Note that we will *not* certify that the spectral gap of the Hamiltonian is at least some Δ, but only that the ground state energy estimate is correct up to ϵ under the promise that the initial state had an overlap with the ground space of at least $\eta > 0$. Certifying the spectral gap would require additional assumptions on the initial state and its support on the first excited state. To make this point clear, consider the case in which the initial state is $\rho = |E_0\rangle\langle E_0|$, i.e. the ground state. In this case, it is not difficult to see that our algorithm already outputs the correct estimate with access to controlled evolution times of order $\mathcal{O}(1)$, as E_0 will be the maximum of $p * n_\sigma$ for all $\sigma > 0$, where n_σ is again a Gaussian with standard deviation σ. Thus, given such an initial state

and enough samples, our algorithm will always output the correct estimate regardless of the circuit depth. On the other hand, it is clear that the state ρ cannot give us any information about the spectral gap of the Hamiltonian. In short, our test will only certify that the estimate of the energy was within a tolerated error.

Intuitively speaking, our algorithm outputs a correct estimate whenever σ is small enough to make sure that the contribution of the other energies in n_σ at E_0 is sufficiently suppressed to make sure that we have a peak close to the ground state. If we have pollution from other energies, then they will make the peak around E_0 wider than if not. Thus, our test will consist on first making sure that we have a peak close to E_0 and then measuring the width of the peak.

We will first collect an auxilliary result about mixtures of Gaussians, as $p * n_\sigma$ is a mixture of Gaussians, i.e. a convex combination of Gaussian random variables. The Lemma below formalizes the idea that if we have pollution from other energies, then the peak around E_0 will be wider than it should be, where we measure the width by the variance.

Lemma D.1. *Let $X = \sum_i p_i N_i$, where $N_i \sim \mathcal{N}(E_i, \sigma^2)$ are normal random variables with the same variance σ^2 but different means E_i. Furthermore, assume $E_0 < E_1 < \cdots < E_m$ and that $p_0 \geq \eta$. If $|\mathbb{E}(X) - E_0| \geq c\epsilon$ for some constants $c > 1, \epsilon > 0$, then:*

$$\mathbb{E}(|X - \mathbb{E}(X)|^2) \geq \sigma^2 + \eta c^2 \epsilon^2 \tag{142}$$

Proof. Recall the standard identity $\mathbb{E}(|X - \mathbb{E}(X)|^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$. Let us compute these expectation values. Clearly, $\mathbb{E}(X) = \sum_i p_i E_i$. Furthermore,

$$\mathbb{E}(X^2) = \sum_i p_i \mathbb{E}(N_i^2) = \sum_i p_i(\sigma^2 + E_i^2). \tag{143}$$

Thus,

$$\mathbb{E}(X^2) - \mathbb{E}(X)^2 = \sigma^2 + \sum_i p_i E_i^2 - (\sum_i p_i E_i)^2.$$

Now note that $\sum_i p_i E_i^2 - (\sum_i p_i E_i)^2$ is nothing but the variance of a discrete random variable Y that takes value E_i with probability p_i and $\mathbb{E}(Y) = \mathbb{E}(X)$. Then

$$\mathbb{E}(|Y - \mathbb{E}(Y)|^2) \geq \eta c^2 \epsilon^2. \tag{144}$$

To see this, note that by our assumption that $|\mathbb{E}(X) - E_0| \geq c\epsilon$, with probability p_0 the difference $|Y - \mathbb{E}(Y)|^2$ will be at least $c^2\epsilon^2$, which yields the claim. □

The result above shows that if the mean of a Gaussian mixture deviates significantly from the minimum mean of its components, then this will result in a variance that is larger than their individual variances σ^2.

Our test will be based on "zooming into" the peak around the estimate we found with our sampling algorithm and testing whether the variance is larger than expected. That is, we will condition the distribution around $\hat{E}_0$ and estimate its variance.

