Hamiltonian simulation for low-energy states with optimal time dependence

Alexander Zlokapa[1,2] and Rolando D. Somma[2]

[1]Center for Theoretical Physics, MIT, 02139, USA
[2]Google Quantum AI, Venice, CA 90291, USA

We consider the task of simulating time evolution under a Hamiltonian H within its low-energy subspace. Assuming access to a block-encoding of $H' := (H - E)/\lambda$, for some $\lambda > 0$ and $E \in \mathbb{R}$, the goal is to implement an ϵ-approximation to the evolution operator e^{-itH} when the initial state is confined to the subspace corresponding to eigenvalues $[-1, -1 + \Delta/\lambda]$ of H', for $\Delta \leq \lambda$. We present a quantum algorithm that requires $\mathcal{O}(t\sqrt{\lambda\Gamma} + \sqrt{\lambda/\Gamma}\log(1/\epsilon))$ queries to the block-encoding for any choice of Γ such that $\Delta \leq \Gamma \leq \lambda$. When the parameters satisfy $\log(1/\epsilon) = o(t\lambda)$ and $\Delta/\lambda = o(1)$, this result improves over generic methods with query complexity $\Omega(t\lambda)$. Our quantum algorithm leverages spectral gap amplification and the quantum singular value transform.

For a given H, the block-encoding of its H' must be prepared efficiently to achieve an asymptotic speedup in simulating the low-energy subspace; we refer to these Hamiltonians as *gap-amplifiable*. We show necessary and sufficient conditions for gap amplifiability in terms of an operationally useful decomposition of H into a sum of squares. Gap-amplifiable Hamiltonians include physically relevant examples such as frustration-free systems, and it encompasses all previously considered settings of low-energy simulation algorithms. Any Hamiltonian can be expressed as a gap-amplifiable Hamiltonian after simple transformations, and our algorithm retains the asymptotic improvement over generic methods as long as the conditions on the parameters are met.

We also provide lower bounds for simulating dynamics of low-energy states. In the worst case, we show that the low-energy condition cannot be used to improve the runtime of Hamiltonian simulation methods. For gap-amplifiable Hamiltonians, we prove that our algorithm is tight in the query model with respect to t, Δ, and λ. In the practically relevant regime where $\log(1/\epsilon) = o(t\Delta)$ and $\Delta/\lambda = o(1)$, we also prove a matching lower bound in gate complexity (up to logarithmic factors). To establish the query lower bounds, we consider oracular problems including search and PARITY $\circ$ OR, and also bounds on the degrees of trigonometric polynomials. To establish the lower bound on gate complexity, we use a circuit-to-Hamiltonian reduction, where a "clock Hamiltonian" acting on a low-energy state can simulate any quantum circuit.

1 Introduction

A leading application for quantum computers is the simulation of the dynamics of quantum systems [23]: given a Hamiltonian H, an initial quantum state $|\psi\rangle$, and an evolution time t, the task is to produce the state at a later time as predicted by Schrödinger's equation,

i.e., $e^{-itH}|\psi\rangle$. Simulating time evolution on a quantum computer, also known as the Hamiltonian simulation problem, is relevant to numerous problems in physics [34, 46], chemistry [3, 34], and many other quantum algorithms (cf., [2, 18, 29]). Extensive effort has thus been made towards obtaining the most efficient quantum algorithms for Hamiltonian simulation in a variety of settings [8–10, 15, 28, 36, 38, 52]. While such quantum algorithms are efficient in that they require polynomial resources in the evolution time and system size for interesting classes of Hamiltonians, many applications introduce considerations that necessitate careful comparison between algorithms. For this reason, recent studies examine improvements for Hamiltonian simulation when given additional knowledge about structures of the Hamiltonian or properties of the initial states. Examples include the locality of interactions [19, 28], Lie-algebraic properties [47], symmetries [51], or particle statistics [5, 6].

Along similar lines, in this article we focus on the Hamiltonian simulation problem for an important class of initial quantum states: *low-energy* states. Given a Hamiltonian H, low-energy states are those supported exclusively on a subspace of energies that are strictly below a given energy (eigenvalue) of H. Hamiltonian simulation of low-energy states is relevant in physically motivated contexts, such as when studying zero-temperature quantum phase transitions in condensed matter systems [43], simulating quantum field theories [32], or computing ground-state energies of molecules with quantum phase estimation [3, 33]. It is also essential for adiabatic quantum computing [1, 22], a generic technique to solve optimization and other problems in quantum computing by preparing the ground state of a complex (interacting) Hamiltonian starting from the ground state of a simpler (non-interacting) one. By ground state we mean the eigenstate (eigenvector) of lowest energy (eigenvalue). Consequently, any enhancement of Hamiltonian simulation techniques under the low-energy consideration of the initial state is anticipated to uncover substantial applications.

Owing to its significance, some recent research has already investigated Hamiltonian simulation of low-energy states. In particular, product formulas may exploit certain structures of the problem (e.g., locality of interactions), especially if the precision requirements are not too stringent. Reference [44] discusses improvements on Hamiltonian simulation methods based on Trotter-Suzuki product formulas when H is a local Hamiltonian presented as a sum of positive semidefinite (PSD) terms. Comparable results were provided in Ref. [26] for additional product-formula based quantum algorithms, such as qDRIFT [15] or other randomized product formulas. While these quantum algorithms show an improvement with respect to the general case that allows an arbitrary initial state, they are generally not optimal in terms of the evolution time, Hamiltonian norm, or the allowed error, which is a common feature of product formulas. Nevertheless, they might still offer some advantages with respect to qubit overheads in comparison to other methods.

Unlike product formulas, methods based on a block-encoding of the Hamiltonian can achieve optimal worst-case bounds (cf. [10, 37]). In principle, these methods do not offer much flexibility for taking advantage of structures in the Hamiltonian or initial state, and additional ideas are needed. To this end, Ref. [27] provides an approach for simulating low-energy states of PSD Hamiltonians that uses quantum phase estimation in combination with the spectral gap amplification technique of Ref. [49]. That algorithm, which was intended for a new demonstration of "fast-forwarding" quantum evolution, is not optimal either because of the considerable overhead of quantum phase estimation with the allowed error.

The scaling of our upper bound more closely resembles that of Ref. [37]. Our results improve upon their upper bounds by logarithmic factors (and we show novel lower bounds

not in Ref. [37]). We summarize the setting of Ref. [37] due to its relevance. For a given Hamiltonian H, we let $H' = (H - E)/\lambda$ be a shifted version of H. Here, $E \in \mathbb{R}$ and $\lambda > 0$ are such that the eigenvalues of H' lie in $[-1, 1]$. Both H and H' share the same eigenstates, and simulating H for time t can be done by simulating H' for "time" $t\lambda$. Assuming access to a unitary T that block-encodes H' (i.e., a unitary that contains H' in one of its blocks), Ref. [37] provides a quantum algorithm for Hamiltonian simulation when the initial state is confined to the subspace corresponding to eigenvalues of H' in $[-1, -1 + \Delta/\lambda]$, where $\Delta \in [0, \lambda]$ specifies the low-energy subspace. The interesting case occurs when $\Delta/\lambda \ll 1$ — formally, when $\Delta = o(\lambda)$ — where the complexity of that algorithm can be an asymptotic improvement over that of more generic Hamiltonian simulation methods, like quantum signal processing (QSP) [36] or related approaches [9, 10]. QSP approximates e^{-itH} using a different block-encoding of H, i.e., a unitary W that contains H/Λ in one of its blocks, for some $\Lambda > 0$. In general, no guarantees are made with regards to the range of eigenvalues associated with the low-energy subspace of H/Λ, other than these being in $[-1, 1]$.

It is unclear how to take advantage of the result in Ref. [37] in general, or whether that algorithm can be further improved. In particular, it is not obvious how to construct a "good" block-encoding T from a given H such that: i) the low energies of H map to eigenvalues of H' in $[-1, -1 + \Delta/\lambda]$ for $\Delta = o(\lambda)$, ii) the process to block-encode H' and implement T is efficient, and iii) the value of λ is not prohibitively large so that the overhead from simulating the shifted Hamiltonian (being linear in λ) does not ruin any other possible improvement coming from the low-energy assumption. For example, if access to the Hamiltonian is given by a block-encoding W of H/Λ rather than T, which is the block-encoding of the shifted version of H, we want to avoid situations in which constructing or implementing T requires too many uses of W. Otherwise methods like QSP, which readily use W, might already be more efficient. Similarly, if after constructing T from W we satisfy $\sqrt{\lambda\Delta} \gg \Lambda$, we will see that QSP using W can again outperform the approach of Ref. [37] that uses T. Hence, if any of these three considerations is not satisfied, it is unclear when — and if — the results of Ref. [37] can be used to successfully exploit the low-energy condition of the initial state.

The goal of this article is to answer some of these open questions. To this end, we first construct a quantum algorithm for Hamiltonian simulation for low-energy states that asymptotically outperforms the time complexity of all known previous methods. Assuming access to the block-encoding T like in Ref. [37], our algorithm computes $e^{-i(t\lambda)H'}|\psi\rangle$ — equivalently, $e^{-itH}|\psi\rangle$ — within specified accuracy. It uses two known primitives, spectral gap amplification [49] and the quantum singular value transform (QSVT) [25], to implement a certain polynomial of T that approximates the correct evolution when acting on $|\psi\rangle$.

To demonstrate the utility of our algorithm, we study when it is possible to efficiently construct the block-encoding T such that our quantum algorithm is useful. Usually, the block-encoding T will have to be constructed from certain operations that provide access to H, like its block-encoding W or a matrix oracle. In these standard access models, we find that there exists a class of Hamiltonians that we deem *gap-amplifiable*, for which Hamiltonian simulation can be performed more efficiently by using our method compared to other techniques like QSP, as long as the initial state is of low energy. Indeed, we will see that there is an equivalence between having efficient access to T with the desired properties and the Hamiltonian H being gap-amplifiable. Roughly, in an interesting regime for the parameters (where the error is not too small), the improvement replaces a cost linear in $t\lambda$ as required by QSP to linear in $t\sqrt{\lambda\Delta}$, where $\Delta \leq \lambda$ is an upper bound on the largest energy. Gap-amplifiable Hamiltonians are PSD and include frustration-free

Hamiltonians as a special case, which appear ubiquitously in physics, quantum computing, and beyond [14, 16, 21, 45]. These Hamiltonians are also known to permit asymptotic improvements in other settings, such as ground state preparation, Gibbs sampling, and quantum linear systems [20, 41, 49, 50]. Moreover, we show that this class of Hamiltonians encompasses all previously considered settings of low-energy simulation studied in the literature [26, 27, 37, 44].

In the asymptotic query complexity of our algorithm, which refers to the number of times the block-encoding T and its inverse are used, the parameters $t\Delta$, $t\lambda$, and $1/\epsilon$ can be arbitrarily large, where $\epsilon > 0$ is the error. Our results show there are only three interesting asymptotic regimes to consider: i) a regime where the evolution time dominates the cost, i.e., where $\log(1/\epsilon) \ll t\Delta \ll t\lambda$; ii) an intermediate regime where the parameters satisfy $t\Delta \ll \log(1/\epsilon) \ll t\lambda$; iii) a regime where error dominates the cost, i.e., where $t\Delta \ll t\lambda \ll \log(1/\epsilon)$. The respective complexities we obtain for the first two regimes are strict improvements over QSP. The time-dominated regime is most natural for applications, since a large $t\Delta \gg 1$ implies that the dynamics in the low-energy sector is not trivial, while still Δ/λ could be vanishingly small. The second regime is distinct in that the resulting query complexity is sublinear in time (we will see that it scales as $\sqrt{t\lambda}$), and hence likely an artificial regime for most applications. The error-dominated regime is subsumed by previous work on general quantum simulation; that is, the complexity matches that of QSP in this regime, which is logarithmic in $1/\epsilon$.

Finally, we study whether our quantum algorithm can be further improved. We prove matching lower bounds for all three asymptotic regimes that show that our quantum algorithm is optimal with respect to t, Δ, and λ, in terms of query complexity. Note that the complexity of our algorithm always produces logarithmic or sublogarithmic scaling in $1/\epsilon$, and previous results readily imply optimality in the error-dominated regime [10]. Our lower bounds are mostly stated in terms of the *query* complexity when given access to the Hamiltonian in standard oracle models. Nevertheless, in the time-dominated regime, which is arguably the most relevant regime to physics and chemistry problems, we go further and also prove a matching lower bound in terms of *gate* complexity.

While an asymptotic benefit is achieved for gap-amplifiable Hamiltonians (equivalently, when given access to T), our algorithm might not offer any improvement with respect to QSP or other known methods in the worst case. That is, if T is not accessible a priori and has to be constructed from some other form of access to H, mapping the low energies of H to eigenvalues near -1 of H' can result in a block-encoding T that is inefficient in general. This observation gives rise to a question of whether our method can be further improved for simulating low-energy states of general Hamiltonians beyond those gap-amplifiable ones. Unfortunately, the answer is negative: we establish matching lower bounds in two common access oracle models — namely, the LCU and sparse matrix model — that show that QSP is already optimal in terms of query complexity, even if the initial state is of low energy. Hence, our results show that the low-energy condition is only useful when additional conditions hold, such as having access to T or being able to construct it efficiently. This is the case for gap-amplifiable Hamiltonians, including frustration-free ones. We emphasize, however, that: i) our lower bounds are for worst-case instances and in the oracle model, and ii) *any* Hamiltonian H (e.g., a sum of Pauli strings) can be transformed into a gap-amplifiable Hamiltonian after simple manipulations (e.g., adding a constant offset), allowing the application of our simulation algorithm. Whether the asymptotic improvement is retained or not for any given H depends on the properties of the resulting parameters after the transformation, especially Δ and λ. For example, our algorithm still gives an asymptotic speedup for Hamiltonians that are perturbatively far from frustration-free Hamiltonians,

even if they are not originally expressed as gap-amplifiable ones.

Last, we note that Ref. [26] recently presented a lower bound for Hamiltonian simulation also within the context of low-energy states. Their proof relies on a state with partial support on an eigenstate of energy $\Theta(\|H\|)$; this high-energy portion of the state solves PARITY, which produces the lower bound. In our work, we show a strictly stronger lower bound in terms of query complexity to a block-encoding of the Hamiltonian; our result applies to a state fully contained in a low-energy subspace of vanishingly small energies compared to $\|H\|$.

We give more details of our contributions next.

1.1 Summary of results

Roughly, our results are four-fold: 1) an improved quantum algorithm for simulating the time evolution of low-energy states, 2) defining a class of Hamiltonians ("gap-amplifiable Hamiltonians") for which our algorithm achieves a speedup in the simulation of low-energy states, 3) a lower bound that shows that no speedup is possible for evolving low-energy states of general Hamiltonians, and 4) a lower bound that shows that our algorithm is optimal for gap-amplifiable Hamiltonians.

To present these results more formally, we let H be the Hamiltonian of an N-dimensional quantum system with Hilbert space $\mathcal{H} \equiv \mathbb{C}^N$, whose evolution we seek to approximate. (We can think of this as being an $n = \log_2(N)$ qubit system.) The ground state energy of H (i.e., its lowest eigenvalue) is E_0 and largest energy (i.e., its largest eigenvalue) is $E_{\max} \geq E_0$. Without loss of generality, we may assume $E_0 \geq 0$, which can always be achieved by shifting the Hamiltonian. We want to simulate the time evolution of an initial state $|\psi\rangle \in \mathcal{H}$, and we say that $|\psi\rangle$ is supported in the low-energy subspace $\mathcal{S}_\Delta$ if and only if it is supported on the subspace corresponding to eigenvalues of H in $[0, \Delta]$. The interesting case for this article is when $\Delta \ll E_{\max}$ and $\Delta > 0$.

Our quantum algorithm uses ancillary systems and is essentially a unitary operation in an enlarged space. In general, we write $\mathbb{1}_M$ to denote the $M \times M$ identity matrix. We begin by stating our results in terms of a block-encoding of $H' = (H - E)/\lambda$, for some $E \in \mathbb{R}$ and $\lambda > 0$, such that the eigenvalues of H' lie within $[-1, 1]$[1]. (To be precise, $H - E$ refers to the matrix $H - E.\mathbb{1}_N$ in our notation.) Moreover, we assume that the low-energy subspace $\mathcal{S}_\Delta$ of H corresponds to eigenvalues of H' within $[-1, -1+\Delta/\lambda]$. This block-encoding is a unitary operator of the form

$$T = \begin{pmatrix} H' & \cdot \\ \cdot & \cdot \end{pmatrix} \tag{1}$$

that acts on an enlarged Hilbert space $\mathbb{C}^M$, $M \geq N$. We also let Π be the orthogonal projector onto $\mathcal{H}$ such that $\Pi T \Pi = H'$; see Sec. 2.3 for details. To use the techniques of the quantum singular value transform (QSVT) [25], we will be interested in quantum circuits that act upon the enlarged Hilbert space associated with a controlled-T operation and its inverse. QSVT can produce a new unitary U that block-encodes a new matrix as $\Pi U \Pi$. For our problem, the new matrix is such that it approximates the time evolution of H when acting on the initial state, i.e., $\Pi U \Pi |\psi\rangle \approx e^{-itH} |\psi\rangle$. Our first result is a quantum algorithm that implements this unitary.

[1]The factor λ also appears in other Hamiltonian simulation methods and is related to an L1 norm of H when presented as a linear combination of unitaries (LCU). Other authors refer to this factor as α.

Theorem 1.1 (Quantum algorithm for time evolution of low-energy states)**.** *Let $\epsilon > 0$ be the error and $t > 0$ be the evolution time. Let T be the unitary of Eq. (1) that block-encodes the Hamiltonian $H' = (H - E)/\lambda$ for some $E \in \mathbb{R}$ and $\lambda > 0$. Let $|\psi\rangle \in \mathcal{S}_\Delta$ be any state supported exclusively on the subspace corresponding to eigenvalues of H' in $[-1, -1 + \Delta/\lambda]$, and assume $t\Delta \geq \epsilon$. Then, for any Γ such that $\Delta \leq \Gamma \leq \lambda$, there exists a quantum algorithm that implements a unitary U and satisfies $|\langle\psi| \Pi U^\dagger \Pi e^{-itH} |\psi\rangle| \geq 1 - \epsilon$, using*

$$\mathcal{O}\left(t\sqrt{\lambda\Gamma} + \sqrt{\frac{\lambda}{\Gamma}} \log \frac{1}{\epsilon}\right) \tag{2}$$

queries to controlled-T and its inverse.

The reason why we use the requirement on the overlap $|\langle\psi| \Pi U^\dagger \Pi e^{-itH} |\psi\rangle| \geq 1 - \epsilon$ to measure the performance of the quantum algorithm is because this expression is amenable to block-encodings. This requirement is equivalent to the quantum circuit producing a state $U |\psi\rangle |0\rangle^{\otimes q} = (1 - \delta) e^{i\alpha} (e^{-itH} |\psi\rangle) |0\rangle^{\otimes q} + |\psi^\perp\rangle$, where $\epsilon \geq \delta \geq 0$, and $|0\rangle^{\otimes q}$ is the initial state of a working register of qubits. The phase α is irrelevant and $|\psi^\perp\rangle$ is an unnormalized state over the full system of $n + q$ qubits such that $\| |\psi^\perp\rangle \| \leq \sqrt{2\epsilon}$.

In the hypothesis, we require $t\Delta \geq \epsilon$; otherwise we could simulate e^{-itH} by applying $\mathbb{1}_N$ (or a global phase) at the cost of no queries, which is uninteresting. Depending on the joint scaling of $t\lambda$, $t\Delta$, and $1/\epsilon$, we may choose different Γ to minimize the query complexity. Because $|\psi\rangle \in \mathcal{S}_\Delta$ also satisfies $|\psi\rangle \in \mathcal{S}_\Gamma$ for any $\Gamma \geq \Delta$, such optimization is feasible. We describe the exhaustive set of possibilities for asymptotic regimes of the parameters; these regimes will be central throughout our work.

1. *Time-dominated regime*: $\log(1/\epsilon) = o(t\Delta)$. We choose $\Gamma = \Delta$, producing an upper bound of $\mathcal{O}(t\sqrt{\lambda\Delta})$.

2. *Intermediate regime*: $\log(1/\epsilon) = o(t\lambda)$ and $t\Delta = o(\log(1/\epsilon))$. We choose $\Gamma = \log(1/\epsilon)/t$, producing an upper bound of $\mathcal{O}(\sqrt{t\lambda \log(1/\epsilon)})$.

3. *Error-dominated regime*: $t\lambda = o(\log(1/\epsilon))$. We choose $\Gamma = \lambda$, producing an upper bound of $\mathcal{O}(\log 1/\epsilon)$.

The time-dominated regime is expected to be the most relevant to applications, where $t\Delta$ must be large enough for interesting dynamics. It gives savings over standard QSP methods, which would instead time-evolve any state $|\psi\rangle \in \mathcal{H}$ (not necessarily of low energy) with access to a block-encoding of H/λ using

$$\mathcal{O}\left(t\lambda + \log \frac{1}{\epsilon}\right) \tag{3}$$

queries (ignoring log log factors) [36]. Since Δ/λ can be asymptotically small, the time-dominated regime shows clear savings over QSP. The intermediate regime represents a fine-tuned region of parameter space, where a particular joint scaling between $t\Delta, 1/\epsilon$, and Δ/λ makes sublinear time dependence possible. For this regime we also obtain an improvement over QSP. In contrast, the error-dominated regime has equivalent cost to QSP.

Our result improves upon Ref. [37] by logarithmic factors in $t\lambda/\epsilon$. It also improves upon the complexity of product formulas in that the asymptotic scaling is linear in t and logarithmic or sublogarithmic in $1/\epsilon$; however, these approaches are not directly comparable. Later, we will also show novel lower bounds that address the time-dominated and

intermediate regimes above to demonstrate optimal scaling with respect to t, Δ, and λ. The error-dominated regime is already addressed by previous lower bounds [9].

Theorem 1.1 and subsequent results assume the initial state to be supported exclusively on the low-energy subspace. However, a similar result still applies even if the support of $|\psi\rangle$ outside $\mathcal{S}_\Delta$ is small, i.e., $\|(\mathbb{1}_N - \mathcal{S}_\Delta)|\psi\rangle\| = \mathcal{O}(\epsilon)$. This is because we are approximating a unitary transformation with unitary operations, and standard bounds for the Euclidean and operator norms can be used to show this result.

We analyze our algorithm in the standard query models considered by QSP and similar algorithms, counting queries to controlled-T and its inverse. Much like QSP and similar algorithms, our quantum algorithm uses a number of other two-qubit gates essentially given by Eq. (2) times the gate complexity needed to perform certain transformation on H or T, as discussed in Sec. 2, including the walk operators to perform polynomial transformations. (Our reference to two-qubit gates also includes one-qubit gates.) This additional factor to the gate complexity will depend on the specific application, but it is expected to be polynomial — and in many cases only constant or linear — in the number of qubits or system size. Detailed analyses of the gate complexity for specific applications are outside the scope of this article.

Our quantum algorithm in Thm. 1.1 is actually constructed from one that simulates a Hamiltonian of the form $H = \lambda A^\dagger A$, for $\lambda > 0$ and A a matrix of dimension $M \times N$ that satisfies $\|A\| \leq 1$. This is an example of what we call a gap-amplifiable Hamiltonian in Sec. 3. The time evolution of H can also be simulated if we have access to a unitary V that is a block-encoding of A and its inverse; that is,

$$V = \begin{pmatrix} A & \cdot \\ \cdot & \cdot \end{pmatrix} . \tag{4}$$

Formally, this block-encoding satisfies $A = \Pi' V \Pi$, where Π' is a projector onto $\mathcal{H}' \equiv \mathbb{C}^M$, since A can be a rectangular matrix. Gap-amplifiable Hamiltonians are PSD ($H \succeq 0$) and readily satisfy $E_0 \geq 0$.

Lemma 1.2 (Quantum algorithm for time evolution of low-energy states of gap-amplifiable Hamiltonians)**.** *Let $\epsilon > 0$ be the error and $t > 0$ be the evolution time. Consider a gap-amplifiable Hamiltonian $H = \lambda A^\dagger A$ acting on $\mathcal{H}$, where $\lambda > 0$ and the dimension of A is $M \times N$. Assume access to a unitary V acting on an enlarged space that block-encodes A as in Eq. (4). Let $|\psi\rangle \in \mathcal{S}_\Delta$ be any state supported exclusively on the subspace corresponding to energies (eigenvalues) of H at most Δ, and assume $t\Delta \geq \epsilon$. Then, for any Γ such that $\Delta \leq \Gamma \leq \lambda$, there exists a quantum algorithm that implements a unitary U and satisfies $|\langle\psi|\Pi U^\dagger \Pi e^{-itH}|\psi\rangle| \geq 1 - \epsilon$, using*

$$\mathcal{O}\left(t\sqrt{\lambda\Gamma} + \sqrt{\frac{\lambda}{\Gamma}}\log\frac{1}{\epsilon}\right) \tag{5}$$

queries to controlled-V and its inverse.

The proof of this lemma is in Sec. 2 and has two steps. First, in Sec. 2.1 we use A and $A^\dagger$ to construct a new Hamiltonian H_{SGA} that acts as the "square root" of H, borrowing ideas from spectral gap amplification [49]. This implies that e^{-itH} can be obtained from $e^{-it(H_{\mathrm{SGA}})^2}$. Second, in Sec. 2.2 we show how to simulate $e^{-it(H_{\mathrm{SGA}})^2}$ using QSVT, which requires the block-encoding of H_{SGA} or A [25]. Hence, Lemma 1.2 already gives an improved Hamiltonian simulation method for low-energy states of the interesting class of Hamiltonians $\lambda A^\dagger A$ when $\Delta = o(\lambda)$.

The main reason why Lemma 1.2 implies Thm. 1.1 is because access to the block-encoding T of Eq. (1) allows us to equivalently express that Hamiltonian H as a gap-amplifiable one, $\lambda A^\dagger A$, up to an irrelevant additive constant. Furthermore, access to A can be efficiently obtained through T, satisfying properties (i)-(iii) discussed in Sec. 1.

Lemma 1.3 (Efficient block-encoding of A)**.** *Let T be the unitary of Eq. (1) that block-encodes Hamiltonian $H' = (H - E)/\lambda$ acting on $\mathcal{H}$, for $E \in \mathbb{R}$ and $\lambda > 0$. Let $|\psi\rangle \in \mathcal{S}_\Delta$ be any state supported exclusively on the subspace associated with eigenvalues of H' in $[-1, -1 + \Delta/\lambda]$ for some $\Delta > 0$. Then, there exists $F \in \mathbb{R}$ and an $M \times N$ matrix A such that $H - F = 2\lambda A^\dagger A$. A block-encoding V of A as in Eq. (4) can be implemented with one query to controlled-$T^\dagger$ and a constant number of two-qubit gates. In addition, $|\psi\rangle$ is supported in the subspace associated with energies at most Δ of $2\lambda A^\dagger A$.*

The proof is given in Sec. 2.3. It is also possible to show the opposite direction of Lemma 1.3: any gap-amplifiable Hamiltonian $\lambda A^\dagger A$ can be efficiently converted into a Hamiltonian H' with the property that, if $|\psi\rangle \in \mathcal{S}_\Delta$, then $|\psi\rangle$ is supported in the subspace of eigenvalues of H' in $[-1, -1 + \Delta/\lambda]$. In addition, a block-encoding T of H' can be constructed from a constant number of queries to the block-encoding of A and $A^\dagger$. This is because we can write $H' = H/\lambda - \mathbb{1}_N = \Pi V^\dagger(\Pi' - \mathbb{1}_M)V\Pi$ in this case, and a block-encoding of H' follows from standard techniques (see Lemma 5.1 for a proof). Hence, our quantum algorithm can give improved results *if and only if* the Hamiltonian can be presented as $H = \lambda A^\dagger A$ (up to additive constants) and certain access to A is given or can be efficiently implemented. The result does not extend to all Hamiltonians because creating access to A can be computationally expensive in general.

More formally, we can write the result of Lemma 1.2 in terms of "general gap-amplifiable Hamiltonians" defined in Sec. 3. These Hamiltonians include the well-known frustration-free Hamiltonians as an example. We will show that they can be equivalently expressed as $\lambda A^\dagger A$, and hence the same quantum algorithm for Lemma 1.2 gives an improved method for the time evolution of low-energy states of these Hamiltonians. This result is captured by Thm. 3.3. We note a difference between Lemma 1.2 and Ref. [49], where the latter work was concerned on amplifying the gap of frustration-free Hamiltonians for faster quantum adiabatic state transformations rather than simulating time-evolution.

Having devised a class of Hamiltonian simulation problems that can benefit from our quantum algorithm, it is natural to ask whether there is an even larger class of Hamiltonians for which simulating the time evolution of their low-energy states can be done more efficiently. To this end, we show a no fast-forwarding result: no improvement over the query complexity of QSP is possible in the worst case, even if the initial state is a low-energy state. We prove this lower bound in two common access models: the LCU model and the sparse matrix model [10, 38]. These models provide a stronger form of access to the Hamiltonian than through its block-encoding, and hence the lower bound automatically applies to the case where only access to the block-encoding is provided. An informal statement of this result in terms of the block-encoding of H/λ follows.

Theorem (No fast-forwarding in the low-energy subspace of generic Hamiltonians, informal)**.** *Let W be a unitary block-encoding of a Hamiltonian H/λ acting on $\mathcal{H}$, such that $\Pi W \Pi = H/\lambda$ for some $\lambda > 0$, and assume $H \succeq 0$. Let $|\psi\rangle \in \mathcal{S}_\Delta$ be any state supported exclusively on the subspace associated with energies (eigenvalues) of H at most Δ, for $\lambda > \Delta > 0$. Then, any quantum algorithm that implements a unitary U such that $|\langle\psi| \Pi U^\dagger \Pi e^{-itH} |\psi\rangle| \geq 2/3$ must use*

$$\Omega(t\lambda) \tag{6}$$

queries to controlled-W and its inverse, even when $\Delta/\lambda = o(1)$.

By assuming U to be unitary, the result also encompasses quantum algorithms where measurements might be interleaved with queries and gates. These can be simulated coherently by enlarging the space using standard methods in quantum computing.

A formal statement of this result that applies to the LCU and sparse matrix models is given in Sec. 4. Even when $\Delta/\lambda = o(1)$, the query complexity in Eq. (6) can be asymptotically worse than the complexity in Thm. 1.1. The proof of this result relies on a reduction from multiple instances of Grover quantum search or, more specifically, from a lower bound on computing PARITY $\circ$ OR of certain bit strings. Among other consequences, this result implies the inefficiency of constructing a "good" block-encoding T from a block-encoding of H/λ in general, as discussed in Sec. 1.

For gap-amplifiable Hamiltonians of the form $\lambda A^\dagger A$ and an initial state with energy at most Δ, it would have been intuitive to suggest that a modification of QSP would have provided an overall query complexity $\mathcal{O}(t\Delta + \log(1/\epsilon))$ for low-energy simulation. This is because the "relevant" energy scale in the problem is Δ rather than λ or $\|H\|$. However, we show that our quantum algorithm is optimal in its dependence on t, Δ, and λ. Informally, our result is as follows.

Theorem (Query complexity lower bound for time evolution of low-energy states of gap-amplifiable Hamiltonians, informal)**.** *Let $\epsilon > 0$ be the error and $t > 0$ be the evolution time. Let H be any Hamiltonian acting on $\mathcal{H}$ and T be a block encoding of $H' = (H - E)/\lambda$, for some $E \in \mathbb{R}$ and $\lambda > 0$, i.e., $\Pi T \Pi = H'$. Let $|\psi\rangle \in \mathcal{S}_\Delta$ be any state supported exclusively on the subspace of eigenvalues of H' in $[-1, -1 + \Delta/\lambda]$ for some $\Delta > 0$, $\Delta \leq \lambda$. Then, any quantum algorithm that implements a unitary U such that $|\langle\psi| \Pi U^\dagger \Pi e^{-itH} |\psi\rangle| \geq 1 - \epsilon$, must use at least Q queries to controlled-T and its inverse given by*

1. $Q = \Omega(t\sqrt{\lambda\Delta})$ *if* $\log(1/\epsilon) = o(t\Delta)$,
2. $Q = \Omega(\sqrt{t\lambda})$ *if* $\log(1/\epsilon) = o(t\lambda)$ *and* $t\Delta = o(\log(1/\epsilon))$,
3. $Q = \Omega(\log(1/\epsilon))$ *if* $t\lambda = o(\log(1/\epsilon))$,

ignoring $\log\log$ *factors in* $1/\epsilon$ *in the third case.*

Formal statements of these results are given in Sec. 5. The above result concerns query complexity given access to T in an oracle model. Since the third regime takes $\Gamma = \lambda$, both upper and lower bounds are already known [10, 36], and no improvement from the low-energy condition follows. In the second regime of intermediate scaling, the lower bound does not exactly match the upper bound $\mathcal{O}(\sqrt{t\lambda \log(1/\epsilon)})$ up to a mild sublogarithmic correction in $1/\epsilon$, and we leave open this discrepancy for error dependence. However, we do not expect a significant improvement in our upper bound in this case, which smoothly transitions to those of the other regimes as we change the asymptotic parameters. The first regime is likely most relevant in terms of applications; examples include quantum phase estimation to determine ground state energies [3, 33] or adiabatic state transformations [11]. Formally, the lower bounds stated above for this regime are proven in two different oracle models, related to the LCU model and the sparse matrix model. In the latter case, we leave open the question of a low-energy simulation algorithm that is optimal in the sparsity of the Hamiltonian; see Ref. [35] for related work. For this regime, we also move outside the oracle model and show our lower bound in terms of gate complexity.

Theorem 1.4 (Gate complexity lower bound for time evolution of low-energy states of gap-amplifiable Hamiltonians, informal)**.** *Let $H = \lambda A^\dagger A$ be any k-local gap-amplifiable Hamiltonian acting on $\mathcal{H}$, where $\lambda > 0$, the dimension of A is $M \times N$, and k is a constant. Let $|\psi\rangle \in \mathcal{S}_\Delta$ be any state supported only on the subspace of eigenvalues of H that are at most Δ for some $\Delta > 0$. For some projector Π, preparing a unitary U satisfying $|\langle\psi|\Pi U^\dagger \Pi e^{-itH}|\psi\rangle| \geq 2/3$, where $t > 0$, must use*

$$\tilde{\Omega}\left(t\sqrt{\lambda\Delta}\right) \tag{7}$$

two-qubit gates, where $\tilde{\Omega}$ ignores logarithmic factors. The algorithm may use an unlimited number of ancilla qubits and the gates may be non-local and come from a possibly infinite gate set.

A formal statement of this result is given in Sec. 6. The proof relies on the Feynman-Kitaev clock Hamiltonian in an argument similar to that of Ref. [31], where the low-energy condition is achieved by controlling the width and momentum of the initial wavepacket. Time evolution evaluates the computation encoded in the Hamiltonian, allowing low-energy simulation to perform arbitrary quantum computations. We show the resulting circuit lower bound using an argument based on computing distinct Boolean functions due to Ref. [28].

The rest of this article is structured as follows. We begin by presenting our quantum algorithm in Sec. 2, where we provide the proof of Thm. 1.1. In Sec. 3, we explain how to improve Hamiltonian simulation of low-energy states of general gap-amplifiable Hamiltonians, including frustration-free Hamiltonians. In Sec. 4, we present formal lower bounds for the complexity of simulating the time evolution of low-energy states of generic Hamiltonians, that is, a no fast-forwarding result. In Sec. 5 we present query lower bounds for the case of gap-amplifiable Hamiltonians. In Sec. 6 we show gate complexity lower bounds for the time-dominated regime of parameters. We give some conclusions in Sec. 7. For completeness, Appendix A presents a detailed discussion on polynomial approximations to the exponential function (corresponding to the evolution operator), and Appendix B gives a lower bound of the K-partition quantum search problem (i.e., computing PARITY$\circ$OR), which we reduced to Hamiltonian simulation to show the lower bounds in Secs. 4 and 5. Appendices C and D contain portions of the proofs for query complexity lower bounds in the sparse model and for gate complexity lower bounds, respectively.

2 Quantum algorithm for time evolution of low-energy states

The proof of Thm. 1.1 is a direct consequence of Lemma 1.2, which is proven in Secs. 2.1 and 2.2 below, and the efficient construction of the block-encoding of A implied by Lemma 1.3, which is proven in Sec. 2.3.

We first sketch the quantum algorithm for the case of gap-amplifiable Hamiltonians of the form $H = \lambda A^\dagger A$, when the initial state is supported on the subspace of energy at most $\Delta > 0$. For any $\Gamma \geq \Delta$, the state is also supported in a subspace of energy at most Γ. Since we will eventually optimize Γ to minimize the query complexity, we will write everything in terms of a state confined to low-energy subspace with energy at most Γ in H, where $0 < \Delta \leq \Gamma \leq \lambda$ restricts the choice of Γ.

We assume access to the block-encoding of A and use it to construct a new Hamiltonian H_{SGA} acting on an enlarged space that functions as the "square root" of H. The quantum algorithm implements the evolution operator $e^{-it\lambda(H_{\mathrm{SGA}}/\sqrt{\lambda})^2}$ that, after tracing out the ancillary systems, gives e^{-itH}. The important property is that we only need to construct

an approximation of $e^{-it\lambda(H_{\mathrm{SGA}}/\sqrt{\lambda})^2}$ when acting on an initial state supported exclusively on a subspace associated with eigenvalues of $H_{\mathrm{SGA}}/\sqrt{\lambda}$ in

$$r_\Gamma := \left[-\sqrt{\frac{\Gamma}{\lambda}}, \sqrt{\frac{\Gamma}{\lambda}}\right] . \tag{8}$$

The problem is analogous to finding a polynomial approximation to $e^{-it\lambda x^2}$, where $x \in r_\Gamma$. This allows us to use a polynomial of lower degree than that needed for the general case (i.e., when the approximation is needed for all $x \in [-1, 1]$). Once the polynomial is obtained, the operation to implement is simply this polynomial where we replace $x \to H_{\mathrm{SGA}}/\sqrt{\lambda}$. QSVT provides the quantum algorithm to implement this operation, which is a sequence of controlled queries to the block-encoding of $H_{\mathrm{SGA}}/\sqrt{\lambda}$ interleaved with other simple two-qubit gates. This is discussed in great detail in Ref. [25]. The block-encoding of $H_{\mathrm{SGA}}/\sqrt{\lambda}$ is implemented using a constant number of queries to the block-encoding of A and its inverse. The query complexity of this algorithm is then determined by the degree of the approximating polynomial.

If access to the Hamiltonian is not initially presented in gap-amplifiable form (i.e., access to the block-encoding of A), but instead access is given to the block-encoding T of H' with low-energy subspace $[-1, -1 + \Delta/\lambda]$, then the Hamiltonian can be expressed in gap-amplifiable form after simple manipulations. This is captured by Lemma 1.3, which states that it is possible to efficiently construct the block-encoding of A using only a constant number of queries to controlled-T and its inverse. Then, the above algorithm can be used for this Hamiltonian as well.

In comparison, Corollary 2 of Ref. [37] describes a quantum algorithm for the same setting that uses

$$\mathcal{O}\left(t\sqrt{\lambda\Gamma}\log^{3/2}\left(\frac{t\lambda}{\epsilon}\right) + \sqrt{\frac{\lambda}{\Gamma}}\log^{5/2}\left(\frac{t\lambda}{\epsilon}\right)\right) \tag{9}$$

queries to controlled-T and its inverse. The additional overhead of $\log^{3/2}(t\lambda/\epsilon)$ originates from Lemma 16 of Ref. [37], which defines a polynomial filter that uniformly maps $[-1, -1+\Gamma/\lambda]$ to $[-1, 0]$. The uniformity of this spectral amplification is not strictly necessary to evaluate time evolution; we instead use a non-uniform method of spectral gap amplification, yielding a more efficient algorithm that removes some logarithmic factors in Eq. (9).