But it is not a given that conditioning around $\hat{E}_0$ will also yield a distribution that is a mixture of Gaussians. Thus, we will show that this is approximately the case if we have a well-defined peak in the distribution.

Before we prove our theorem showing how conditioning on an interval yields another mixture of Gaussians, let us again give some intuition. Given some peak $\hat{E}_0$ that we identified, we will truncate to some interval $[\hat{E}_0 - L, \hat{E}_0 + L]$, with

$$L = \sqrt{2c \log(\sigma^{-1}\tau^{-1})}\sigma \tag{145}$$

for some constants $c > 1, 1 > \tau > 0$ to be specified later. The mass of energies that are located to the right of $E_0 + 3L$ is negligible on the interval $[\hat{E}_0 - L, \hat{E}_0 + L]$, as it is at most $\mathcal{O}(L\tau^{2c})$. We see that we can safely discard these random variables when truncating to the interval. Furthermore, we will also approximate well all Gaussians whose means are located in $[\hat{E}_0 - L/2, \hat{E}_0 + L/2]$, as their mass outside this interval is $\mathcal{O}(\tau^{\frac{c}{4}})$.

Thus, the only energies that "pose a threat" of pollution are those that have substantial weight on the edge of our interval $[\hat{E}_0 - L, \hat{E}_0 + L]$. This is where we use the assumption that we found a peak: we will impose that the mass going around the edges of the intervals is small. Note that only the right edge of the interval has to be considered, as we assume in our algorithm that we know the ground state energy up to σ and, thus, there are no energies to the left of the interval. More formally, we want:

$$\int_{\hat{E}_0 + \frac{L}{2}}^{\hat{E}_0 + 2L} (p * n_\sigma)(x) dx \leq c' \epsilon^2 \eta^2. \tag{146}$$

for some constant c' we will choose later. We will call distributions that satisfy this property *peaked around* $\hat{E}_0$.

Definition D.2 (Distribution peaked around value)**.** *Given a point distribution p on $\mathbb{R}$, $\epsilon > 0$, $\eta > 0$ and n_σ the probability density function of a Gaussian random variable with standard deviation σ, we say that $p * n_\sigma$ is peaked around a point $\hat{E}_0$ up to ϵ with parameters $c, c', \tau > 0$ if for $L = \sigma\sqrt{2c \log(\sigma^{-1}\tau^{-1})}$ we have that:*

$$\int_{\hat{E}_0 + \frac{L}{2}}^{\hat{E}_0 + 2L} (p * n_\sigma)(x) dx \leq c' \epsilon^2 \eta^2. \tag{147}$$

To further simplify notation we will define $n_{\sigma,i}$ to be the density of the random variable $\mathcal{N}(E_i, \sigma^2)$. We will start by proving that whenever Eq. (146) holds, the mass on $[\hat{E}_0 - L, \hat{E}_0 + L]$ is mostly given by energies in the interval $[\hat{E}_0 - L/2, \hat{E}_0 + L/2]$.

Lemma D.3. *Let $X = \sum_i p_i N_i$, where $N_i \sim \mathcal{N}(E_i, \sigma^2)$ are normal random variables with the same variance σ^2 but different means E_i and L be defined as in Eq.* (145)*. For some $\hat{E}_0$ that is whithin σ of E_0, group the means of the N_i into three subsets:*

$$G = \{i : |E_i - \hat{E}_0| \leq \tfrac{L}{2}\}, \tag{148}$$

$$B = \{i : \tfrac{L}{2} \leq |E_i - \hat{E}_0| \leq 2L\}, \tag{149}$$

$$F = \{i : 2L \leq |E_i - \hat{E}_0|\}. \tag{150}$$

If $p * n_\sigma$ is peaked around $\hat{E}_0$ with $c = 4$ and $\tau = \frac{c' \epsilon^2 \eta^2}{4}$ then:

$$\left| \int_{\hat{E}_0 - L}^{\hat{E}_0 + L} (p * n_\sigma)(x) dx - \sum_{i \in G} p_i \right| \leq 4c' \epsilon^2 \eta^2. \tag{151}$$