2.1 Spectral gap amplification and block-encoding of H_{SGA}

For the first step of the proof of Lemma 1.2, we use the spectral gap amplification technique of Ref. [49], which we revisit here. Consider a gap-amplifiable Hamiltonian that can be expressed as $H = \lambda A^\dagger A$ for some operator A that can be represented as an $M \times N$ matrix. (In Sec. 3 we generalize this definition.) We define a new $(M+N)\times(M+N)$ Hamiltonian

$$H_{\mathrm{SGA}} := \sqrt{\lambda}\begin{pmatrix}\mathbf{0} & A^\dagger \\ A & \mathbf{0}\end{pmatrix} . \tag{10}$$

We write $\mathbf{0}$ to denote the all-zero matrices whose dimensions are implied from context. The central reason that H_{SGA} is useful is the property

$$(H_{\mathrm{SGA}})^2 = \begin{pmatrix}H & \mathbf{0} \\ \mathbf{0} & \lambda AA^\dagger\end{pmatrix} . \tag{11}$$

The first block of $(H_{\mathrm{SGA}})^2$ contains the Hamiltonian to be simulated, i.e., $\Pi(H_{\mathrm{SGA}})^2\Pi = H$. Hence, we can think of H_{SGA} as the "square root" of H. We diagonalize H_{SGA} explicitly to show that its eigenvalues are the square roots of the eigenvalues of H[2]. Let H have eigenstates $|\psi_j\rangle \in \mathcal{H}$ and eigenvalues $\gamma_j \geq 0$, $j \in [N] := \{1, \ldots, N\}$. Let $|\Psi_j\rangle \in \mathbb{C}^{M+N}$ be the corresponding states in the enlarged space by setting the last $M - N$ amplitudes in the computational basis to zero. Define

$$|\phi_j^{\pm}\rangle = \frac{1}{\sqrt{2}} \left(|\Psi_j\rangle \pm \frac{1}{\sqrt{\gamma_j}} H_{\mathrm{SGA}} |\Psi_j\rangle \right) . \tag{12}$$

These are eigenstates of H_{SGA}, since

$$H_{\mathrm{SGA}} |\phi_j^{\pm}\rangle = \frac{1}{\sqrt{2}} \left(H_{\mathrm{SGA}} |\Psi_j\rangle \pm \frac{1}{\sqrt{\gamma_j}} H_{\mathrm{SGA}}^2 |\Psi_j\rangle \right) \tag{13}$$

$$= \frac{1}{\sqrt{2}} \left(H_{\mathrm{SGA}} |\Psi_j\rangle \pm \sqrt{\gamma_j} |\Psi_j\rangle \right) \tag{14}$$

$$= \pm\sqrt{\gamma_j} |\phi_j^{\pm}\rangle . \tag{15}$$

It is possible to simulate certain transformations on H, such as the time evolution e^{-itH}, using H_{SGA}. For instance, for any $|\psi\rangle \in \mathcal{H}$, a useful identity for our quantum algorithm is

$$\Pi e^{-it(H_{\mathrm{SGA}})^2} |\Psi\rangle = \Pi e^{-it\lambda(H_{\mathrm{SGA}}/\sqrt{\lambda})^2} |\Psi\rangle = e^{-itH} |\psi\rangle , \tag{16}$$

where $|\Psi\rangle \in \mathbb{C}^{M+N}$ is the version of $|\psi\rangle \in \mathcal{H}$ in the enlarged space (after setting the last $M - N$ amplitudes in the computational basis to zero). If access to a block-encoding of $H_{\mathrm{SGA}}/\sqrt{\lambda}$ is given, the operator e^{-itH} can be simulated via QSVT using such a block-encoding.

A technical challenge to construct this block-encoding is that A and $A^\dagger$ can be, in principle, rectangular matrices. It is possible, however, to pad these matrices with zeroes to make them square. Without loss of generality, we can take $M \geq N$ by adding zeros when $M < N$. Rather than considering A, we can then consider the $M \times M$ matrix

$$\mathbf{A} := \left(\begin{array}{c|ccc} A & 0 & \ldots & 0 \\ & \vdots & \ldots & \vdots \\ & 0 & \ldots & 0 \end{array} \right) , \tag{17}$$

where the last $M-N$ columns are zero. Accordingly, instead of using $H_{\mathrm{SGA}}/\sqrt{\lambda}$ to simulate the time evolution of H as in Eq. (16), we can use $\mathbf{H}_{\mathrm{SGA}}/\sqrt{\lambda} := |0\rangle\langle 1| \otimes \mathbf{A}^\dagger + |1\rangle\langle 0| \otimes \mathbf{A}$, where $\mathbf{H}_{\mathrm{SGA}}$ is dimension $2M \times 2M$. To describe the quantum algorithm as a sequence of operations acting on qubit systems, we assume without loss of generality that $M = 2^m$ and $N = 2^n$ are powers of two, for $m \geq n$. Then, $\mathbf{H}_{\mathrm{SGA}}$ acts on the space of $m + 1$ qubits and Eq. (16) implies

$$e^{-it(\mathbf{H}_{\mathrm{SGA}})^2} |0\rangle^{\otimes m+1-n} |\psi\rangle = e^{-it\lambda(\mathbf{H}_{\mathrm{SGA}}/\sqrt{\lambda})^2} |0\rangle^{\otimes m+1-n} |\psi\rangle = |0\rangle^{\otimes m+1-n} \left(e^{-itH} |\psi\rangle\right) . \tag{18}$$

[2]Among other consequences, this property implies that the spectral gap of H_{SGA} is larger than that of H, allowing for faster adiabatic state transformations, which was the problem studied in Ref. [49]

Time evolution with $H = \lambda A^\dagger A$ can then be simulated via QSVT when access to a block-encoding $V_{\rm SGA}$ of $\mathbf{H}_{\rm SGA}/\sqrt{\lambda}$ is allowed. The $m+1-n$ ancillas can be discarded at the end.

Let V be the block-encoding of A in Eq. (4) that satisfies $\Pi' V \Pi = A$, where Π' and Π are the orthogonal projectors onto $\mathcal{H}' \equiv \mathbb{C}^M$ and $\mathcal{H} \equiv \mathbb{C}^N$, respectively. Assume that V acts on the space $\mathbb{C}^Q$ of $q = \log_2 Q \geq m$ qubits. Let $P' := |0\rangle\langle 0|^{\otimes m-n} \otimes \mathbb{1}_N$ be the projector associated with Π' that acts on $\mathcal{H}'$ and $P'' := |0\rangle\langle 0|^{\otimes q-n} \otimes \mathbb{1}_N$ a projector in $\mathbb{C}^Q$. Then,

$$\mathbf{A} = \Pi' V \Pi' P' = \Pi' V P'' \Pi' \,, \tag{19}$$

where $\Pi' V \Pi'$ corresponds to the first $M \times M$ dimensional block of V and VP'' is such that the first N columns are those of V and the last $Q-N$ columns are zero; essentially, we constructed a block-encoding of $\mathbf{A}$, which is the matrix in Eq. (17). We can write

$$\mathbf{H}_{\rm SGA}/\sqrt{\lambda} = \left(|0\rangle\langle 0| \otimes \mathbf{A}^\dagger + |1\rangle\langle 1| \otimes \mathbf{A}\right)(X \otimes \mathbb{1}_M) \tag{20}$$

$$= \Pi'' \left(|0\rangle\langle 0| \otimes P'' V^\dagger + |1\rangle\langle 1| \otimes V P''\right)(X \otimes \mathbb{1}_M)\Pi'' \,, \tag{21}$$

where X is the one-qubit Pauli flip operator, and Π'' is now the orthogonal projector onto the space of $m+1$ qubits, i.e., $\mathbb{C}^{2M}$. This is the projector associated with the subspace where the m qubits associated with $\mathbb{C}^M$ are in $|0\rangle^{\otimes m}$. Equation (21) is not yet a proper block-encoding of $\mathbf{H}_{\rm SGA}/\sqrt{\lambda}$ because P'' is not unitary. However, P'' can be expressed as an LCU of the form $P'' = (U_{P''} + \mathbb{1}_Q)/2$, where $U_{P''} := 2P'' - \mathbb{1}_Q$ is a unitary acting on $\mathbb{C}^Q$. There exist standard techniques to construct a block-encoding of an LCU; see, for example, the results in Lemma 6 of Ref. [18] or Ref. [38]. More specifically, we can define the unitary that acts on a space of $q+2$ qubits

$$V' := |00\rangle\langle 00| \otimes U_{P''} V^\dagger + |01\rangle\langle 01| \otimes V^\dagger + |10\rangle\langle 10| \otimes V U_{P''} + |11\rangle\langle 11|\, V \,, \tag{22}$$

and also define

$$V_{\rm SGA} := (\mathbb{1}_2 \otimes \mathrm{H} \otimes \mathbb{1}_Q) V' (X \otimes \mathrm{H} \otimes \mathbb{1}_Q) \,, \tag{23}$$

where H is the one-qubit Hadamard gate that implements $\mathrm{H}|0\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and $\mathrm{H}|1\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. Note that the dimension of V' and $V_{\rm SGA}$ is $4Q \times 4Q$ in this example. Then, if Π'' is the orthogonal projector onto $\mathbb{C}^{2M}$, which is the space associated with $\mathbf{H}_{\rm SGA}$, we obtain

$$\Pi'' V_{\rm SGA} \Pi'' = \mathbf{H}_{\rm SGA}/\sqrt{\lambda} \,. \tag{24}$$

The subspace associated with Π'' is now that where qubits 2 through $m+2$ are in $|0\rangle^{\otimes m+1}$.

Equation (23) is the desired block-encoding of $\mathbf{H}_{\rm SGA}/\sqrt{\lambda}$. To implement it, we need to apply V' once, which requires applying controlled-V and controlled-$V^\dagger$ once. The additional gate complexity is dominated by that of implementing the controlled-$U_{P''}$ operation. This gate complexity is $\mathcal{O}(q)$ using standard techniques.

Having constructed a block-encoding of $\mathbf{H}_{\rm SGA}/\sqrt{\lambda}$, it remains to show: i) how to simulate the time evolution of $H = \lambda A^\dagger A$ using Eq. (18), and ii) how to equivalently express any Hamiltonian as $\lambda A^\dagger A$ from its block-encoding T. These are discussed next.

2.2 Polynomial approximation of e^{-itx^2} and QSVT

For the second step of the proof of Lemma 1.2, we follow an approach similar to Hamiltonian simulation via QSVT. Briefly, we review the standard approach for simulating e^{-itH} given access to a block-encoding of H/λ, for any Hamiltonian H [25]. QSVT starts from proper even and odd degree polynomial approximations to the trigonometric functions $\cos(t\lambda x)$ and $\sin(t\lambda x)$, which can be obtained using a Jacobi-Anger expansion (Thm. 57 of Ref. [25]). Here, $x \in \mathbb{R}$ plays the role of an eigenvalue of H/λ, and after replacing $x \to H/\lambda$ in the polynomials, we produce approximations of $\cos(tH)$ and $\sin(tH)$, respectively. QSVT can then implement these polynomials using the block-encoding of H/λ a number of times that depends on the degree of the polynomials. They can be later combined through an LCU, ultimately implementing an approximation of $e^{-itH} = \cos(tH) - i\sin(tH)$. See Thm. 58 of Ref. [25] for more details.

We will now modify this procedure to accommodate a block-encoding of $\mathbf{H}_{\mathrm{SGA}}/\sqrt{\lambda}$ instead of H/λ, applicable to gap-amplifiable Hamiltonians. For our problem, we start from even and odd polynomial approximations to $\cos(t\lambda x^2)$ and $\sin(t\lambda x^2)$ instead. Here, $x \in \mathbb{R}$ plays the role of a certain eigenvalue of $\mathbf{H}_{\mathrm{SGA}}/\sqrt{\lambda}$, which coincides with an eigenvalue of $\pm\sqrt{H/\lambda}$. After replacing $x \to \mathbf{H}_{\mathrm{SGA}}/\sqrt{\lambda}$, these polynomials produce approximations of $\cos(t(\mathbf{H}_{\mathrm{SGA}})^2)$ and $\sin(t(\mathbf{H}_{\mathrm{SGA}})^2)$. QSVT can then implement these polynomials using the block-encoding of $\mathbf{H}_{\mathrm{SGA}}/\sqrt{\lambda}$. These can be later combined through an LCU, ultimately implementing an approximation of $e^{-it(\mathbf{H}_{\mathrm{SGA}})^2} = \cos(t(\mathbf{H}_{\mathrm{SGA}})^2) - i\sin(t(\mathbf{H}_{\mathrm{SGA}})^2)$. This operation also simulates e^{-itH} due to Eq. (18). Importantly, since we assume the initial state $|\psi\rangle$ to be supported only on the subspace corresponding to eigenvalues of $H = \lambda A^\dagger A$ at most $\Gamma > 0$, we only need to consider these polynomial approximations for the domain $x \in r_\Gamma$ in Eq. (8). That is, for given $\epsilon \geq 0$, we seek polynomials $P_{\epsilon,\cos}$ and $P_{\epsilon,\sin}$ satisfying

$$|P_{\epsilon,f}(x)| \leq 1 \;\forall\; x \in [-1,1]\,, \tag{25}$$

$$\max_{x\in r_\Gamma} |P_{\epsilon,f}(x) - f(x)| \leq \epsilon\,, \tag{26}$$

for $f(x) = \cos(t\lambda x^2)$ and $f(x) = \sin(t\lambda x^2)$. The techniques of piecewise polynomial constructions outlined in Ref. [25] are useful for this task. In particular, we use a slightly modified version of Corollary 66 of Ref. [25].

Lemma 2.1 (Piecewise polynomial approximations based on a local Taylor series)**.** *Let $r \in (0,1]$, $\delta \in (0,r]$, and let $f : [-r-\delta, r+\delta]$ be such that $f(x) = \sum_{k=0}^{\infty} a_k x^k$ for all $x \in [-r-\delta, r+\delta]$. Choose any $B > 0$ such that $B \geq \sum_{k=0}^{\infty} |a_k|(r+\delta)^k$. For $\epsilon \in (0, 3B/2)$, there is an efficiently computable polynomial P of degree $\mathcal{O}(\frac{1}{\delta}\log\frac{B}{\epsilon})$ such that*

$$\|f(x) - P(x)\|_{[-r,r]} \leq \epsilon \quad \text{and} \quad \|P(x)\|_{[-1,1]} \leq \epsilon + \|f(x)\|_{[-r-\frac{\delta}{2}, r+\frac{\delta}{2}]}\,. \tag{27}$$

We provide a self-contained proof of this Lemma in Appendix A. Here, we neglect the cost of classical processing (e.g., for finding the polynomial coefficients), which remains efficient; it is discussed briefly in Appendix A and in depth in Ref. [25].

For the specific case of the sine and cosine functions, Lemma 2.1 implies the following.

Lemma 2.2. *Let $\epsilon > 0$ be the error, $t \in \mathbb{R}$, $\Gamma > 0$, and $\lambda \geq \Gamma$. Then, there exists polynomial approximations $P_{\epsilon,f}(x)$ to $f(x) = \sin(t\lambda x^2)$ and $f(x) = \cos(t\lambda x^2)$ of degree*

$$\mathcal{O}\left(t\sqrt{\lambda\Gamma} + \sqrt{\frac{\lambda}{\Gamma}}\log\frac{1}{\epsilon}\right) \tag{28}$$

that satisfy Eqs. (25) and (26) with error 2ϵ.

Proof. Let $r = \delta = \sqrt{\Gamma/\lambda}$. The Taylor expansions of $\sin(t\lambda x^2)$ and $\cos(t\lambda x^2)$ at $x = r + \delta$ are such that

$$B_{\sin} = \sinh\left[\lambda t\left(2\sqrt{\Gamma/\lambda}\right)^2\right] < e^{4\Gamma|t|}, \quad B_{\cos} = \cosh\left[\lambda t\left(2\sqrt{\Gamma/\lambda}\right)^2\right] < e^{4\Gamma|t|}\,. \tag{29}$$

We can then choose $B = \exp(4\Gamma|t|)$ for these cases. Applying Lemma 2.1 to polynomial approximations of $\sin(t\lambda x^2)$ and $\cos(t\lambda x^2)$ with error ϵ, we obtain $\frac{1}{\delta}\log(B/\epsilon) = \sqrt{\frac{\lambda}{\Gamma}}\log\frac{e^{4\Gamma t}}{\epsilon}$. Hence, for $t > 0$, a polynomial of degree

$$\mathcal{O}\left(t\sqrt{\lambda\Gamma} + \sqrt{\frac{\lambda}{\Gamma}}\log\frac{1}{\epsilon}\right) \tag{30}$$

suffices to produce the approximating polynomials for the sine and cosine functions.

However, Lemma 2.1 only provides normalization up to $\|P(x)\|_{[-1,1]} \leq 1 + \epsilon$ since sin and cos satisfy $\|f(x)\|_{[-r-\delta/2,r+\delta/2]} \leq 1$. Since QSVT can implement polynomials normalized to 1 or less, we rescale these polynomial approximations by $\frac{1}{1+\epsilon}$. By the triangle inequality

$$\left\|\frac{1}{1+\epsilon}P(x) - f(x)\right\|_{[-r,r]} \leq \frac{1}{1+\epsilon}\|P(x) - f(x)\|_{[-r,r]} + \frac{\epsilon}{1+\epsilon}\|f(x)\|_{[-r,r]} \leq \frac{2\epsilon}{1+\epsilon} < 2\epsilon\,. \tag{31}$$

Consequently, this rescaling only increases the approximation error from ϵ to 2ϵ in the worst case. The result is two polynomials that approximate $\cos(t\lambda x^2)$ and $\sin(t\lambda x^2)$ within error 2ϵ in the domain $x \in r_\Gamma$. □

We now have suitable polynomial approximations of $\sin(t\lambda x^2)$ and $\cos(t\lambda x^2)$. We can then use QSVT to simulate the Hamiltonian as explained above. This requires access to the block-encoding V_{SGA} of $\mathbf{H}_{\mathrm{SGA}}/\sqrt{\lambda}$; see Eq. (23). To apply the corresponding polynomials, QSVT uses the controlled-V_{SGA} and its inverse a number of times determined by the degree of the polynomials, which was shown to be $\mathcal{O}(t\sqrt{\lambda\Gamma} + \sqrt{\frac{\lambda}{\Gamma}}\log\frac{1}{\epsilon})$. Together with Sec. 2.1, this completes the proof of Lemma 1.2. We refer to Ref. [25] for more details on QSVT.

2.3 Efficient implementation of V

The main result in Thm. 1.1 is stated in terms of the block-encoding T, but our quantum algorithm is essentially that for gap-amplifiable Hamiltonians, i.e., Lemma 1.2. Now we prove Lemma 1.3, which relates T to the block-encoding V of a corresponding A.

The block-encoding T is an $M \times M$ unitary matrix, where we assume $M \geq N$ without loss of generality. It satisfies $\Pi T \Pi = H' = (H - E)/\lambda$, for some $E \in \mathbb{R}$ and $\lambda > 0$; see Eq. (1). Here, Π is the projector onto $\mathcal{H} = \mathbb{C}^N$. Let

$$A := \left(\frac{T^\dagger + \mathbb{1}_M}{2}\right)\Pi \tag{32}$$

be of dimension $M \times N$. It follows that,

$$2\lambda A^\dagger A = 2\lambda\Pi\left(\frac{T + \mathbb{1}_M}{2}\right)\left(\frac{T^\dagger + \mathbb{1}_M}{2}\right)\Pi \tag{33}$$

$$= \lambda(\mathbb{1}_N + (H - E)/\lambda) \tag{34}$$

$$= H - F\,, \tag{35}$$

where $F = E - \lambda$. This proves that, when access to T and $T^\dagger$ is given, the Hamiltonian can be equivalently expressed as a gap-amplifiable Hamiltonian (up to an additive constant).

Then, to simulate H, or equivalently $2\lambda A^\dagger A$, using the approach for gap-amplifiable Hamiltonians of Lemma 1.2 (also discussed in Secs. 2.1 and 2.2), we need to construct the block-encoding V of A defined in Eq. (32). This satisfies $\Pi' V \Pi = A$ for corresponding projectors Π' and Π. The only remaining challenge is that the term in brackets in Eq. (32) is an LCU rather than unitary. Nevertheless, we can again use standard techniques to block-encode an LCU. For example, we can define

$$V := (\mathrm{H} \otimes \mathbb{1}_M) \left(|0\rangle\langle 0| \otimes T^\dagger + |1\rangle\langle 1| \otimes \mathbb{1}_M \right) (\mathrm{H} \otimes \mathbb{1}_M) \,, \tag{36}$$

where H is the Hadamard gate, which implies

$$\langle 0| \, V \, |0\rangle = \left(\frac{T^\dagger + \mathbb{1}_M}{2} \right) . \tag{37}$$

Equivalently, $\Pi' V \Pi = A$, which is the desired result. In the notation of Sec. 2.1, the unitary V of Eq. (36) acts on the space $\mathbb{C}^Q$ of $q = m + 1$ qubits, and Π' is the projector that takes $\mathbb{C}^{2M}$ to $\mathbb{C}^M$.

Implementing V requires using a controlled-$T^\dagger$ operation once, in addition to a constant number of two-qubit gates. Also, if the initial state $|\psi\rangle$ is supported on the subspace associated with eigenvalues of H' in $[-1, -1 + \Delta/\lambda]$, then it is supported on the subspace associated with energies at most Δ of the Hamiltonian $2\lambda A^\dagger A$. This proves Lemma 1.3 that, together with Lemma 1.2, imply Thm. 1.1. We note that the factor of 2 that appears in the block-encoding $2\lambda A^\dagger A$ occurs because Δ is restricted to $[0, 2\lambda]$ to ensure the block-encoding T has eigenvalues in $[-1, 1]$. Although we took $H = \lambda A^\dagger A$ in the previous subsections, the constant factor does not affect the reported query complexity of $\mathcal{O}(t\sqrt{\lambda\Gamma} + \sqrt{\lambda/\Gamma}\log(1/\epsilon))$ in Thm. 1.1.

3 Quantum algorithm for time evolution of low-energy states of general gap-amplifiable Hamiltonians

Our quantum algorithm for simulating time evolution of low-energy states is based on that for simulating Hamiltonians that can be expressed as $H = \lambda A^\dagger A$, for $\lambda > 0$. In this case, if the initial state is of low-energy $\Delta \ll \lambda$ (e.g., $\Delta = o(\lambda)$), Lemma 1.2 can give an asymptotic improvement over generic methods like QSP. We called these "gap-amplifiable Hamiltonians", but this definition can be extended to also address a more general case where $H = \sum_{l=1}^{L} \lambda_l A_l^\dagger A_l$, $\lambda_l > 0$ for all $l \in [L]$, and A_l are matrices of dimension $M \times N$. Examples of such Hamiltonians are frustration-free Hamiltonians [14, 16, 21, 45], where each term $\lambda_l A_l^\dagger A_l$ might correspond to a local interaction between spins or other physical systems. In this section we show how to use the quantum algorithm of Lemma 1.2 to also give an improved simulation method for these "general gap-amplifiable Hamiltonians".

3.1 General gap-amplifiable Hamiltonians

We begin by providing a general definition of a (λ, L)-gap-amplifiable Hamiltonian.

Definition 3.1 (General gap-amplifiable Hamiltonian)**.** *A Hamiltonian $H \succeq 0$ acting on $\mathcal{H} \equiv \mathbb{C}^N$ is said to be* (λ, L)-gap-amplifiable *if there exists $L \geq 1$ such that: i) H can be expressed as $H = \sum_{l=0}^{L-1} \lambda_l A_l^\dagger A_l$, where $\lambda_l > 0$ for all $l \in [L]$, ii) efficient access to block-encodings of each A_l are provided, and iii) $\sum_{l=0}^{L-1} \lambda_l \leq \lambda$.*

The case $L = 1$ was analyzed in Sec. 2. In general, the operators A_l can be expressed as rectangular matrices of dimension $M \times N$. A block-encoding V_l of A_l might involve two different projectors Π and Π' onto $\mathbb{C}^N$ and $\mathbb{C}^M$, respectively, such that $\Pi' V_l \Pi = A_l$. The unitaries V_l might act on a larger space $\mathbb{C}^Q$, where $Q \geq \max(M, N)$.

Each block-encoding V_l will in practice incur some cost to construct. This cost can be measured as a query complexity to an oracle that provides some other form of access to A_l. Later, in Secs. 4.1 and 5.1, we will discuss two common access models for quantum algorithms:

i) the LCU model, which provides coefficients $\beta_{l,j} > 0$ and access to unitaries $U_{l,j}$, $l \in [L]$ and $j \in [J]$, such that $A_l = \frac{1}{\beta_l} \sum_{j=1}^{J} \beta_{l,j} U_{l,j}$, for $\beta_l := \sum_{j=1}^{J} \beta_{l,j}$, and

ii) the sparse matrix model, which provides access to the positions and values of nonzero matrix elements of each sparse A_l, $l \in [L]$.

In our definition, we assume *efficient access* to the block-encoding of A_l; this refers to requiring only a constant number of queries to the oracles in the LCU or sparse matrix model to construct each V_l.

3.2 Quantum algorithm and complexity

To use the same algorithm as that of Lemma 1.2, we begin by showing a simple reduction from general (λ, L)-gap-amplifiable Hamiltonians to $(\lambda, 1)$-gap-amplifiable Hamiltonians, i.e., $\lambda A^\dagger A$.

Lemma 3.2 (Reduction to $(\lambda, 1)$-gap amplifiable Hamiltonians)**.** *Consider a (λ, L)-gap-amplifiable Hamiltonian H acting on $\mathcal{H} \equiv \mathbb{C}^N$ and assume access to the unitary $W = \sum_{l=0}^{L-1} |l\rangle\langle l| \otimes V_l$ and its inverse. Each unitary V_l is the block-encoding of the $M \times N$ dimensional A_l and satisfies $A_l = \Pi' V_l \Pi$, where Π' and Π are orthogonal projectors onto $\mathcal{H}' \equiv \mathbb{C}^M$ and $\mathcal{H} \equiv \mathbb{C}^N$, respectively. Then, there exists a matrix A of dimension $(LM) \times N$ such that H can be alternatively expressed as $H = \lambda A^\dagger A$. A block-encoding V of A can be constructed using controlled-W once, and additional $\mathcal{O}(L)$ two-qubit gates. Moreover, if $|\psi\rangle \in \mathcal{S}_\Delta$ is any state supported exclusively on the subspace corresponding to energies (eigenvalues) of H in $[0, \Delta]$, then $|\psi\rangle$ is supported exclusively on a subspace corresponding to energies at most Δ of $\lambda A^\dagger A$.*

Proof. The statement on the support of $|\psi\rangle$ follows directly from the fact that $\lambda A^\dagger A$ is also an expression for H, which is PSD by assumption. Since this is (λ, L)-gap-amplifiable, it can be expressed as $H = \sum_{l=0}^{L-1} \lambda_l A_l^\dagger A_l$, where $\lambda_l > 0$ and each A_l is of dimension $M \times N$. We define

$$A := \sum_{l=0}^{L-1} \sqrt{\frac{\lambda_l}{\lambda}} |l\rangle \otimes A_l \,, \tag{38}$$

which is of dimension $M' \times N$, $M' := LM$. Here, $\lambda = \sum_{l=0}^{L-1} \lambda_l$ and note that

$$A^\dagger A = \frac{1}{\lambda} \sum_{l=0}^{L-1} \lambda_l A_l^\dagger A_l \,. \tag{39}$$

This implies $\|A\| \leq \frac{1}{\lambda} \sum_l \lambda_l \|A_l^\dagger A_l\| \leq 1$ and

$$H = \lambda A^\dagger A \,. \tag{40}$$

We seek a block-encoding of this A, which would allow us to simulate H as explained in Sec. 2.1. Assume $L = 2^l$ without loss of generality and let G be a unitary that performs the map

$$G : |0\rangle \to \sum_{l=0}^{L-1} \sqrt{\frac{\lambda_l}{\lambda}} |l\rangle \ , \tag{41}$$

where $|0\rangle = |0\rangle^{\otimes l}$ in this case. This can be implemented with $\mathcal{O}(L)$ two-qubit gates using standard methods. We also assume that each block-encoding V_l of A_l acts on the space of q qubits, i.e., on $\mathbb{C}^Q$ for $Q = 2^q$. Note that

$$W(G|0\rangle \otimes \mathbb{1}_Q) = \sum_{l=0}^{L-1} \sqrt{\frac{\lambda_l}{\lambda}} |l\rangle \otimes V_l \ . \tag{42}$$

This implies that

$$\Pi' W(G \otimes \mathbb{1}_Q)\Pi = \sum_{l=0}^{L-1} \sqrt{\frac{\lambda_l}{\lambda}} |l\rangle \otimes A_l = A \ . \tag{43}$$

The block-encoding is then $V = W(G \otimes \mathbb{1}_Q)$, which requires one use of W and G. □

The ability to simulate low-energy states of a general (λ, L)-gap-amplifiable Hamiltonian follows as a consequence of the above results. While Thm. 1.1 is phrased in terms of queries to a Hamiltonian block-encoding with low-energy subspace $[-1, -1 + \Delta/\lambda]$, the phrasing below is more natural for settings such as frustration-free systems.

Theorem 3.3 (Quantum algorithm for time evolution of low-energy states of general gap-amplifiable Hamiltonians)**.** *Let $\epsilon > 0$ be the error and $t > 0$ be the evolution time. Consider a (λ, L)-gap-amplifiable Hamiltonian $H = \sum_{l=0}^{L-1} \lambda_l A_l^\dagger A_l$ acting on $\mathcal{H} \equiv \mathbb{C}^N$, where $\lambda_l > 0$, and each A_l has dimension $M \times N$. Assume access to unitaries V_l block-encoding each A_l. Let $|\psi\rangle \in \mathcal{S}_\Delta$ be any state supported exclusively on the subspace associated with energies (eigenvalues) of H at most $\Delta > 0$, and assume $t\Delta \geq \epsilon$. Then, for any Γ such that $\Delta \leq \Gamma \leq \lambda$, there exists a quantum algorithm that implements a unitary U and satisfies $|\langle\psi| \Pi U^\dagger \Pi e^{-itH} |\psi\rangle| \geq 1 - \epsilon$, using*

$$\mathcal{O}\left(t\sqrt{\lambda\Gamma} + \sqrt{\frac{\lambda}{\Gamma}} \log \frac{1}{\epsilon}\right) \tag{44}$$

queries to each controlled-V_l and its inverse.

Proof. The result is an immediate consequence of Lemma 3.2, to construct a block-encoding of A when expressing the Hamiltonian as $H = \lambda A^\dagger A$, and Lemma 1.2 to simulate the latter. The number of queries to $V = W(G \otimes \mathbb{1}_Q)$ appearing in Lemma 3.2 is given by Eq. (5). Each such query uses W once, and this uses each controlled-V_l once. □

When the initial state is such that $\Delta \ll \lambda$, or $\Delta = o(\lambda)$, and the Hamiltonian is gap-amplifiable, the result is again an improvement over QSP. Because any Hamiltonian in $\mathcal{H}$ can be expressed as a sum of PSD terms $\sum_{l=0}^{L-1} \lambda_l A_l^\dagger A_l$ up to a constant, the algorithm of Thm. 3.3 can always be applied. However, in the following section, we will see that the low-energy assumption does not result in an improvement over QSP for *all* Hamiltonians, even in the LCU and sparse matrix access models. Our low-energy simulation algorithm does

not provide any benefit in general, because offsetting each term in a given H to make it PSD may place the low-energy sector of H to a relatively large energy sector of the shifted Hamiltonian. In other words, Δ/λ can become a constant after the shift, preventing any fast-forwarding from the low-energy assumption alone.

Nonetheless, physically relevant classes of Hamiltonians are already (λ, L)-gap-amplifiable with reasonably small Δ/λ. We consider the example of a frustration-free Hamiltonian $H_{\mathrm{FF}} = \sum_{l=0}^{L-1} \pi_l$. Each $\pi_l \succeq 0$ might refer to a local term in the Hamiltonian, e.g., involving a few qubits. The ground state energy of H_{FF} is $E_0 = 0$. With some preprocessing, it is possible to find the λ_l's and A_l's such that $\pi_l = \lambda_l A_l^\dagger A_l$ and construct the block-encodings V_l, which are also local operators in this example. Implementing the operation W that applies the controlled-V_l will incur in some gate cost, which is $\mathcal{O}(L)$ for the case of local Hamiltonians. Hence, the total gate complexity to simulate the time evolution of a low-energy state of H_{FF} using our quantum algorithm is $\mathcal{O}(Lt\sqrt{\lambda\Delta})$, assuming constant precision. Applying QSP instead to this problem would have resulted in gate complexity $\mathcal{O}(Lt\lambda)$ for this case.

Similarly, if a Hamiltonian is nearly frustration-free, some benefit can be achieved. To illustrate this, we consider the example where $H = \delta P + H_{\mathrm{FF}}$, for some Pauli string P, and $0 \leq \delta \ll 1$. Hence, $H + \delta = \delta(P + \mathbb{1}_N) + H_{\mathrm{FF}}$ is a sum of PSD terms and a $(\lambda + 2\delta, L + 1)$-gap-amplifiable Hamiltonian. For example, we can choose $A_0 = (P + \mathbb{1}_N)/2$ for the first operator, so that $A_0^\dagger A_0 = (P + \mathbb{1}_N)/2$. For an initial state $|\psi\rangle \in \mathcal{S}_{\Delta_0}$ of H, the low-energy cutoff must now satisfy $\Delta \geq \Delta_0 + \delta$ due to the offset. Since $\lambda \approx \lambda + 2\delta$, we observe that this nearly frustration-free system can still benefit from our quantum algorithm. The additional query complexity would only be $\mathcal{O}(t\delta\sqrt{\lambda\Delta_0})$ for $\delta \ll 1$. This observation hints at improvements over QSP for Hamiltonians that are not necessarily gap-amplifiable; although the asymptotic complexity might not change for a particular example, the actual gate complexity by using our algorithm could still give an improvement.

4 No fast-forwarding theorem for simulating time evolution of low-energy states in general

In general, the no fast-forwarding theorems of Refs. [4, 8, 28] establish a lower bound of $\Omega(t\|H\|)$ queries or gates needed to simulate the time evolution e^{-itH}. The result of Ref. [8], in particular, follows from a reduction of the parity problem to Hamiltonian simulation. It is shown that simulating the dynamics of a particular quantum system in one dimension (i.e., a quantum walk) allows one to compute the parity of a bit string of increasing length n after time $t = \Theta(n)$. A lower bound on the query complexity of parity implies a lower bound on the number of queries needed for Hamiltonian simulation. However, that proof cannot directly extend to simulating low-energy states, since the initial state is not fully confined to a low-energy subspace.

Without relying on the reduction from parity, we may gain some intuition for the simulation lower bound as follows. Hamiltonian simulation methods like QSP implement a polynomial of a unitary, and the query complexity coincides with the degree of this polynomial. In the worst case, the state $|\psi\rangle$ can be supported on states of energy as high as $\|H\|$ or as small as $-\|H\|$. If we use QSP, we would need to start from a polynomial approximation of $e^{-i\phi}$, where $|\phi|$ can be as large as $t\|H\|$. For constant error, this is achieved by a polynomial of degree $\Omega(t\|H\|)$. Then, QSP must query the Hamiltonian $\Omega(t\|H\|)$ times. When confined to the low-energy subspace with largest eigenvalue Δ, this analysis suggests a Hamiltonian simulation lower bound of $\Omega(t\Delta)$ instead. However, we will show that this intuition is incorrect.

As discussed in the previous section, our quantum algorithm can apply to any Hamiltonian. However, Thm. 1.1 improves upon QSP when the block-encoding T of $H' = (H - E)/\lambda$ can be accessed and has the property that the low-energy subspace of H corresponds to the subspace associated with eigenvalues of H' near -1. From simple inspection, it is not obvious how to efficiently construct such T (and H') when query access to the original H is given (e.g., in form of a block-encoding of H/λ). Indeed, we find that no speedup is possible for simulating the time evolution of low-energy states of general Hamiltonians, and that the asymptotic complexity of QSP in Eq. (3) remains optimal. Our result applies to any Hamiltonian simulation method that assumes access to H in the LCU model or in the sparse matrix model, which we now formally introduce.

4.1 Access models

We seek to prove that when access to H is given through standard access models it is not generically possible to improve the query complexity of methods like QSP under the low-energy assumption. To this end, we discuss two of the most common models used in Hamiltonian simulation to provide access to H: the LCU and sparse matrix models.

The LCU model (cf., [9]) describes Hamiltonians of the form $H = \sum_{l=0}^{L-1} \beta_l U_l$, where $\beta_l > 0$ and U_l is unitary acting on $\mathbb{C}^N$ for all l. It assumes access to SELECT and PREPARE oracles defined by

$$\mathrm{SELECT}(H) = \sum_{l=0}^{L-1} |l\rangle\langle l| \otimes U_l, \quad \mathrm{PREPARE}(H) : |0\rangle \to \sum_{l=0}^{L-1} \sqrt{\frac{\beta_l}{\beta}} |l\rangle \ , \tag{45}$$

where $\beta = \sum_{l=0}^{L-1} \beta_l$. Without loss of generality, we can assume $L = 2^l$, so that $|0\rangle$ above refers to the l-qubit state $|0\rangle^{\otimes l}$. A useful property used by some quantum algorithms is

$$\frac{1}{\beta} H = \langle 0|^{\otimes l} \left(\mathrm{PREPARE}(H)\right)^\dagger \mathrm{SELECT}(H) \mathrm{PREPARE}(H) \, |0\rangle^{\otimes l} \ , \tag{46}$$

which gives essentially a block-encoding of H/β. The query complexity in the LCU model is measured by the number of times these oracles are used, together with their inverses and controlled versions.

In contrast, the sparse matrix model gives access to the elements of a d-sparse matrix H [7]. It uses an oracle O_H that allows to perform the transformations

$$|j\rangle\,|k\rangle\,|z\rangle \to |j\rangle\,|k\rangle\,|z \oplus H_{jk}\rangle\,, \quad |j\rangle\,|\ell\rangle \to |j\rangle\,|\nu(j,\ell)\rangle \ . \tag{47}$$

Here, $j, k \in [N] := \{1, \ldots, N\}$ label a row and column of H, and H_{jk} is the corresponding matrix entry (specified in binary form). Also, $\ell \in [d]$ and $\nu : [N] \times [d] \to [N]$ computes the row index of the ℓ^{th} nonzero entry of the j^{th} row in place. Like in the LCU model, having oracle access to the matrix allows one to construct a block-encoding of the Hamiltonian — for example, a block-encoding of $H/(d\|H\|_{\max})$ (cf. Sec. 4.1 of Ref. [38] or Ref. [18]). The query complexity in the sparse matrix model is measured by the number of times O_H is used, together with its inverse and controlled version.

A block-encoding of H/λ, for some $\lambda > 0$, can be constructed using either the LCU or the sparse matrix model. Compared to simply accessing the block-encoding directly, both the LCU and sparse access models provide stronger access to H. Our lower bounds for these models are accordingly stronger than restricting access to the block-encoding model. Since the LCU and sparse matrix access models are incomparable, we must provide lower bounds in both models.

4.2 Lower bound: formal statement

Given a Hamiltonian H, we wish to show that even if the initial state is supported on a subspace associated with energies that are vanishingly small, e.g., $\Delta = o(\|H\|)$ or $\Delta = o(\lambda)$, time evolution of low-energy states requires $\Omega(t\|H\|)$ queries in the worst case. We will construct a Hamiltonian H that can be efficiently block-encoded as H/λ for $\lambda = \Theta(\|H\|)$. In our algorithm (Thm. 1.1), the query complexity can be written in terms of independent and asymptotically large parameters $t\Delta, \lambda/\Delta$, and $1/\epsilon$. Here, we only show a lower bound for the term that depends linearly on t, and we disregard the dependence on $1/\epsilon$; that is, we are assuming constant error or instances where $t\lambda = \Omega(\log(1/\epsilon))$. A lower bound in terms of $1/\epsilon$ in the opposite regime $t\lambda = \mathcal{O}(\log(1/\epsilon))$ for general Hamiltonians is already given in Ref. [10].

In order to vary the parameters $t\Delta$ and λ/Δ independently of each other, we introduce a Grover-like problem defined by two integer parameters, $K \geq 1$ and $N \geq 1$. To show the general no fast-forwarding theorem, we will evolve a Hamiltonian $H_{K,N}$ and examine its low-energy subspace such that $t\Delta = \Theta(K)$ and $\lambda/\Delta = \Theta(\sqrt{N})$. In the following section, to show the optimality of our quantum algorithm for gap-amplifiable Hamiltonians, we will use the same task but instead take $t\Delta = \Theta(K)$ and $\lambda/\Delta = \Theta(N)$.

Before stating the no fast-forwarding theorem, we briefly introduce the Grover-like problem, which we refer to as PARITY $\circ$ $\mathrm{OR}_{K,N}$. Informally, the problem consists of performing K instances of quantum search, each over N items. There is a single marked element in each instance. The output is the parity of a bit string of size K, where the k^{th} bit is determined by the k^{th} instance, and is 0 or 1 whether the marked element is in the first $N/2$ items or not, respectively. As shown in Appendix B via the adversary method, this problem requires $\Omega(K\sqrt{N})$ queries to a Grover-like oracle. We defer the formal definition of this problem to Appendix B as it is not essential for the analysis in this section. However, we note that this problem is related to a quantum state preparation problem, which we now describe; see Lemma B.2 for its proof.

Lemma 4.1 (Quantum state preparation lower bound)**.** *For all $k \in [K]$ define functions $f_k : [N] \to \{0,1\}$ such that $f_k(x_k) = 1$ for some $x_k \in [N]$ and $f_k(x \neq x_k) = 0$ otherwise. Then, at least $\Omega(K\sqrt{N})$ queries to an oracle $O_f : |k\rangle|x\rangle \to (-1)^{f_k(x)}|k\rangle|x\rangle$ acting on $\mathbb{C}^{K\times N}$ are required to prepare a quantum state*

$$|\phi\rangle = \bigotimes_{k=1}^{K} |\phi_k\rangle \tag{48}$$

satisfying $|\langle\phi_k|x_k\rangle|^2 \geq \alpha^2$ for constant $\alpha \in (0,1)$. Here, $|\phi_k\rangle \in \mathbb{C}^N$ and $|\phi\rangle \in \mathbb{C}^{N^K}$.