Proof. Clearly:

$$\int_{\hat{E}_0 - L}^{\hat{E}_0 + L} (p * n_\sigma)(x) dx = \sum_{i \in G} p_i \int_{\hat{E}_0 - L}^{\hat{E}_0 + L} n_{\sigma,i}(x) dx + \sum_{i \in B} p_i \int_{\hat{E}_0 - L}^{\hat{E}_0 + L} n_{\sigma,i}(x) dx + \sum_{i \in F} p_i \int_{\hat{E}_0 - L}^{\hat{E}_0 + L} n_{\sigma,i}(x) dx. \tag{152}$$

We will estimate each term in the sum separately. Let us start with energies in F. As, by definition, all energies in that set are at least $2L$ away from $\tilde{E}_0$, they are at least L away from the upper end of the integration limits. Thus, by our choice of L we have for $i \in F$:

$$\int_{\hat{E}_0-L}^{\hat{E}_0+L} n_{\sigma,i}(x)dx \leq 2L\frac{e^{c\log(\tau\sigma)}}{\sigma\sqrt{2\pi}} = \mathcal{O}(\tau^c \log\left(\tau^{-1}\right)) \tag{153}$$

and so

$$\sum_{i\in F} p_i \int_{\hat{E}_0-L}^{\hat{E}_0+L} n_{\sigma,i}(x)dx = \mathcal{O}(\tau^c \log\left(\tau^{-1}\right)). \tag{154}$$

Let us now estimate the contribution of terms in G. We have for $i \in G$:

$$\int_{\hat{E}_0-L}^{\hat{E}_0+L} n_{\sigma,i}(x)dx = 1 - \int_{-\infty}^{\hat{E}_0-L} n_{\sigma,i}(x)dx - \int_{\hat{E}_0+L}^{+\infty} n_{\sigma,i}(x)dx. \tag{155}$$

By the property that all energies in G are $\frac{L}{2}$ away from $\hat{E}_0$, we see that both integrals above are at most $\mathcal{O}(\tau^{\frac{c}{2}})$. Thus:

$$\left|\sum_{i\in G} p_i \int_{\hat{E}_0-L}^{\hat{E}_0+L} n_{\sigma,i}(x)dx - \sum_{i\in G} p_i\right| = \mathcal{O}(\tau^{\frac{c}{2}}). \tag{156}$$

By our choice of L we have $\tau^{\frac{c}{2}} = \left(\frac{c'\epsilon^2\eta^2}{4}\right)^2$. Thus, both Eq. (156) and Eq. (154) are bounded by $c'\epsilon^2\eta^2$.

To finish the proof, we need to consider the contribution of the energies in B. Note that for $i \in B$:

$$\sum_{i\in B} p_i \int_{\hat{E}_0-L}^{\hat{E}_0+L} n_{\sigma,i}(x)dx \leq 2\sum_{i\in B} p_i \int_{\hat{E}_0+\frac{L}{2}}^{\hat{E}_0+2L} n_{\sigma,i}(x)dx \leq \sum_i 2p_i \int_{\hat{E}_0+\frac{L}{2}}^{\hat{E}_0+2L} n_{\sigma,i}(x)dx \leq 2c'\epsilon^2\eta^2. \tag{157}$$

To see this, note that by definition of B the mean of these Gaussians are greater than $\hat{E}_0 + \frac{L}{2}$ and the integral on the r.h.s. is over a larger interval that is closer to their mean. Thus, as the mass of the Gaussian is maximal around the mean, the inequality follows. By our assumption in Eq. (146) we conclude that:

$$\sum_{i\in B} p_i \int_{\hat{E}_0-L}^{\hat{E}_0+L} n_{\sigma,i}(x)dx \leq 2c'\epsilon^2\eta^2. \tag{158}$$

Putting together inequalities in Eqs. (154), (156) and (158), Eq. (151) follows.