Here, x_k is the single marked element of the k^{th} instance of Grover search with N items. Accordingly, preparation of $|\phi\rangle$ would allow us to solve $\mathcal{O}(K)$ instances of Grover search with high probability (determined by α). In this sense, the lower bound $\Omega(K\sqrt{N})$ is not surprising, but the technical analysis of Appendix B is required to formally prove it.

The above state preparation problem reduces to a Hamiltonian simulation problem, which we will use to prove lower bounds on time evolution. These apply to the LCU and sparse matrix models discussed in Sec. 4.1. The main result of this section follows.

Theorem 4.2 (No fast-forwarding in the low-energy subspace of generic Hamiltonians, formal)**.** *Let $K \geq 1$ and $N \geq 1$ be integers. Then, there exists a sequence of PSD Hamiltonians $\{H_{K,N}\}_{K,N}$, each acting on a N^K-dimensional Hilbert space $\mathcal{H}_{K,N} := \mathbb{C}^{N^K}$, with the following properties.*

- *In the LCU model, each Hamiltonian can be expressed as* $H_{K,N} = \sum_{l=0}^{L_K-1} \beta_l U_l$, *where* $\beta_l > 0$ *and* U_l *is unitary. Also,* $\lambda_K = \sum_{l=0}^{L_K-1} \beta_l$ *satisfies* $\lambda_K = \Theta(K)$ *and* $L_K = \Theta(K)$.
- *In the sparse matrix model, each Hamiltonian can be expressed a* d_K*-sparse matrix, with* $d_K = \Theta(K)$. *Also,* $\lambda_K = d_K \|H_{K,N}\|_{\max}$ *satisfies* $\lambda_K = \Theta(K)$.

Assume access to the corresponding oracles for each model. Then, for each $H_{K,N}$, *there exists* $\Delta_{K,N} = \mathcal{O}(K/\sqrt{N})$, *a low-energy state* $|\psi_{K,N}\rangle \in \mathcal{S}_{\Delta_{K,N}}$, *and a time* $t_N = \mathcal{O}\left(\sqrt{N}\right)$, *such that preparing a unitary* U *satisfying* $|\langle\psi_{K,N}| \Pi U^\dagger \Pi e^{-itH_{K,N}} |\psi_{K,N}\rangle| \geq 2/3$ *requires at least*

$$\Omega(t_N \lambda_K) \tag{49}$$

queries to either the SELECT *and* PREPARE *oracles in the LCU model, or the* $O_{H_{K,N}}$ *oracle in the sparse matrix model. In addition,* $\Delta_{K,N}/\lambda_K = \mathcal{O}(1/\sqrt{N})$.

The result proves that only having standard methods of access to H might prevent one from using the low-energy condition of the initial state to improve upon methods like QSP. Independently, it gives a no fast-forwarding result for Hamiltonian simulation in the LCU model, complementing the result of Ref. [8] that applies to the sparse matrix model. (In general, Hamiltonians in the LCU model are not sparse.)

Given the result of no fast-forwarding in the LCU model, we can already conclude that access to a block-encoding of a general Hamiltonian does not allow fast-forwarding of time evolution in the low-energy subspace either. Otherwise there would be a contradiction, since Eq. (46) details how to efficiently build the block-encoding in this model.

We give the proof of Thm. 4.2, first in the LCU model and then in the sparse matrix model, below.

4.3 No fast-forwarding in the low-energy subspace for the LCU model

Consider an n-qubit system in $\mathbb{C}^N$, where $N = 2^n$. For some $x \in \{0,1\}^n$, let $|x\rangle$ be a marked quantum state and $|s\rangle := \frac{1}{\sqrt{N}} \sum_{x'=0}^{N-1} |x'\rangle$ be the uniform superposition state over all computational basis states (bitstrings). The Hamiltonian

$$H_x = (1 + 1/\sqrt{N})\mathbb{1}_N - |x\rangle\langle x| - |s\rangle\langle s| \tag{50}$$

solves one instance of quantum search as follows [40]: i) initially prepare $|s\rangle$, ii) let it evolve for time $t_N = \Theta(\sqrt{N})$, and iii) perform a measurement to obtain x with high probability. The Hamiltonian acts non-trivially in a subspace of dimension 2 spanned by $\{|x\rangle, |s\rangle\}$. It has two distinct eigenvalues $\{0, \frac{2}{\sqrt{N}}\}$ in that subspace and the eigenvalue $1 + \frac{1}{\sqrt{N}}$ is of multiplicity $N - 2$. It follows that $H_x \succeq 0$, $\|H_x\| = 1 + \frac{1}{\sqrt{N}} \leq 2$ and the spectral gap is $\Delta = \frac{2}{\sqrt{N}}$.

As described above, we will consider K independent copies of this search problem; choosing arbitrary K and N permits the low-energy and evolution time to be set independently. The total Hamiltonian is then

$$H_{K,N} := \sum_{k=1}^{K} \mathbb{1}_N^{\otimes k-1} \otimes H_{x_k} \otimes \mathbb{1}_N^{\otimes K-k}\,, \tag{51}$$

where $x_k \in \{0,1\}^n$ for all $k \in [K]$, referring to a marked item over a set of N items in the k^{th} instance of quantum search.

We briefly comment on the assumed LCU access for $H_{K,N}$. Observe that each H_{x_k} can be easily expressed as an LCU since H_x can be expressed as

$$H_x = \frac{1}{\sqrt{N}}\mathbb{1}_N + \left(\frac{\mathbb{1}_N - 2\,|x\rangle\langle x|}{2}\right) + \left(\frac{\mathbb{1}_N - 2\,|s\rangle\langle s|}{2}\right) . \tag{52}$$

Each term above is associated with a unitary; for example, the first term in brackets is a reflection over the state $|x\rangle$ which corresponds to the Grover oracle that applies $O_x\,|y\rangle = (-1)^{\delta_{xy}}\,|y\rangle$. Then, if we let $\mathrm{U}_1 = \mathbb{1}_N$, $\mathrm{U}_2(x) = \mathbb{1}_N - 2\,|x\rangle\langle x|$, $\mathrm{U}_3 = \mathbb{1}_N - 2\,|s\rangle\langle s|$, we have $H_x = \frac{1}{\sqrt{N}}\mathrm{U}_1 + \frac{1}{2}\mathrm{U}_2(x) + \frac{1}{2}\mathrm{U}_3$ and

$$H_{K,N} = \sum_{k=1}^{K} \mathbb{1}_N^{\otimes k-1} \otimes \left(\frac{1}{\sqrt{N}}\mathrm{U}_1 + \frac{1}{2}\mathrm{U}_2(x_k) + \frac{1}{2}\mathrm{U}_3\right) \otimes \mathbb{1}_N^{\otimes K-k} , \tag{53}$$

which is itself an LCU. That is, $H_{K,N} = \sum_{l=0}^{L_K-1} \beta_l U_l$, where each U_l uses U_1, $\mathrm{U}_2(x_k)$, or U_3 once, $L_K = 3K$, $\beta_l = \mathcal{O}(1)$ for all l, and $\lambda_K = \sum_{l=0}^{L_K-1} \beta_l = \mathcal{O}(K)$. We then assume access to the SELECT($H_{K,N}$) and PREPARE($H_{K,N}$) oracles of Eq. (45), where the β_l's and U_l's are specified from Eq. (53).

Importantly, the corresponding SELECT($H_{K,N}$) can be implemented with a single call to an oracle that implements $O_f\,|k\rangle\,|x\rangle = (-1)^{x_k}\,|k\rangle\,|x\rangle$. This is the oracle O_f used in Lemma 4.1. To see this, note that $\mathrm{U}_2(x_k) = \sum_x (-1)^{f_k(x)}\,|x\rangle\langle x|$ and then the oracle can be expressed as $O_f = \sum_k |k\rangle\langle k| \otimes \mathrm{U}_2(x_k)$. The SELECT($H_{K,N}$) oracle implements each $\mathrm{U}_2(x_k)$ on a different subsystem conditional on $|k\rangle$. This can be achieved by conjugating one $O_f \otimes \mathbb{1}_N^{K-1}$ with corresponding controlled-SWAP operations, which do not use the oracle.

Each Grover instance can be solved within the low-energy subspace $[0, \Delta_{K,N} = \frac{2K}{\sqrt{N}}]$ of the Hamiltonian $H_{K,N}$. Let

$$|\psi_{K,N}\rangle := |S\rangle = |s\rangle^{\otimes K} \tag{54}$$

be now the equal superposition state in $\mathbb{C}^{N^K}$. Each $|s\rangle$ is confined to the low-energy subspace of each H_{x_k}, and hence $|S\rangle$ is confined to the low-energy subspace of $H_{K,N}$.

$$|\phi_{k,0}\rangle = \sqrt{\frac{1+\sqrt{\frac{1}{N}}}{2}}\,|x_k\rangle + \sqrt{\frac{1-\sqrt{\frac{1}{N}}}{2}}\,|\perp_k\rangle , \tag{55}$$

$$|\phi_{k,1}\rangle = \frac{1}{\sqrt{2\left(1+\frac{1}{\sqrt{N}-1}\right)}}\,|x_k\rangle - \sqrt{\frac{1+\sqrt{\frac{1}{N}}}{2}}\,|\perp_k\rangle , \tag{56}$$

where we introduced quantum state

$$|\perp_k\rangle = \frac{1}{\sqrt{N-1}} \sum_{j\neq x_k} |j\rangle . \tag{57}$$

The low-energy eigenstates of $H_{K,N}$ are then tensor products of the eigenstates above. In particular, we decompose $|S\rangle$ over its eigenbasis as

$$|S\rangle = \bigotimes_{k=1}^{K} \sqrt{\frac{1+\frac{1}{\sqrt{N}}}{2}}\,|\phi_{k,0}\rangle - \sqrt{\frac{1-\frac{1}{\sqrt{N}}}{2}}\,|\phi_{k,1}\rangle . \tag{58}$$

To apply the state preparation lower bound of Lemma 4.1, it suffices to consider the preparation of

$$|\phi\rangle = \bigotimes_{k=1}^{K} |x_k\rangle \ , \tag{59}$$

which corresponds to $\alpha = 1$ in Lemma 4.1. We show that $|\phi\rangle$ can be prepared *exactly* via time evolution with $H_{K,N}$ from the initial state $|S\rangle$. For simplicity, we isolate the analysis to a single subsystem H_{x_k} with a marked element x_k, since the subsystems do not interact:

$$e^{-itH_{x_k}} |s\rangle = \sqrt{\frac{1+\frac{1}{\sqrt{N}}}{2}} |\phi_{k,0}\rangle - e^{-\frac{2it}{\sqrt{N}}} \sqrt{\frac{1-\frac{1}{\sqrt{N}}}{2}} |\phi_{k,1}\rangle \ . \tag{60}$$

The probability of producing the marked element after time t is given by

$$|\langle x_k| e^{-iH_{x_k}t} |s\rangle|^2 = \left| \frac{1+\sqrt{\frac{1}{N}}}{2} - e^{-2it/\sqrt{N}} \frac{1-\sqrt{\frac{1}{N}}}{2} \right|^2 \geq \sin\left(\frac{t}{\sqrt{N}}\right)^2 \ , \tag{61}$$

for $0 \leq t \leq \frac{\pi}{2}\sqrt{N}$. Hence, at time $t_N = \frac{\pi}{2}\sqrt{N}$, we have $e^{-it_N H_{K,N}} |\psi_{K,N}\rangle = e^{-it_N H_{K,N}} |S\rangle = |\psi\rangle$. By Lemma 4.1, this requires $\Omega(K\sqrt{N})$ queries to the oracle, and hence $\Omega(K\sqrt{N})$ queries to SELECT$(H_{K,N})$. Written in terms of λ_K and t_N, this lower bound is $\Omega(t_N \lambda_K)$, which applies even when $\Delta_{K,N}/\lambda_K = \mathcal{O}(1/\sqrt{N})$ is asymptotically small. Thus, we obtain a lower bound of $\Omega(t\lambda)$ at time $t = \Theta(\lambda/\Delta)$ in general, when access to the LCU oracles is given.

4.4 No fast-forwarding in the low-energy subspace for the sparse matrix model

We now consider the sparse matrix access model, which is incomparable to the LCU model. We also consider a Hamiltonian that uses the $(\text{PARITY} \circ \text{OR})_{K,N}$ lower bound, but the Hamiltonian is made sparse following Ref. [17] on spatial quantum walks. Each H_{x_k} is now defined over a d-dimensional periodic lattice with N vertices, where we choose d to be a constant larger than 4. As shown in Ref. [17] (see their Eq. 96), choosing constant $\gamma > 0$ and $c = \mathcal{O}(1/\sqrt{N})$ yields a PSD $N \times N$ Hamiltonian

$$H_{x_k} = \gamma(D - A) - |x_k\rangle\langle x_k| + c \tag{62}$$

with diagonal matrix D, adjacency matrix A, and marked state $x_k \in \{0,1\}^n$. Moreover,

$$|\langle x_k| e^{-iH_{x_k}t} |s\rangle|^2 = \alpha^2 \sin\left(\frac{\alpha t}{\sqrt{N}}\right)^2 \ , \tag{63}$$

for some constant $\alpha > 0$ and $|s\rangle = \frac{1}{\sqrt{N}} \sum_{x'=0}^{N-1} |x'\rangle$. To apply a similar argument as in Sec. 4.3, we need to show that $|s\rangle$ is (almost) in the low-energy subspace of H_{x_k}, and that the relevant energies are $\mathcal{O}(1/\sqrt{N})$ when the Hamiltonian is normalized so that $\|H_{x_k}\| = \mathcal{O}(1)$. Define the integral $I_{j,d}$ by

$$I_{j,d} = (2d)^{-j} \int_0^\infty dx \frac{x^{j-1}e^{-x}}{\Gamma(x)} [\mathcal{I}_0(x/d)]^d \tag{64}$$

for modified Bessel function of the first kind $\mathcal{I}_0$. For a lattice with dimension $d > 4$, the constant α in the success probability is given by $I_{1,d}/\sqrt{I_{2,d}}$; moreover, $3/4 < \alpha < 1$ for all choices of d. It was also shown in Ref. [17] that H_{x_k} with

$$\gamma = I_{1,d}, \quad c = \frac{\alpha}{\sqrt{N}} \ , \tag{65}$$

has ground and first excited state energies

$$E_0 = 0, \quad E_1 = \frac{2\alpha}{\sqrt{N}} , \tag{66}$$

while the spectral norm of H_{x_k} is upper-bounded by a constant. Hence, $H_{x_k} \succeq 0$. The N eigenstates $|\phi_{k,j}\rangle$ of H_{x_k} satisfy, for all $j = 0, \ldots, N-1$,

$$|\langle s\,|\phi_{k,j}\rangle|^2 = \frac{1}{N}\frac{1}{(E_j - c)^2 F'(E_j - c)} , \tag{67}$$

where in the large-N limit,

$$F'(E) = \frac{1}{NE^2} + \frac{I_{2,d}}{\gamma^2} . \tag{68}$$

Evaluating the support of $|s\rangle$ on the two eigenstates of lowest energy E_0 and E_1, we obtain

$$|\langle s_k|\phi_{k,0}\rangle|^2 = |\langle s_k|\phi_{k,1}\rangle|^2 = \frac{1}{N}\left(\frac{1}{N} + \frac{I_{2,d}}{\gamma^2}\frac{I_{1,d}^2}{I_{2,d}N}\right)^{-1} \to \frac{1}{2} \tag{69}$$

to leading order. Consequently, $|s\rangle$ mostly is supported in a subspace associated with eigenvalues at most $2\alpha/\sqrt{N}$ of H_{x_k} and has vanishingly small support outside this subspace. Formally, we can introduce the low-energy state

$$|\psi\rangle := \frac{\Pi_\Delta\,|s\rangle}{\|\Pi_\Delta\,|s\rangle\|} \tag{70}$$

for projector Π_Δ onto the low-energy subspace of the Hamiltonian defined by $\Delta = 2\alpha/\sqrt{N}$. Eq. (69) implies that $|\psi\rangle$ satisfies $|\,\langle\psi|s\rangle\,| \geq 1 - \delta$ for any (arbitrarily small) $\delta > 0$ that is constant with respect to N. Consequently, a lower bound on the query complexity for preparing a state of the form $e^{-itH_{x_k}}\,|s\rangle$ with constant fidelity also implies a lower bound on the query complexity for preparing the state $e^{-itH_{x_k}}\,|\psi\rangle$ with constant fidelity. We continue our analysis in terms of $|s\rangle$ and apply it to the low-energy state $|\psi\rangle$ at the end.

These properties of H_{x_k} and $|\psi\rangle$ are similar to the case studied in Sec. 4.3. The rest of the proof is similar to that of the LCU model. Time evolution of the initial state $|S\rangle = |s\rangle^{\otimes K}$ for some time $t_N = \Theta(\sqrt{N})$ under the $\mathcal{O}(Kd)$-sparse Hamiltonian

$$H_{K,N} = \sum_{k=1}^{K} \mathbb{1}_N^{\otimes k-1} \otimes H_{x_k} \otimes \mathbb{1}_N^{\otimes K-k} , \tag{71}$$

will produce a state $|\phi\rangle = \bigotimes_{k=1}^{K} |\phi_k\rangle$ that satisfies

$$|\,\langle\phi_k|x_k\rangle\,|^2 = \alpha^2 , \tag{72}$$

for the constant α in Eq. (63). Since $d = \mathcal{O}(1)$, note that the sparsity of $H_{K,N}$ is $d_K = \mathcal{O}(K)$ and also $\|H_{K,N}\|_{\max} = \mathcal{O}(1)$, since $\|H_{x_k}\|_{\max} = \mathcal{O}(1)$ for Eq. (62). Then, $\lambda_K = d_K\|H_{K,N}\|_{\max} = \mathcal{O}(K)$. Furthermore, the oracle $O_{H_{K,N}}$ that provides access to the nonzero matrix elements of $H_{K,N}$ and their positions can be constructed with a single call to an oracle that implements $O_f\,|k\rangle\,|x\rangle = (-1)^{x_k}\,|k\rangle\,|x\rangle$. This is similar to the oracle in Lemma 4.1, which now implies a lower bound $\Omega(K\sqrt{N})$ for the number of uses of $O_{H_{K,N}}$. To see this, note that each entry of $H_{K,N}$ can be labeled by (k, x), and the positions of the nonzero entries are readily obtained from those of H_{x_k}, which are known due to Eq. (62). In

terms of λ_K and t_N, the lower bound is $\Omega(t_N \lambda_K)$. Since the state $|s\rangle$ has vanishingly small support outside the low-energy subspace $\mathcal{S}_\Delta$, the state $|S\rangle$ similarly has vanishingly small support outside the low-energy subspace $\mathcal{S}_{\Delta_{K,N}}$, for $\Delta_{K,N} = \mathcal{O}(K/\sqrt{N})$. The lower bound also holds for the state $|\psi_{K,N}\rangle := \Pi_{\Delta_{K,N}} |S\rangle / \|\Pi_{\Delta_{K,N}} |S\rangle\|$, which is entirely contained within the low-energy subspace, and obtained by projecting each $|s\rangle$ into their low-energy subspaces.

Hence, we proved a lower bound $\Omega(t\lambda)$ at time $t = \Theta(\lambda/\Delta)$ in general, when access oracle access to a sparse matrix H is given. Note that our result takes constant sparsity d; a more careful analysis would be needed to obtain a lower bound in terms of the sparsity of H.

5 Query lower bounds for simulating time evolution of low-energy states of gap-amplifiable Hamiltonians

We show lower bounds on the query complexity of simulating low-energy states of gap-amplifiable Hamiltonians $H = \lambda A^\dagger A$, given some form of oracle access to A. As shown by Thm. 1.1 and Sec. 4, this proves to be more powerful than having access to H alone.

Lemma 1.3 and Lemma 5.1 below show an equivalence between having access to a gap-amplifiable Hamiltonian $H = \lambda A^\dagger A$ and having access to a block-encoding of H' with low-energy subspace corresponding to eigenvalues in $[-1, -1 + \Delta/\lambda]$. For this reason, we will perform the analysis for gap-amplifiable Hamiltonians here. For a state $|\psi\rangle$ with energy at most Δ of such H, we summarize all possible asymptotic regimes and results:

1. *Time-dominated regime*: $\log(1/\epsilon) = o(t\Delta)$, producing an upper bound of $\mathcal{O}(t\sqrt{\lambda\Delta})$. In this section, we will show matching query lower bounds in two different oracle models: the LCU model and the sparse matrix model. In the following section, we will show matching gate complexity lower bounds for this case.

2. *Intermediate regime*: $\log(1/\epsilon) = o(t\lambda)$ and $t\Delta = o(\log(1/\epsilon))$, producing an upper bound of $\mathcal{O}(\sqrt{t\lambda \log 1/\epsilon})$. In this section, we will show a query lower bound in the LCU model of $\Omega(\sqrt{t\lambda})$.