□

We can show in a similar fashion that the following holds:

Corollary D.4. *Under the same conditions as Lemma D.3, assume further that σ is small enough to ensure that $|\hat{E}_0 \pm L| \leq 1$. If $p * n_\sigma$ is peaked around $\hat{E}_0$ with $c = 4$ and $\tau = \frac{c'\epsilon^2\eta^2}{4}$ then:*

$$\left|\int_{\hat{E}_0-L}^{\hat{E}_0+L} x(p * n_\sigma)(x)dx - \sum_{i\in G} p_i E_i\right| \leq 4c'\epsilon^2\eta^2. \tag{159}$$

and

$$\left| \int_{\hat{E}_0-L}^{\hat{E}_0+L} x^2 (p * n_\sigma)(x) dx - \sum_{i \in G} p_i (\sigma^2 + E_i^2) \right| \leq 4c' \epsilon^2 \eta^2. \tag{160}$$

Proof. We can apply the same estimates as we used to show Lemma D.3 to show that the contributions from energies in B and F are small. This is because we assume that $[\hat{E}_0 - L, \hat{E}_0 + L] \subset [-1, 1]$ and for $x \in [-1, 1]$:

$$|x n_{\sigma,i}| \leq n_{\sigma,i}, \quad x^2 n_{\sigma,i} \leq n_{\sigma,i} \tag{161}$$

Furthermore, it is also not difficult to show that for $i \in G$ the value of the first two moments are approximated if we integrate over $[\hat{E}_0 - L, \hat{E}_0 + L]$. □

This shows that under Eq. (146) the moments of the random variable conditioned on the interval $[\hat{E}_0 - L, \hat{E}_0 + L]$ are indeed well-approximated by those of the energies in the interval $[\hat{E}_0 - L/2, \hat{E}_0 + L/2]$. Thus, if we check that Eq. (146) holds and we estimate the moments conditioned on the interval, then we can approximate the quantity in Eq. (142) and check that the variance is larger than expected if we got the energy right up to $c\epsilon$.

D.1 Certification algorithm and its correctness

With these technical results at hand, let us give a description of our energy certification in Algorithm 3, which we refer to as $GSEE\_CERT$. We deliberately formulated it in a way independent

Algorithm 3 Energy certification algorithm algorithm $GSEE\_CERT$.

procedure GSEE_CERT(ϵ, $\hat{E}_0$, σ, η,p)
  $\tau \leftarrow \frac{\epsilon^2 \eta}{160}$
  Let $X \sim \sum_i p_i N(E_i, \sigma^2)$ with pdf $p * n_\sigma$
  $L = \sqrt{8 \log(\tau^{-1})} \sigma$
  $\tau \leftarrow$ estimate of $\int_{\hat{E}_0 + \frac{L}{2}}^{\hat{E}_0 + 2L} p * n_\sigma(x) dx$ up to $\frac{\epsilon^2 \eta}{160}$.
  if $\tau < \frac{\epsilon^2 \eta}{80}$ **then**
    $M \leftarrow$ estimate $\mathbb{E}(X | X \in [\hat{E}_0 - L, \hat{E}_0 - L])$ up to $\frac{\epsilon^2 \eta}{160}$
    $S \leftarrow$ estimate $\mathbb{E}((X - M)^2 | X \in [\hat{E}_0 - L, \hat{E}_0 - L])$ up to $\frac{\epsilon^2 \eta}{160}$.
    if $|S - \sigma^2| \leq 2\epsilon^2 \eta$ **then**
      return Accept estimate and return M.
    else
      return Reject estimate.
    end if
  else
    return Reject estimate.
  end if
end procedure

of how we obtained the estimates for the relevant quantities. This is because different algorithms for GSEE have different access models to the function $p * n_\sigma$, be it samples, as in this work, or to estimates of the density, as in Ref. [40]. Furthermore, we assume that we are certain about the

correctness of the estimates used in the algorithm for simplicity. But it is straightforward to extend the analysis to the case were the estimates are only correct up to some given failure probability. We then have:

Theorem D.5. *If $\hat{E}_0$ is such that $|\hat{E}_0 - E_0| \leq \sigma$, Algorithm 3 accepts, and all the estimates used were correct, then we have:*

$$|M - E_0| \leq 4\epsilon. \tag{162}$$

Proof. If the algorithm accepted, then we already know that the density $p * n_\sigma$ is peaked around $\hat{E}_0$. Otherwise we would have already rejected in the first at Line 15. Let us assume that

$$|\hat{E}_0 - E_0| > 4\epsilon. \tag{163}$$

We will now show that the test will reject. As before, let $X = \sum_i p_i N(E_,\sigma^2)$ and $\tilde{X}$ be the same random variable conditioned on the interval $[\hat{E}_0 - L, \hat{E}_0 - L]$. Furthermore, also let as before G be the set of energies $E_i \in [\hat{E}_0 - L/2, \hat{E}_0 - L/2]$. By Lemma D.3 and Cor. D.4, we have that the variance of $\tilde{X}$ approximates up to $\frac{\epsilon^2\eta}{160}$ that of the random variable $Y = \left(\sum_{i\in G} p_i N(E_i, \sigma^2)\right) / \left(\sum_{i\in G} p_i\right)$.

By Lemma D.1, if Eq. (163) holds, then the variance of X conditioned on $[\hat{E}_0 - L, \hat{E}_0 + L]$ must exceed

$$\mathbb{E}(|X - \hat{E}_0|^2 | X \in [\hat{E}_0 - L, \hat{E}_0 - L]) \geq \sigma^2 + 16\eta\epsilon^2 - \frac{\eta\epsilon^2}{160} > \sigma^2 + 2\epsilon^2\eta, \tag{164}$$

which means that the test will have rejected. Thus, if we accept, then Eq. (162) must hold. □

Let us now show that the test will always accept if we evolve for a time proportional to the gap over the precision:

Lemma D.6. *Let Δ_{true} be the spectral gap of H, $p_0 \geq \eta$, a desired precision $\epsilon > 0$ and $|\hat{E}_0 - E_0| \leq \sigma$. Let σ_0 be the minimal $\sigma > 0$ such that Algorithm 3 accepts for a given p. Then:*

$$\sigma_0 \leq \frac{\Delta_{\text{true}}}{10\sqrt{\log\left(2\eta^{-1}\epsilon^{-1}\right)}}. \tag{165}$$

Proof. Let us first show that for our value of σ_0 in Eq. (165), the distribution will be peaked around E_0. By our choice of σ_0, it is easy to see that the contribution of the other energies on $[E_0 - 2L, E_0 + 2L]$ will be of order $\mathcal{O}((\epsilon\eta)^{50})$, which immediately implies that the distribution is peaked. Furthermore, the probability distribution on $[E_0 - 2L, E_0 + 2L]$ is well-approximated by a single Gaussian up to corrections of order $\mathcal{O}((\epsilon\eta)^{50})$. Thus, the variance conditioned on $[E_0 - L, E_0 + L]$ will deviate from σ by less than $\epsilon^2\eta$. As our estimate S of the variance deviates from the true one by at most $\frac{\epsilon^2\eta}{160}$, we have $|S - \sigma^2| \leq 2\epsilon^2\eta$, which means we will accept. □

We conclude that the test is sound: in the worst case, if we pick σ of similar magnitude as the spectral gap, we are guaranteed to accept. But it can be the case that $\sigma_0 \ll \Delta_{\text{true}}$ depending on the initial state. To illustrate this, take once again the case where $p_0 = 1$. Then the test will always accept, irrespective of the value of the gap. This is because we will always only have one Gaussian component in n_σ and, thus, the variance will always be σ^2.

D.2 Sample complexity of $GSEE\_CERT$ with rejection sampling

Let us analyze the sample complexity of the test above with access to samples and with a success probability of correct verification of $1-\delta$.