3. *Error-dominated regime*: $t\lambda = o(\log(1/\epsilon))$, producing an upper bound of $\mathcal{O}(\log(1/\epsilon))$. A matching lower bound (up to log log factors) is known [10], based on parity.

Hence, we will obtain query lower bounds that show our algorithm is optimal with respect to t, Δ, and λ; we do not prove optimality with respect to $1/\epsilon$ in the intermediate scaling regime but note that the algorithm has a mild sublogarithmic dependence on error.

5.1 Access models

Similar to Sec. 4.1, we assume that access to $M \times N$ matrix A is given through the LCU or sparse matrix model. Since A is not necessarily a square matrix, some small modifications are needed.

The matrix A can be expressed as $A = (\sum_{l=0}^{L-1} \beta_l U_l)\Pi$, where each $\beta_l > 0$ and each U_l is unitary acting on $\mathbb{C}^M$ for all l, and Π is the projector onto $\mathbb{C}^N$. The LCU model then assumes access to a SELECT and a PREPARE oracle defined by

$$\text{SELECT}(A) = \sum_{l=0}^{L-1} |l\rangle\langle l| \otimes U_l, \quad \text{PREPARE}(A) : |0\rangle \to \sum_{l=0}^{L-1} \sqrt{\frac{\beta_l}{\beta}} |l\rangle \ , \tag{73}$$

where $\beta = \sum_{l=0}^{L-1} \beta_l$. As before, we can assume $L = 2^l$, so that $|0\rangle$ above refers to the l-qubit state $|0\rangle^{\otimes l}$.

The sparse matrix model uses an oracle O_A that allows to perform the transformations

$$|j\rangle\,|k\rangle\,|z\rangle \to |j\rangle\,|k\rangle\,|z \oplus A_{jk}\rangle\,, \quad |j\rangle\,|\ell\rangle \to |j\rangle\,|\nu(j,\ell)\rangle\ . \tag{74}$$

Here, $j \in [M]$ and $k \in [N]$ label a row and column of d-sparse A, and A_{jk} is the corresponding matrix entry (specified in binary form). Also, $\ell \in [d]$ and $\nu : [M] \times [d] \to [N]$ computes the row index of the ℓ^{th} nonzero entry of the j^{th} row in place.

A block-encoding V of A can be obtained with one use of SELECT(A), one use of PREPARE(A), and one use of PREPARE(A)† following a similar procedure to the one given in Sec. 4.1. It can also be constructed using O_A and other two-qubit gates (cf. Lemma 6 of Ref. [38] or Ref. [18]). In both models, the block-encoding is a unitary acting on an enlarged space that contains A/α in one of its blocks. The factor $\alpha > 0$ can be different for the LCU and sparse models. In the results below, α is defined appropriately for either model.

In Lemma 1.3, we showed that a constant number of queries to a block-encoding T of a Hamiltonian $H' = (H - E)/\lambda$ with the low-energy subspace contained in $[-1, -1 + \Delta/\lambda]$ provides access to a block-encoding V of A such that $H - F = 2\lambda A^\dagger A$, where $E, F \in \mathbb{R}$ and $\lambda > 0$. Since our algorithm uses the block-encoding of A, Lemma 1.3 implies that our algorithm can be applied when access to a Hamiltonian is given through the block-encoding T. Here, we will show our lower bounds for gap-amplifiable Hamiltonians. For these lower bounds to apply when given access to T instead of V, we need to prove the opposite direction compared to Lemma 1.3: we must show that a constant number of queries to V suffice to prepare T. This result follows from standard techniques.

Lemma 5.1 (Efficient block-encoding of H')**.** *Let V be the unitary that block-encodes an $M \times N$ matrix A as in Eq. (4) and let the Hamiltonian acting on $\mathcal{H}$ be $H = \lambda A^\dagger A$, for some $\lambda > 0$. Let $|\psi\rangle \in \mathcal{S}_\Delta$ be any state supported exclusively on the subspace associated with eigenvalues of H at most Δ, for $\lambda \geq \Delta > 0$. Then, a block-encoding T of $H' = (H - \lambda)/\lambda$ can be implemented with a constant number of queries to controlled-V and controlled-$V^\dagger$. In addition, $|\psi\rangle$ is supported in the subspace associated with eigenvalues of H' in $[-1, -1 + \Delta/\lambda]$.*

Proof. The statement on the support of $|\psi\rangle$ immediately follows from the definition of H'. We can alternatively write

$$H' = A^\dagger A - \mathbb{1}_N = \Pi V^\dagger \Pi' V \Pi\,. \tag{75}$$

Let $M = 2^m$ and $N = 2^n$. Above, the projector Π' can be replaced by

$$P' = |0\rangle\langle 0|^{\otimes m-n} \otimes \mathbb{1}_N\ ; \tag{76}$$

see Sec. 2.1. The projector P' can be expressed as $(U_{P'} + \mathbb{1}_M)/2$, where $U_{P'} := 2P' - \mathbb{1}_M$ is unitary. Then,

$$H' = \Pi V^\dagger \left(\frac{U_{P'} + \mathbb{1}_M}{2}\right) V \Pi\,. \tag{77}$$

Standard techniques give a block encoding of $V^\dagger \left(\frac{U_{P'}+\mathbb{1}_M}{2}\right) V$, which is a sum of two unitaries. For example, if

$$\tilde{V} := |+\rangle\langle +| \otimes U_{P'} + |-\rangle\langle -| \otimes \mathbb{1}_M\,, \tag{78}$$

where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$, we obtain

$$H' = \Pi V^\dagger \tilde{V} V \Pi \,, \tag{79}$$

and hence

$$T = V^\dagger \tilde{V} V \,. \tag{80}$$

To construct it, we needed V and $V^\dagger$ once, and $\mathcal{O}(m)$ two-qubit gates to implement $U_{P'}$. □

We can now show lower bounds for simulating general gap-amplifiable Hamiltonians and conclude that the same lower bounds hold when given access to a block-encoding of H' instead.

5.2 Lower bounds: formal statements

We first address the lower bounds in the time-dominated scaling regime. Our argument is once again based on the PARITY $\circ$ $\mathrm{OR}_{K,N}$ lower bound of K partitions of independent Grover searches presented in Sec. 4. We will construct a Hamiltonian such that the asymptotic parameters $t\Delta$ and λ/Δ are independently controlled as $t\Delta = \mathcal{O}(K)$ and $\lambda/\Delta = \mathcal{O}(N)$, allowing t and Δ to be arbitrarily set with respect to λ. We present our lower bounds formally in terms of these parameters.

Theorem 5.2 ($t\sqrt{\lambda\Delta}$ lower bound on the query complexity of simulating low-energy states of gap-amplifiable Hamiltonians)**.** *Let $K \geq 1$ and $N \geq 1$ be integers. Then, there exists a sequence of gap-amplifiable Hamiltonians $\{H_{K,N} = \lambda_K A_{K,N}^\dagger A_{K,N}\}_{K,N}$, each acting on an N^K-dimensional Hilbert space $\mathcal{H}_{K,N} := \mathbb{C}^{N^K}$, $\|A_{K,N}\| \leq 1$, with the following properties.*

- *In the LCU model, each term can be expressed as $A_{K,N} = (\sum_{l=0}^{L_K-1} \beta_l U_l)\Pi$, where $\beta_l > 0$, U_l is unitary, and Π is a projector onto $\mathcal{H}_{K,N}$. Also, $\sum_{l=0}^{L_K-1} \beta_l \leq 1$, $L_K = \Theta(K)$, and $\lambda_K = \Theta(K)$.*
- *In the sparse matrix model, each term $A_{K,N}$ and $H_{K,N}$ can be expressed as a d_K-sparse matrix, with $d_K = \Theta(K)$. Also, $\lambda_K = d_K \|H_{K,N}\|_{\max}$ satisfies $\lambda_K = \Theta(K)$.*

Assume access to the corresponding oracles for each model. Then, for each $H_{K,N}$, there exists $\Delta_{K,N} = \mathcal{O}(K/N)$, a low-energy state $|\psi_{K,N}\rangle \in \mathcal{S}_{\Delta_{K,N}}$, and a time $t_N = \mathcal{O}(N)$, such that preparing a unitary U satisfying $|\langle\psi_{K,N}| \Pi U^\dagger \Pi e^{-it_N H_{K,N}} |\psi_{K,N}\rangle| \geq 2/3$ requires at least

$$\Omega\left(t_N \sqrt{\lambda_K \Delta_{K,N}}\right) \tag{81}$$

queries to each of either the $\mathrm{SELECT}(A_{K,N})$ *and* $\mathrm{PREPARE}(A_{K,N})$ *oracles in the LCU model, or the $O_{A_{K,N}}$ oracles in the sparse matrix model. In addition, $\Delta_{K,N}/\lambda_K = \mathcal{O}(1/N)$.*

In Sec. 1.1, we report our lower bound in terms of access to a block-encoding of a Hamiltonian H' with low-energy subspace $[-1, -1 + \Delta/\lambda]$. That result is implied by Theorem 5.2: the LCU model can provide stronger access to A or A_l compared to only providing access to their block-encodings, since these can be efficiently constructed from the LCU oracles. In addition, providing access to the block-encoding of A is equivalent to providing access to the block-encoding of H' (due to Lemmas 1.3 and 5.1).

To show the lower bound in the intermediate regime where $\log(1/\epsilon) = o(t\lambda)$ and $t\Delta = o(\log(1/\epsilon))$, we instead use a version of the polynomial method applicable to trigonometric polynomials. This allows us to show that time-dependence in the LCU model cannot be improved from $\sqrt{t\lambda}$.

Theorem 5.3 ($\sqrt{t\lambda}$ lower bound on the query complexity of simulating low-energy states of gap-amplifiable Hamiltonians)**.** *Let $t > 0$, $\Delta > 0$, and $\epsilon > 0$ be such that $t\Delta = o(\log(1/\epsilon))$ and $\log(1/\epsilon) = o(t\lambda)$, where $\lambda = 1$. Let $\theta \in [0, \theta_M]$, where $\theta_M := \arcsin\sqrt{\Delta}$. Then, there exists a continuous family of gap-amplifiable Hamiltonians $\{H_\theta = A_\theta^\dagger A_\theta\}_\theta$, each acting on a 2-dimensional Hilbert space $\mathbb{C}^2$, $A_\theta = (A_\theta)^\dagger$, $\|A_\theta\| \leq 1$, with the following properties. In the LCU model, $A_\theta = (U_\theta + (U_\theta)^\dagger)/2$ is a linear combination of two unitaries U_θ and $U_\theta^\dagger$. Assume access to these unitaries and their controlled versions. Then, for each H_θ, there exists a low-energy state $|\psi_\theta\rangle \in \mathcal{S}_\Delta$, such that preparing a unitary U satisfying $|\langle\psi| \Pi U^\dagger \Pi e^{-itH_\theta} |\psi\rangle| \geq 1 - \epsilon$ requires at least*

$$\Omega\left(\sqrt{t\lambda}\right) \tag{82}$$

uses of the unitaries. Similarly, if access to A_θ is given through its block-encoding, preparing U requires $\Omega(\sqrt{t\lambda})$ queries to this block-encoding.

5.3 $t\sqrt{\lambda\Delta}$ lower bound for the LCU model

We provide the proof of Thm. 5.2, which concerns the time-dominated regime, for the LCU model. The proof in this case is similar to that in Sec. 4.3, with the modification that we use the square of the Hamiltonians H_x to obtain a gap-amplifiable Hamiltonian. This modifies that analysis slightly, as we discuss below.

We consider the gap-amplifiable Hamiltonian acting on $\mathcal{H}_{K,N} = \mathbb{C}^{N^K}$

$$H_{K,N} = \sum_{k=1}^{K} \mathbb{1}_N^{\otimes k-1} \otimes (H_{x_k})^2 \otimes \mathbb{1}_N^{\otimes K-k} \,, \tag{83}$$

where $H_x = (1 + 1/\sqrt{N})\mathbb{1}_N - |x\rangle\langle x| - |s\rangle\langle s|$, $|s\rangle = \frac{1}{\sqrt{N}} \sum_{x'=0}^{N-1} |x'\rangle$. Each of the K bit strings $x_k \in \{0,1\}^n$ is associated with an instance of Grover search over N items. Since $\|H_{x_k}\| = \mathcal{O}(1)$, the $H_{K,N}$ above is a (λ_K, K)-gap-amplifiable Hamiltonian with $\lambda_K = \mathcal{O}(K)$. That is, using Lemma 3.2, the Hamiltonian can also be expressed as $H_{K,N} = \lambda_K A_{K,N}^\dagger A_{K,N}$. The operators $A_{K,N}$ of dimension $(KN) \times N$ can be constructed from the individual operators $A_k^\dagger = A_k = H_{x_k}/(1 + 1/\sqrt{N})$ of dimension $N \times N$:

$$A_{K,N} := \frac{1}{\sqrt{K}} \sum_{k=1}^{K} |k\rangle \otimes (\mathbb{1}_N^{\otimes k-1} \otimes A_k \otimes \mathbb{1}_N^{\otimes K-k}) \,. \tag{84}$$

This can also be expressed as

$$A_{K,N} = \tilde{A}_{K,N} \Pi \,, \tag{85}$$

where $\tilde{A}_{K,N} := \sum_{k=1}^{K} |k\rangle\langle k| \otimes (\mathbb{1}_N^{\otimes k-1} \otimes A_k \otimes \mathbb{1}_N^{\otimes K-k})$ is of dimension $(KN) \times (KN)$ and Π is the projector onto $\mathcal{H}_{K,N}$; explicitly, we can project onto $\frac{1}{\sqrt{K}} \sum_{k=1}^{K} |k\rangle$. Note that $\tilde{A}_{K,N}$ can be expressed as an LCU with $L_K = \Theta(K)$ unitaries that use U_1, $\mathrm{U}_2(x_k)$, and U_3 in Sec. 4.3.

Importantly, the corresponding SELECT$(A_{K,N})$ can be implemented with a single call to an oracle that implements $O_f |k\rangle |x\rangle = (-1)^{x_k} |k\rangle |x\rangle$. This is the oracle O_f of Lemma 4.1. SELECT$(A_{K,N})$ requires applying $\mathrm{U}_2(x_k)$ on a different subsystem conditioned in $|k\rangle$. However, like in Sec. 4.3, this can be achieved by using O_f and conjugating it with corresponding controlled-SWAP gates.

The low-energy state $|\psi_{K,N}\rangle := |S\rangle = |s\rangle^{\otimes K}$ is now exclusively supported on the subspace of eigenvalues $[0, \Delta_{K,N} = \frac{4K}{N}]$ of $H_{K,N}$. The eigenstates of $H_{K,N}$ are the same as those discussed in Sec. 4.3, and to analyze the time evolution of $|S\rangle$ we consider a single system to obtain

$$|\langle x_k| e^{-it(H_{x_k})^2} |s\rangle|^2 = \left| \frac{1+\sqrt{\frac{1}{N}}}{2} - e^{-4it/N} \frac{1-\sqrt{\frac{1}{N}}}{2} \right|^2 \geq \sin\left(\frac{2t}{N}\right)^2 . \tag{86}$$

At time $t_N = \frac{\pi}{4}N$, each $|s\rangle$ is evolved exactly to $|x_k\rangle$. That is, if $|\phi\rangle = \bigotimes_{k=1}^{K} |x_k\rangle$, then $e^{-it_N H_{K,N}} |S\rangle = |\phi\rangle$.

By Lemma 4.1, simulating this time evolution requires $\Omega(K\sqrt{N})$ queries to the oracle O_f, and hence $\Omega(K\sqrt{N})$ queries to SELECT$(A_{K,N})$. Written in terms of λ_K, $\Delta_{K,N}$, and t_N, this lower bound is $\Omega(t_N\sqrt{\lambda_K \Delta_{K,N}})$, which applies even when $\Delta_{K,N}/\lambda_K = \mathcal{O}(1/N)$ is asymptotically small. Then, we proved a lower bound of $\Omega(t\sqrt{\lambda\Delta})$ at time $t = \Theta(\lambda/\Delta)$ for gap-amplifiable Hamiltonians, when access to the LCU oracles for $A_{K,N}$ is given.

5.4 $t\sqrt{\lambda\Delta}$ lower bound for the sparse matrix model

We provide the proof of Thm. 5.2, which concerns the time-dominated regime, for the sparse matrix model. We introduce a gap-amplifiable Hamiltonian $H_{K,N}$ acting on $\mathcal{H}_{K,N} = \mathbb{C}^{N^K}$ built from an expander graph similar to a construction in Ref. [49]. Once again, time evolution under this Hamiltonian solves $(\text{PARITY} \circ \text{OR})_{K,N}$. The Hamiltonian is also a sum of gap-amplifiable Hamiltonians acting on K different subsystems of dimension N. Each of these Hamiltonians can be expressed as $H_{x_k} = \sum_{l=1}^{L} A_{l,x_k}^\dagger A_{l,x_k} = \lambda A_k^\dagger A_k$ for constant L, $k \in [K]$, and $\lambda = \mathcal{O}(1)$, and $\|A_{l,x_k}\|_{\max} \leq 1$. Explicitly, the gap-amplifiable Hamiltonian under consideration is

$$H_{K,N} = \sum_{k=1}^{K} \mathbb{1}_N^{\otimes k-1} \otimes \left(\sum_{l=1}^{L} A_{l,x_k}^\dagger A_{l,x_k} \right) \otimes \mathbb{1}_N^{\otimes K-k} = \lambda_K A_{K,N}^\dagger A_{K,N} . \tag{87}$$

In particular, in this construction, if each H_{x_k} has constant sparsity then $H_{K,N}$ is $d_K = \mathcal{O}(K)$-sparse. Moreover, if $\lambda = \mathcal{O}(1)$, then $\lambda_K = \mathcal{O}(K)$.

To show that $H_{K,N}$ produces the required state of Lemma 4.1, we can largely rely on our proof for the LCU model. It suffices to analyze the individual gap-amplifiable Hamiltonians H_x and show that the state $|s\rangle$ is mostly confined to the low-energy subspace corresponding to $\Delta = \mathcal{O}(1/N)$. We also require that evolving this low-energy state for time $t_N = \mathcal{O}(N)$ solves an instance of quantum search.

We introduce a suitable gap-amplifiable Hamiltonian with these properties. The marked element is x. For organizational purposes, we divide the bulk of the proof into three pieces: we introduce a Hamiltonian H_x and show it is gap-amplifiable; we show that $|s\rangle$ is largely confined in the low-energy in H_x and replace it with a similar state $|\psi\rangle$ that is fully confined in the low-energy subspace; finally, we show that time evolution $e^{-it_N H_x} |\psi\rangle$ produces the marked state with constant probability in time $t_N = \mathcal{O}(N)$.

Lemma 5.4. *Let $x \in \{0,1\}^n$ and $O_x : |y\rangle \to (-1)^{\delta_{x,y}} |y\rangle$ be Grover's oracle. Then, there exists a d-sparse (λ, L)-gap-amplifiable Hamiltonian $H_x = \sum_{l=0}^{L-1} A_{l,x}^\dagger A_{l,x}$ acting on $\mathbb{C}^N$, such that $d = \mathcal{O}(1)$, $\lambda = \mathcal{O}(1)$, and $L = d + 1 = \mathcal{O}(1)$. The operators $A_{l,x}$ are sparse and of dimension $N \times N$, and satisfy $A_{l,x} = A_{l,x}^\dagger = \Pi_{l,x}$, for some projectors $\Pi_{l,x}$. Moreover, access to the nonzero entries of $A_{l,x}$ and their positions can be performed with one use of O_x. Finally, H_x has lowest eigenvalue 0 and the first nonzero eigenvalue is at least $1/(4N)$.*

The proof is in Appendix A of Ref. [49] and also contained in Appendix C. To provide a similar proof as we did for the Hamiltonians in Sec. 5.3, we would like $|s\rangle$ to reside within a low-energy subspace of H_x associated with $\Delta = \mathcal{O}(1/N)$. Unfortunately, $|s\rangle$ includes vanishingly small support outside this low-energy subspace. We accordingly introduce a state $|\psi\rangle$ that has high overlap with $|s\rangle$ and is fully confined in $\mathcal{S}_\Delta$.

Lemma 5.5. *Let Π_Δ be the projector onto the low-energy subspace of the Hamiltonian H_x defined in Lemma 5.4, and let $\Delta = \mathcal{O}(1/N)$. The state $|\psi\rangle$ defined by*

$$|\psi\rangle = \frac{\Pi_\Delta |s\rangle}{\|\Pi_\Delta |s\rangle\|} \tag{88}$$

satisfies $|\langle\psi|s\rangle| \geq 1 - \delta$ for any (arbitrarily small) constant $\delta > 0$. Moreover, for eigenstates $|\phi_j\rangle$ of H_x, for $j = 0, \ldots, N-1$, with ground state $|\phi_0\rangle$, the following holds:

$$|\psi\rangle = c_0 |\phi_0\rangle + \sum_{j\geq 1} c_j |\phi_j\rangle , \quad c_0 > \frac{1}{2}, \quad |\phi_0\rangle = \frac{1}{\sqrt{2}} |x\rangle + \frac{1}{\sqrt{2(N-1)}} \sum_{y\neq x} |y\rangle . \tag{89}$$

The proof of this lemma is also contained in Appendix C.

We seek to show that time evolution of $|\psi\rangle$ under H_x produces the marked state $|x\rangle$ with constant probability if the evolution time is $t_N = \mathcal{O}(N)$. As shown in Eq. (89), the ground state $|\phi_0\rangle$ has constant overlap with both the marked state $|x\rangle$ and the state $|\psi\rangle$. Hence, we can obtain the marked state with high probability by projecting from $|\psi\rangle$ onto $|\phi_0\rangle$. We implement this measurement with random time evolution, as given by the following known result.

Lemma 5.6 (Ground state projection by random time evolution, from Ref. [11])**.** *Let H be a Hamiltonian with eigenstates $|\phi_0\rangle, |\phi_1\rangle, \ldots$ of eigenvalues $\lambda_0 \leq \lambda_1 \leq \ldots$. The time evolution channel $\mathcal{N}$ defined by drawing evolution times T randomly from a distribution with characteristic function Φ satisfies*

$$\|(\mathcal{N} - \mathcal{M})(|s\rangle\langle s|)\|_1 \leq 2\sqrt{\sum_{j\geq 1} |\Phi(\lambda_j) c_0 c_j^*|^2} \leq \sup_j |\Phi(\lambda_j)| , \tag{90}$$

where $\|\cdot\|_1$ indicates trace norm, and

$$\mathcal{M}(\rho) = P_0 \rho P_0 + \mathcal{E}((1 - P_0)\rho(1 - P_0)) \tag{91}$$

is the channel that performs a perfect measurement of the ground state. Here, $P_0 = |\phi_0\rangle\langle\phi_0|$ is the projector onto the ground state and $\mathcal{E}$ is some quantum channel.

Since for $j \geq 1$ we have $\lambda_j > 1/(4N)$ for H_x, choosing the distribution to be the uniform interval $[0, \alpha N]$ for constant $\alpha > 0$ gives the characteristic function

$$\epsilon := \sup_j |\Phi(\lambda_j)| = \max_{1/(4N)\leq\lambda\leq d+1} \left| \frac{e^{i\lambda\alpha N} - 1}{i\lambda\alpha N} \right| \leq \frac{8}{\alpha} . \tag{92}$$

The corresponding channel $\mathcal{M}$ outputs a state

$$\sigma = \mathcal{M}(|\psi\rangle\langle\psi|) = c_0^2 |\phi_0\rangle\langle\phi_0| + (1 - c_0^2)\rho_\perp , \tag{93}$$

where $\langle\psi| \rho_\perp |\psi\rangle = 0$. If the state $\tilde{\sigma} = \mathcal{N}(|\psi\rangle\langle\psi|)$ is ϵ-close to σ, then

$$|\langle\phi_0| (\tilde{\sigma} - \sigma) |\phi_0\rangle| \leq \|\tilde{\sigma} - \sigma\| \leq \|\tilde{\sigma} - \sigma\|_1 \leq \epsilon . \tag{94}$$

Hence, by the triangle inequality

$$|\langle\phi_0|\tilde{\sigma}|\phi_0\rangle| \geq c_0^2 - \epsilon \geq \frac{1}{4} - \epsilon\,. \tag{95}$$

Due to the definition of the ground state, evolving for random times of at most αN will yield the marked string x with probability at least

$$|\langle x|\mathcal{N}(|\psi\rangle\langle\psi|)|x\rangle|^2 \geq \frac{1}{2}\left(\frac{1}{4} - \frac{8}{\alpha}\right)\,. \tag{96}$$

By choosing, for example $\alpha = 64$, we conclude that $t = \mathcal{O}(N) = \mathcal{O}(1/\Delta)$ is sufficient to solve quantum search with constant probability using the Hamiltonian H_x.

The full Hamiltonian over K subsystems that solves $(\mathrm{PARITY} \circ \mathrm{OR})_{K,N}$ is the one in Eq. (87). Each H_{x_k} is given in Lemma 5.4, that is,

$$H_{x_k} = \sum_{l=0}^{d} A_{l,x_k}^\dagger A_{l,x_k} = \lambda A_k^\dagger A_k\,. \tag{97}$$

The sparsity of H_{x_k} is $d = \mathcal{O}(1)$ and

$$\lambda = d+1, \quad A_k := \frac{1}{\sqrt{d+1}}\left(\sum_{l=0}^{d}|l\rangle\otimes A_{l,x_k}\right)\,. \tag{98}$$

The operators $A_{K,N}$ are defined via

$$\lambda_K = K(d+1), \quad A_{K,N} = \frac{1}{\sqrt{K}}\sum_{k=1}^{K}|k\rangle\otimes A_k\,, \tag{99}$$

which allow us to express the Hamiltonian as $H_{K,N} = \lambda_K A_{K,N}^\dagger A_{K,N}$. The operators $A_{K,N}$ are $\mathcal{O}(K)$-sparse. Each entry can be labeled by (k,x), and constructing the matrix oracle that provides access to these entries and their positions can be done with one use of the oracle of Lemma 4.1 that implements $O_f|k\rangle|x\rangle = (-1)^{x_k}|k\rangle|x\rangle$. The initial state is $|\psi_{K,N}\rangle := \Pi_{\Delta_{K,N}}|S\rangle / \|\Pi_{\Delta_{K,N}}|S\rangle\|$ for $\Delta_{K,N} = \mathcal{O}(K/N)$ by Lemma 5.5.

Hence, using the query complexity lower bound $\Omega(K\sqrt{N})$ for this problem in Lemma 4.1, it implies a lower bound $\Omega(t_N\sqrt{\lambda_K\Delta_{K,N}})$ on the number of queries to the matrix oracle for $A_{K,N}$. As in the LCU case, this applies even when $\Delta_{K,N}/\lambda_K = \mathcal{O}(1/N)$ is asymptotically small.

5.5 $\sqrt{t\lambda}$ lower bound for the LCU model

We now address the intermediate regime of parameters satisfying $\log(1/\epsilon) = o(t\lambda)$ and $t\Delta = o(\log(1/\epsilon))$, and provide the proof to Thm. 5.3. In this regime, the optimization over $\Gamma \in [\Delta, \lambda]$ in Lemma. 1.2 produces $\Gamma = \log(1/\epsilon)/t$, and the query complexity of our algorithm is

$$\mathcal{O}\left(\sqrt{t\lambda\log(1/\epsilon)}\right)\,. \tag{100}$$

We will show a lower bound of $\Omega(\sqrt{t\lambda})$ applicable to the LCU model.

Specifically, to produce problem instances in the intermediate regime, we choose parameters such that $\lambda = 1$, $1 > \Delta > 0$, and $1 \geq \sqrt{\epsilon} \geq t\Delta \geq 8\epsilon > 0$. For example,

we can let $\Delta = \epsilon$ and $t = 1/\sqrt{\epsilon}$, and satisfy $\log(1/\epsilon) = o(1/\sqrt{\epsilon})$ and $\sqrt{\epsilon} = o(\log 1/\epsilon)$, which are the desired conditions since $t\lambda = 1/\sqrt{\epsilon}$ and $t\Delta = \sqrt{\epsilon}$ in this example. Simple modifications of our proofs below would allow one to consider other classes of instances in this regime. We will show that simulating certain instances in our setting is related to realizing a trigonometric polynomial approximation to $t\sin^2\theta$, for sufficiently large $t > 0$ and sufficiently small $|\theta|$. The proof of Thm. 5.3 uses the following result on the degree for such a polynomial.

Lemma 5.7 (Lower bound on the degree of trigonometric-polynomial approximations)**.** *Let $t > 0$, $\epsilon > 0$, and $\Delta < 1$ satisfy $1 \geq \sqrt{\epsilon} \geq t\Delta \geq 8\epsilon$, and define $\theta_M := \arcsin\left(\sqrt{\Delta}\right)$. Then, the lowest degree K of a complex trigonometric polynomial $P_K(\theta) := \sum_{k=-K}^{K} a_k e^{ik\theta}$, $a_k \in \mathbb{C}$, satisfying*

$$|P_K(\theta)| \leq 1,\ \forall\, \theta \in (-\pi, \pi]\ , \tag{101}$$

$$|e^{-it\sin^2\theta} - P_K(\theta)| \leq \epsilon,\ \forall\, \theta \in [-\theta_M, \theta_M]\ , \tag{102}$$

is $K \geq \frac{1}{\pi}\sqrt{2t}$.

Proof. In the Taylor series expansion of the exponential

$$e^{-it\sin^2\theta} = 1 - it\sin^2\theta + r(\theta), \tag{103}$$

the remainder satisfies, for $|\theta| \leq \theta_M$,

$$|r(\theta)| \leq \sum_{l=2}^{\infty} \frac{(t\sin^2\theta)^l}{l!} \leq (t\Delta)^2 \leq \epsilon\ , \tag{104}$$

where the last inequality follows from the assumption $\sqrt{\epsilon} \geq t\Delta$. By the triangle inequality, the condition $|e^{-it\sin^2\theta} - P_K(\theta)| \leq \epsilon$ implies that

$$|1 - it\sin^2\theta - P_K(\theta)| \leq 2\epsilon,\ \forall\, \theta \in [-\theta_M, \theta_M]. \tag{105}$$

Consider instead the trigonometric polynomial $Q_K(\theta) := i(1 - P_K(\theta))$. Applying the triangle inequality, this polynomial is required to satisfy

$$|Q_K(\theta)| \leq 2,\ \forall\, \theta \in (-\pi, \pi]\ , \tag{106}$$

$$|t\sin^2\theta - Q_K(\theta)| \leq 2\epsilon,\ \forall\, \theta \in [-\theta_M, \theta_M]\ . \tag{107}$$

We will show a lower bound on K for $Q_K(\theta)$, which implies the same lower bound on the degree of $P_K(\theta)$.

Let $f'(\theta) := \frac{\mathrm{d}f(\theta)}{\mathrm{d}\theta}$. Bernstein's inequality implies (cf. [12])

$$\sup_{\theta\in(-\pi,\pi]} |Q_K''(\theta)| \leq K^2 \sup_{\theta\in(-\pi,\pi]} |Q_K(\theta)| \leq 2K^2\ . \tag{108}$$

Then,

$$|Q_K(\theta_M)| = \left|Q_K(0) + \int_0^{\theta_M} d\theta \left(Q_K'(0) + \int_0^{\theta} d\theta' Q_K''(\theta')\right)\right| \tag{109}$$

$$\leq |Q_K(0)| + \theta_M\, |Q_K'(0)| + K^2\theta_M^2. \tag{110}$$

Due to our condition in Eq. (107), we can impose $|Q_K(0)| \leq 2\epsilon$. Also, $Q_K(\theta)$ can in principle be a combination of even and odd degree polynomials. However, we note that

the even part $\frac{1}{2}(Q_K(\theta) + Q_K(-\theta))$ readily satisfies Eqs. (106) and (107) if $Q_K(\theta)$ does. Hence, we can assume that $Q_K(\theta)$ is an even polynomial and place a lower bound on its degree. This allow us to set $Q'_K(0) = 0$. Since $|Q_K(\theta_M)|$ must be within 2ϵ of $t\sin^2\theta_M$, we obtain

$$t\sin^2\theta_M - 2\epsilon \leq 2\epsilon + K^2\theta_M^2. \tag{111}$$

The condition $\Delta \leq 1$ implies $\theta_M \leq \frac{\pi}{2}\sin\theta_M = \frac{\pi}{2}\sqrt{\Delta}$, and the condition $t\Delta \geq 8\epsilon$ implies $t\Delta - 4\epsilon \geq t\Delta/2$, giving the desired result on the polynomial degree:

$$K \geq \frac{\sqrt{t\Delta - 4\epsilon}}{\theta_M} \geq \frac{2}{\pi}\sqrt{\frac{t\Delta}{2\Delta}} \geq \frac{1}{\pi}\sqrt{2t}. \tag{112}$$

□

The lower bound $\Omega(\sqrt{t})$ automatically generalizes to bound the degree of a trigonometric polynomial that approximates, for example, $\frac{1}{2}e^{-it\sin^2\theta}$ for $\theta \in [-\theta_M, \theta_M]$.

We now describe how this trigonometric polynomial degree bound can be adapted to lower bound the query complexity of low-energy simulation in the LCU model. To this end, we need the following result, which is similar to Claim 5.3 in Ref. [39].

Lemma 5.8 (Adapted from Ref. [39])**.** *Let U_θ be a unitary acting on $\mathbb{C}^2$ of the form*

$$U_\theta := \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{i\theta_M} \end{pmatrix}, \tag{113}$$

where θ_M is given and fixed, and $\theta \in (-\pi, \pi]$. Let $K > 0$ be a positive integer. Consider a quantum circuit on $q \geq 2$ qubits that has starting state $|0\rangle^{\otimes q}$, uses an arbitrary number of θ-independent two-qubit gates, uses K applications of controlled-U_θ and its inverse in total, and performs no intermediate measurements. Then the amplitudes of basis states before the final measurement are degree-K trigonometric polynomials in θ.

We provide the proof for completeness.

Proof. This is shown by induction. Let $V_0, \ldots, V_K$ be the θ-independent unitaries interleaved with the controlled-U_θ's and their inverses. Let $Q = 2^q$ be the dimension of the Hilbert space associated with the q qubits. For $K = 0$, we have $V_0|0\rangle^{\otimes q} = \sum_{i=0}^{Q-1} V_0^{i,0}|i\rangle$, where $V_k^{i,j}$ is the (i,j)-th matrix entry of V_k. These entries do not depend on θ, and then the claim is readily true for $K = 0$.

Let $|\psi_k\rangle$ be the state before the aplication of the $(k+1)$-th application of controlled-U_θ (or its inverse). Assume it can be written in the computational basis as

$$|\psi_k\rangle = \sum_{j\in\{0,1\}} \sum_{b\in\{0,1\}} \sum_{w\in\{0,1\}^{q-2}} P^{j,b,w}(\theta)\,|j\rangle\,|b\rangle\,|w\rangle\ , \tag{114}$$

where the first register is the $\mathbb{C}^2$ space U_θ acts on, the second register is the controlled qubits, and the third register corresponds to the remaining $q-2$ qubits used as workspace. Each $P^{j,b,w}(\theta)$ is a trigonometric polynomial of degree at most k by assumption.

We now apply the next controlled-U_θ. Note that

$$\text{controlled}{-}U_\theta\,|j\rangle\,|b\rangle\,|w\rangle \begin{cases} e^{i\theta}\,|0\rangle\,|1\rangle\,|w\rangle & \text{if } j=0, b=1\ , \\ e^{i\theta_M}\,|1\rangle\,|1\rangle\,|w\rangle & \text{if } j=1, b=1\ , \\ |j\rangle\,|0\rangle\,|w\rangle & \text{if } b=0\ . \end{cases} \tag{115}$$

Using this in Eq. (114) increases the degree of each $P_{j,b,w}(\theta)$ by at most 1. Applying V_{k+1} does not increase the degree of the polynomial because the transformation is θ-independent. A similar result applies when acting with the inverse of controlled-U_θ. Hence, we conclude that the state output by this quantum circuit takes the form

$$|\psi_K\rangle = \sum_{j\in\{0,1\}} \sum_{b\in\{0,1\}} \sum_{w\in\{0,1\}^{q-2}} P_K^{j,b,w}(\theta)\,|j\rangle\,|b\rangle\,|w\rangle \ , \tag{116}$$

where each $P_K^{j,b,w}(\theta)$ is a trigonometric polynomial of degree at most K. □

We are now ready to prove Thm. 5.3. Define the gap-amplifiable Hamiltonians

$$H_\theta := \left(\frac{U_\theta + (U_\theta)^\dagger}{2i}\right)^2 \ , \tag{117}$$

where θ_M is given, $0 < \theta_M \leq \pi/2$, and $0 \leq \theta \leq \theta_M$. Note that we can write $H_\theta = A_\theta^\dagger A_\theta$ (i.e., $\lambda = 1$), where

$$A_\theta^\dagger = A_\theta = \begin{pmatrix} \sin\theta & 0 \\ 0 & \sin\theta_M \end{pmatrix} \ . \tag{118}$$

Consider, for example, the initial state $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. This is supported on the subspace of eigenvalues at most $\Delta = \sin^2\theta_M$ of H_θ. Time evolution implies

$$\langle 0|\, e^{-itH_\theta}\,|\psi\rangle = \frac{1}{2} e^{-it\sin^2\theta} \ . \tag{119}$$

In the LCU model we assume access to U_θ, its inverse, and controlled versions. If the operator e^{-itH_θ} acting on $|\psi\rangle$ is approximated using controlled-U_θ and its inverse K times, Lemma 5.8 implies that the approximation to Eq. (119) is given by a trigonometric polynomial of degree at most K. If the approximation is done within error ϵ, and assuming the parameters of Lemma 5.7, then this degree must satisfy $K = \Omega(\sqrt{t})$. Note that while θ_M was given and fixed (similarly for Δ), the phase θ is unknown a priori and the simulation method is required to produce an ϵ-approximation for all $\theta \in [0, \theta_M]$. Also, note that the SELECT and PREPARE oracles in the LCU model for this example, which allow to block-encode A_θ, are simply obtained from a single controlled-U_θ and its inverse, and a Hadamard gate. This gives the desired result on the query complexity in the LCU model stated in Thm. 5.3.

6 Gate complexity lower bounds for simulating time evolution of low-energy states of gap-amplifiable Hamiltonians

We depart from the oracle model and instead show a gate complexity lower bound that addresses the time-dominated regime. In this regime, a state $|\psi\rangle$ has energy at most Δ in a (λ, L)-gap-amplifiable Hamiltonian, and the simulation parameters satisfy $\log(1/\epsilon) = o(t\Delta)$ and $t\Delta = o(t\lambda)$. We present a more formal version of Theorem 1.4.

Theorem 6.1 (Gate complexity lower bound for time evolution of low-energy states of gap-amplifiable Hamiltonians)**.** *For any n-qubit quantum circuit with G two-qubit gates such that $1 \leq n \leq G/n \leq 2^n$, there exists a sequence of 9-sparse and 4-local ($\lambda =$*

$4, L = 1)$*-gap-amplifiable Hamiltonians* $\{H_G\}_G$ *on the space* $\mathcal{H}_G$ *of* $\Theta(G)$ *qubits and a state* $|\psi_G\rangle \in \mathcal{S}_\Delta$ *for* $\Delta/\lambda = \mathcal{O}(1/G^2)$*, such that preparing a unitary* U *satisfying*

$$|\langle\psi_G| \Pi U^\dagger \Pi e^{-itH_G} |\psi_G\rangle| \geq 2/3 \,, \tag{120}$$

for time $t = \mathcal{O}(G^2)$ *and some projector* Π *onto* $\mathcal{H}_G$*, requires at least*

$$\tilde{\Omega}\left(t\sqrt{\lambda\Delta}\right) \tag{121}$$

two-qubit gates, where $\tilde{\Omega}$ *ignores logarithmic factors. The unitary* U *may use an unlimited number of ancilla qubits and the gates may be non-local and come from a possibly infinite gate set.*

To show Thm. 6.1, we use ideas developed in Refs. [31] and [28]. We time-evolve a Gaussian wavepacket (i.e., an initial state with Gaussian-like amplitudes in the computational basis), whose width and initial momentum are chosen such that it is a low-energy state of the so-called Feynman-Kitaev clock Hamiltonian. This Hamiltonian can be shown to be gap-amplifiable. After a sufficient amount of time, the wavepacket spreads and propagates sufficiently far to encode the result of a computation that performs G two-qubit gates. Using an argument based on counting distinct Boolean functions, we show that this implies that simulating the time evolution itself must have required at least G two-qubit gates. In terms of the parameters of the problem instances, this coincides with the claimed lower bound of $\tilde{\Omega}(t\sqrt{\lambda\Delta})$.