Corollary D.7. *Under the same conditions as Theorem 3.6, assume further that $|E_0 - \hat{E}_0| \leq \sigma$. Then we can perform the run Algorithm 3 with probability of correct rejection or acceptance $1-\delta$ with access to*

$$\tilde{\mathcal{O}}(\epsilon^{-4}\eta^{-3}\log\left(\delta^{-1}\right)) \tag{166}$$

uses of the block-encoding of $g_{\sigma,\xi;\epsilon''}(H)$.

Proof. There are two steps that need to be performed for running Algorithm 3. First, we need to determine if the distribution is peaked around $\hat{E}_0$. If it is peaked, we need to determine the variance up to an error $\eta\epsilon^2$ on the interval $[\hat{E}_0 - L, \hat{E}_0 + L]$. Let us start by showing how to determine if the distribution is peaked. For that we will perform rejection sampling with the Gaussian condtioned on $[\hat{E}_0 + \frac{L}{2}, \hat{E}_0 + 2L]$. As explained in Section 3 in Eq. (15), the probability of accepting is given by

$$\frac{\sigma\sqrt{2\pi}}{3L\int_{\hat{E}_0+\frac{L}{2}}^{\hat{E}_0+2L}(p*n_\sigma)(x)dx} = \frac{c\log\left(\tau^{-1}\right)}{\int_{\hat{E}_0+\frac{L}{2}}^{\hat{E}_0+2L}(p*n_\sigma)(x)dx}, \tag{167}$$

for an explicitly known constant c. Thus, by running $\mathcal{O}(\epsilon^{-2}\eta^{-1})$ rounds of the rejection sampling, we can decide if the distribution is peaked around $\hat{E}_0$ or not with probability of failure at most $\frac{\delta}{2}$ by checking how many times we accepted the sample.

If it is peaked, then we can perform rejection sampling again, but now on the interval $[\hat{E}_0 - L, \hat{E}_0 + L]$. With access to $\mathcal{O}(\eta^{-2}\epsilon^{-4}\log\left(\delta^{-1}\right))$ samples from the conditioned distribution, we can estimate the variance of the conditioned random variable up to a precision of $\mathcal{O}(\epsilon^2\eta)$ up to a failure probability of at most $\frac{\delta}{2}$.

The probability of accepting a sample from the interval is given by

$$\frac{\sigma\sqrt{2\pi}}{2L\int_{\hat{E}_0-L}^{\hat{E}_0+L}(p*n_\sigma)(x)dx}. \tag{168}$$

As before, note that for our choice of L we have that this is lower-bounded by

$$\frac{C\log\left(\tau^{-1}\right)}{\int_{\hat{E}_0-L}^{\hat{E}_0+L}(p*n_\sigma)(x)dx} \tag{169}$$

for some constant $C > 0$. Now, as $E_0 \in [\hat{E}_0 - L, \hat{E}_0 + L]$ by our choice of L and the assumption that $|E_0 - \hat{E}_0| \leq \sigma$, we have that $\int_{\hat{E}_0-L}^{\hat{E}_0+L}(p*n_\sigma)(x)dx = \Omega(\eta)$. Thus, we need to run the rejection sampling an expected $\tilde{\mathcal{O}}(\eta^{-1})$ number of times to accept a sample. This gives a total complexity for this step of $\tilde{\mathcal{O}}(\eta^{-3}\epsilon^{-4}\log\left(\delta^{-1}\right))$, which dominates the overall sample complexity of the certification algorithm. Furthermore, we see that the probability of obtaining wrong estimates is at most δ. As the algorithm always outputs the correct answer if given correct estimates, this gives the claim. □

Note that the assumption that we are given an estimate that satisfies $|E_0 - \hat{E}_0| \leq \sigma$ does not increase the overall depth of circuits required for the certification method. This is because we can always obtain such an estimate with circuits of depth $\mathcal{O}(\sigma^{-1})$ using e.g. Algorithm 1.