We schematically describe the argument here before introducing modifications required to realize the necessary technical details. In standard circuit-to-Hamiltonian reductions, the Feynman-Kitaev clock Hamiltonian reads

$$H = \sum_{x=1}^{G} \left(-U_x \otimes |x\rangle\langle x-1| - U_x^\dagger \otimes |x-1\rangle\langle x| + \mathbb{1} \otimes |x\rangle\langle x| + \mathbb{1} \otimes |x-1\rangle\langle x-1|\right) \,. \tag{122}$$

(We assume periodic boundary conditions were $|x = G\rangle \equiv |x = 0\rangle$.) It encodes a circuit $\mathcal{U} = U_G U_{G-1} \cdots U_1$ consisting of G two-qubit gates acting on n qubits ($N = 2^n$), and H acts on the space $\mathbb{C}^{2^{n+G}}$, where we choose a clock encoding implemented over G qubits. Up to discretization, the dynamics of H can be understood by considering a free particle on a line, since

$$H_{\mathrm{FP}} := W^\dagger H W = \mathbb{1}_N \otimes \sum_{x=1}^{G} \left(-|x\rangle\langle x-1| - |x-1\rangle\langle x| + |x\rangle\langle x| + |x-1\rangle\langle x-1|\right) \tag{123}$$

for unitary

$$W := \sum_{x=1}^{G} U_x \cdots U_1 \otimes |x\rangle\langle x| \,. \tag{124}$$

Note that H_{FP} is a translationally invariant tight-binding model; it can be interpreted as the discretized version of $\tilde{H}_{\mathrm{FP}} = -\partial_x^2/2$, the free-particle Hamiltonian in the continuum (with mass $m = 1$). We will show that there exists an initial state $|\psi_\Delta\rangle \in \mathcal{S}_\Delta$ within the low-energy subspace of H given by $\Delta = \Theta(1/G^2)$, such that evolution for time $t = \Theta(G^2)$ with constant error produces large support on the state that encodes the output of the circuit $\mathcal{U}$, after tracing out the clock space spanned by $\{|x\rangle\}_x$. We will then show a lower bound requiring $\tilde{\Omega}(G)$ gates to simulate the evolution of $|\psi_\Delta\rangle$.

To show that this result is tight with our upper bound, we first rewrite H in the form of a gap-amplifiable Hamiltonian to apply our algorithm. Using a binary encoding for $|x\rangle$ with $g = \log_2 G$ bits, note that

$$H_{\mathrm{FP}} = \mathbb{1}_N \otimes (\mathbb{1}_G - X)^\dagger (\mathbb{1}_G - X) \,, \tag{125}$$

where X is the $G \times G$ matrix that implements the cyclic permutation $X|x\rangle \to |x-1\rangle$. (We assumed that G is a power of two without loss of generality.) The lowest eigenvalue of H_{FP} is 0 and the largest is bounded by 2. We can define $A_{\mathrm{FP}} := \mathbb{1}_N \otimes (\mathbb{1}_G - X)/2$ so that $H_{\mathrm{FP}} = 4A_{\mathrm{FP}}^\dagger A_{\mathrm{FP}}$ is a $(\lambda = 4, L = 1)$-gap-amplifiable Hamiltonian. Note that our quantum algorithm can simulate $e^{-itH_{\mathrm{FP}}}$ acting on $|\psi_{\mathrm{FP}}\rangle$, which is a state of supported in the subspace of energies at most Δ, using $O(t\sqrt{\Delta} + 1/\sqrt{\Delta}) = \mathcal{O}(G)$ queries to the block-encoding of A_{FP}. This block-encoding can be constructed with polylog(G) two-qubit gates because of the simple structure of the operator A_{FP}. For example, applying X can be done using the quantum Fourier transform with $\mathcal{O}(g^2)$ two-qubit gates. Since H is a similarity transformation of H_{FP} (Eq. (123)), then H is also a $(\lambda = 4, L = 1)$-gap-amplifiable Hamiltonian: $H = 4A^\dagger A$, with $A := WA_{\mathrm{FP}}W^\dagger$. To simulate e^{-itH} on $|\psi_\Delta\rangle := W|\psi_{\mathrm{FP}}\rangle$, we simply compute $e^{-itH_{\mathrm{FP}}}|\psi_{\mathrm{FP}}\rangle$ and conjugate by W, since $e^{-itH} = We^{-itH_{\mathrm{FP}}}W^\dagger$. The gate complexity to simulate $e^{-itH}|\psi\rangle$ remains $\tilde{\mathcal{O}}(G)$, since the circuit to realize W requires only an additional $\tilde{\mathcal{O}}(G)$ two-qubit gates using standard techniques. This coincides with the lower bound, up to mild polylogarithmic corrections. The same result applies for other encodings of $|x\rangle$, like a unary encoding, as long as transforming from one encoding to the other can be done with $\tilde{\mathcal{O}}(G)$ two-qubit gates. Note that QSP and other generic methods that do not use the low-energy condition would have produced an algorithm of gate complexity $\mathcal{O}(G^2)$ or worse in this case, since $t = \Theta(G^2)$.

Next, we present some technical details required for the proof of Thm. 6.1.

6.1 Clock Hamiltonian and low-energy states

We now provide details of the clock Hamiltonian used in our construction, define the initial state, and show that this is confined to low-energy subspace. As before, we let $\mathcal{U} = U_G U_{G-1} \cdots U_1$ be a quantum circuit constructed from a universal gate set of two-qubit gates (which includes one-qubit gates), where $\mathcal{U}$ acts on n qubits. To obtain a local Hamiltonian, we encode the clock register in unary as

$$|0\rangle = |100\ldots0\rangle \,, \quad |1\rangle = |010\ldots0\rangle \,, \ldots, \quad |G'-1\rangle = |000\ldots1\rangle \tag{126}$$

using G' qubits, for some $G' \geq G$. This encoding will also allow us to show the result for a sparse Hamiltonian. We will specify the overhead $G' - G$ later; for now, it suffices to introduce integers $c_1, c_2 \geq 2$ constrained such that $G' = (c_1 + c_2)G$. Define unitaries V_x acting on the n qubits that pad the original circuit with identity gates, so that

$$V_x = \begin{cases} \mathbb{1}_N & \text{if } \ 1 \leq x \leq c_1 G \,, \\ U_{x-c_1 G} & \text{if } \ c_1 G + 1 \leq x \leq (c_1+1)G \,, \\ \mathbb{1}_N & \text{if } \ (c_1+1)G + 1 \leq x \leq (c_1+c_2)G \,, \end{cases} \tag{127}$$

with periodic boundary conditions (i.e., $x = 0$ is equivalent to $x = G'$). We consider the PSD clock Hamiltonian H acting on Hilbert space $\mathcal{H}_G := \mathbb{C}^{2^{n+G'}}$ of $n + G'$ qubits, of periodic boundary conditions, defined by

$$H := \sum_{x=1}^{G'} \left(-V_x \otimes |x\rangle\langle x-1| - V_x^\dagger \otimes |x-1\rangle\langle x| + \mathbb{1}_N \otimes |x\rangle\langle x| + \mathbb{1}_N \otimes |x-1\rangle\langle x-1| \right) . \tag{128}$$

This Hamiltonian is at most 9-sparse because the V_x's and $V_x^\dagger$'s are 4-sparse in the worst case (two-qubit gates), and H contains these unitaries in the off-diagonal blocks. It is also 4-local because the elements $|x\rangle\langle x|$ and $|x\rangle\langle x-1|$ can be replaced by one-qubit and two-qubit interactions, respectively. The padding of identities $V_1 = \cdots = V_{c_1 G} = \mathbb{1}_N$ allows us to initialize a wavepacket (state) over the clock register without interfering with the computation. That is, if the computation on the n qubits starts in $|0\rangle^{\otimes n}$, we can choose *any* initial superposition over the clock register like

$$|\psi\rangle = |0\rangle^{\otimes n} \otimes \sum_{x=0}^{c_1 G} \psi(x) |x\rangle \ , \tag{129}$$

without requiring any information from $\mathcal{U}$. To additionally restrict $|\psi\rangle$ to be of low energy, we choose a Gaussian wavepacket centered at site $x_0 = (c_1/2)G$ and with initial rightwards momentum $p_0 > 0$, i.e.,

$$\psi(x) = \eta \exp\left[-\frac{(x-x_0)^2}{2\sigma^2} + ip_0 x\right], \quad \eta = \left(\sum_{x=0}^{c_1 G} \exp\left[-\frac{(x-x_0)^2}{\sigma^2}\right]\right)^{-1/2} . \tag{130}$$

Below, we will choose the wavepacket to have width σ linear in G and initial rightwards momentum p_0 linear in $1/G$. We wish to show that these choices imply that $|\psi\rangle$ is mostly confined to a low-energy subspace with cutoff $\Delta = \mathcal{O}(1/G^2)$. (The choice of an initial momentum being nonzero is not necessary, but we use it to relate our results with Ref. [31] more directly.)

In the continuum limit, where $x \in \mathbb{R}$, these facts are apparent. The Hamiltonian becomes that of a free particle, i.e., $\tilde{H}_{\mathrm{FP}} = p^2/2m = -\partial_x^2/2$, with $m = 1$. Its eigenfunctions are e^{-ipx}. The spatial width $\sigma \sim G$ of a Gaussian wavepacket implies, under a Fourier transform, a width of $\sim 1/G$ in momentum space. The initial momentum $p_0 \sim 1/G$ scales identically with respect to G. Hence, the initial state of Gaussian amplitudes in the continuum is easily shown to be mostly supported in the subspace of energies at most $\Delta = \mathcal{O}(1/G^2)$ of $\tilde{H}_{\mathrm{FP}}$.

In practice, our initial state $|\psi\rangle$ is both defined on a discrete finite-dimensional system and has truncated tails compared to a true Gaussian (due to Eq. (129)). Nevertheless, we show that it has high overlap with a state $|\psi_\Delta\rangle$ that is fully confined to a low-energy subspace of H specified by some $\Delta = \mathcal{O}(1/G^2)$, by explicitly evaluating the energy of the wavepacket under the clock Hamiltonian and applying Markov's inequality.

Lemma 6.2 (Low-energy state of clock Hamiltonian)**.** *Let H be the clock Hamiltonian for a G-gate quantum circuit $\mathcal{U} = U_G \dots U_1$ as defined in Eq. (128); it contains $c_1 G$ initial identity gates and $(c_2 - c_1 - 1)G$ final identity gates, for even positive integers c_1, c_2. Let $|\psi\rangle$ be the state defined in Eq. (129), with amplitudes determined by Eqs. (130), and centered at $x_0 = (c_1/2)G$. Fix the width of the wavepacket to $\sigma = \hat{\sigma} x_0$ and the momentum to $p_0 = \hat{p}_0/x_0$ for constants $\hat{\sigma} > 0$, $\hat{p}_0 > 0$. Then, for any (arbitrarily small) constant $\delta > 0$, there exists $\Delta = \mathcal{O}(1/G^2)$ such that the state*

$$|\psi_\Delta\rangle := \frac{\Pi_\Delta |\psi\rangle}{\|\Pi_\Delta |\psi\rangle\|}, \tag{131}$$

where Π_Δ is the projector onto the low-energy subspace specified by Δ, satisfies

$$|\langle\psi_\Delta|\psi\rangle| \geq 1 - \delta. \tag{132}$$

Proof. Direct computation gives

$$\langle\psi| H |\psi\rangle = \sum_{x=1}^{2x_0} |\psi(x) - \psi(x-1)|^2 \tag{133}$$

$$= \eta^2 \sum_{x=1}^{2x_0} \exp\left[-\frac{(x-(x_0+1/2))^2}{\sigma^2} - \frac{1}{4\sigma^2}\right] \times \left(\exp\left[\frac{x-x_0-1/2}{\sigma^2}\right] + \exp\left[-\frac{x-x_0-1/2}{\sigma^2}\right] - 2\cos(p_0)\right) . \tag{134}$$

For $x_0 = c_1 G/2 = \Theta(G)$, we choose wavepacket width $\sigma = \hat{\sigma} x_0 = \Theta(G)$ and momentum $p_0 = \hat{p}_0/x_0 = \Theta(1/G)$ for constants $\hat{\sigma}, \hat{p}_0$. The terms of the sum are independent of x up to $1/G$ corrections, i.e.,

$$\langle\psi| H |\psi\rangle = \eta^2 \sum_{x=1}^{2x_0} \left(\exp\left[-\frac{1}{\hat{\sigma}^2}\right] + \mathcal{O}(G^{-1})\right)\left(\frac{\hat{p}_0^2}{x_0^2} + \frac{1}{\hat{\sigma}^2 x_0^2} + \mathcal{O}(G^{-3})\right) \tag{135}$$

$$= \mathcal{O}\left(\frac{1}{G^2}\right) , \tag{136}$$

where in the last step we used the fact that $\eta = \Theta(G^{-1/2})$. Let H have eigenstates $|\varphi_k\rangle$ with corresponding eigenvalues $\gamma_k \geq 0$. By Markov's inequality, the support of $|\psi\rangle$ on eigenstates with energy at least a/G^2 is bounded by

$$\sum_{k\,:\,\gamma_k \geq a/G^2} |\langle\varphi_k|\psi\rangle|^2 = \mathcal{O}\left(\frac{1}{a}\right) , \tag{137}$$

for any $a > 0$. Consequently, for any $\delta > 0$, one can always choose $\Delta = \mathcal{O}(1/G^2)$ such that the state $|\psi_\Delta\rangle$ defined in Eq. (131) satisfies

$$|\langle\psi_\Delta|\psi\rangle| \geq 1 - \delta. \tag{138}$$

□

The state $|\psi_\Delta\rangle$ corresponds to $|\psi_G\rangle$ in Thm. 6.1. Nevertheless, we will continue our analysis to obtain a gate complexity lower bound on time-evolving $|\psi\rangle$ with constant fidelity; due to Lemma 6.2, this implies an equivalent gate complexity lower bound on time-evolving $|\psi_\Delta\rangle$.

6.2 Evaluating $\mathcal{U}$ via time evolution of low-energy states

We now show that time evolution of the low-energy state $|\psi\rangle$ or $|\psi_\Delta\rangle$ for time $t = \Theta(G^2)$ produces large support on the state $U_G \cdots U_1 |0\rangle^{\otimes n}$ of the n qubits. We sketch this by first considering the continuum limit, where $x \in \mathbb{R}$. The result in this case is readily available by the dynamics under the free particle Hamiltonian $\tilde{H}_{\mathrm{FP}} = -\partial_x^2/2$. A Gaussian wavepacket at time $t = 0$, centered at x_0 and with momentum p_0,

$$\psi(x,0) \propto \exp\left[-\frac{(x-x_0)^2}{2\sigma^2} + i p_0 x\right] , \tag{139}$$

evolves under Schrödinger's equation $i\frac{\partial\psi(x,t)}{\partial t} = -\frac{1}{2}\frac{\partial^2\psi(x,t)}{\partial x^2}$ to

$$\psi(x,t) \propto \exp\left[-\frac{(x-x_0(t))^2}{2\sigma(t)^2} + i\phi(x,t)\right] , \tag{140}$$

for phase $\phi(x,t) \in \mathbb{R}$ and

$$x_0(t) = x_0 + p_0 t, \quad \sigma(t) = \sqrt{\sigma^2 + \left(\frac{t}{\sigma}\right)^2} \,. \tag{141}$$

That is, the center of the Gaussian wavepacket is spatially translated from x_0 to $x_0 + tp_0$, and its width increases in time to $\sigma(t)$. As discussed above, our choice of initial state $|\psi\rangle$ is, in the corresponding continuum limit, a Gaussian wavepacket with spatial width $\sigma \sim G$ centered at $x_0 \sim G$ and with initial momentum $p_0 \sim 1/G$. The clock Hamiltonian is a discretized version of $\tilde{H}_{\mathrm{FP}}$. To perform a computation that evaluates $\mathcal{U} = U_G U_{G-1} \cdots U_1$, we need t to be sufficiently large to apply the next G terms in the clock Hamiltonian; that is, we choose the parameters so that $x_0(t) - x_0 \sim G$. Additionally, if we choose $tp_0 \sim t/G$ to "advance" G steps, we obtain a time $t \sim G^2$. This widens the wavepacket by a constant factor (i.e., $\sigma(G^2)/\sigma(0) > 1$), allowing the evolved state to be supported in a sufficiently large subspace related to having applied the corresponding non-trivial G gates.

To formally prove that time evolution evaluates $U_G \dots U_1 |0\rangle^{\otimes n}$ with high probability, we can start from the analysis in the continuum limit, and then bound errors due to the discretization and truncation of the initial state $|\psi\rangle$. The proof is based on techniques established in Ref. [48] for analyzing the dynamics of one-dimensional systems. Due to the length of this technical analysis, we provide the proof in Appendix D.

Lemma 6.3 (Time evolution of initial Gaussian-like state under the clock Hamiltonian)**.** *Let H be the clock Hamiltonian for the G-gate unitary $\mathcal{U} = U_G U_{G-1} \dots U_1$ acting on n qubits as defined in Eq. (128); it contains $c_1 G$ initial identity gates and $(c_2 - c_1 - 1)G$ final identity gates, for even integers c_1, c_2. Let $|\psi\rangle$ be the state defined in Eq. (129), with amplitudes determined by Eqs. (130), and centered at site $x_0 = (c_1/2)G$. Fix the width of the wavepacket to $\sigma = \hat{\sigma} x_0$ and the momentum to $p_0 = \hat{p}_0/x_0$ for positive constants $\hat{\sigma}$, and $\hat{p}_0$. Consider the time-evolved state $e^{-itH} |\psi\rangle$ for time $t = \hat{t} x_0^2 = \mathcal{O}(G^2)$ for constant $\hat{t} > 0$. Then, for any $G \geq 1$, there exists a choice of constants c_1, c_2, $\hat{\sigma}$, $\hat{p}_0$, and $\hat{t}$, such that*

$$\frac{1}{2} \left\| \mathcal{U} |0\rangle\langle 0|^{\otimes n} \mathcal{U}^\dagger - \Pi_n e^{-itH} |\psi\rangle\langle\psi| e^{itH} \Pi_n \right\|_1 \leq \frac{1}{4} \,, \tag{142}$$

where Π_n is the projector onto the space of the n qubits where the computation occurs, and $\frac{1}{2}\|\cdot\|_1$ indicates the trace distance.

In particular, this implies

$$\langle 0|^{\otimes n} \mathcal{U}^\dagger \rho(t) \mathcal{U} |0\rangle^{\otimes n} \geq 3/4 \,, \tag{143}$$

for $\rho(t) = \mathrm{tr}_{\mathrm{clock}}(e^{-itH} |\psi\rangle\langle\psi| e^{itH})$, i.e., the state of the n qubits after evolving with H and tracing out the clock register. By making this overlap to be larger than 2/3, we can allow other errors to be small constants and still satisfy Eq. (120).

6.3 Gate complexity lower bound

We summarize the analysis and results thus far. For an arbitrary circuit $\mathcal{U} = U_G U_{G-1} \cdots U_1$ acting on n qubits with G two-qubit gates, we introduced a gap-amplifiable clock Hamiltonian $H = \lambda A^\dagger A$, with $\lambda = 4$. We defined a state $|\psi_\Delta\rangle \in \mathcal{S}_\Delta$, for $\Delta = \mathcal{O}(1/G^2)$, that has overlap $|\langle\psi_\Delta|\psi\rangle| \geq 1 - \delta$ for arbitrarily small constant $\delta > 0$, where $|\psi\rangle$ is the state with Gaussian-like amplitudes in Eqs. (129) and (130). Consequently, a gate complexity lower bound on time-evolving $|\psi\rangle$ under H with constant fidelity implies a lower bound

on time-evolving the low-energy state $|\psi_\Delta\rangle$. We showed that evolution for time $t = \Theta(G^2)$ produces a state close to $\mathcal{U}|0\rangle^{\otimes n}$. We must now show that this implies a lower bound on the gate complexity of time evolving low-energy states of gap-amplifiable Hamiltonians. Specifically, we will show that preparing $\mathcal{U}|0\rangle^{\otimes n}$ with constant probability requires $\tilde{\Omega}(G) = \tilde{\Omega}(t\sqrt{\lambda\Delta})$ gates, implying the same lower bound for time-evolving $|\psi_\Delta\rangle$, which is the state $|\psi_G\rangle$ in Thm. 6.1.

We follow a similar argument to Ref. [28]. To this end, we constrain the circuits $\mathcal{U}$ to those that compute distinct Boolean functions $f : \{0,1\}^n \to \{0,1\}$. Suppose briefly that $\mathcal{U}$ acts on n qubits and an unlimited number of ancilla qubits. We say that $\mathcal{U}$ computes a Boolean function f with high probability if measuring the first output qubit of $\mathcal{U}|x_1x_2\cdots x_n0\cdots0\rangle$ yields $f(x)$ with probability at least $2/3$. Restricting only to $\mathcal{U}$ that act on a total of n qubits, we count the number of distinct Boolean functions that $\mathcal{U}$ can encode in terms of its depth T.

Lemma 6.4 (Lower bound on distinct Boolean functions computed by $\mathcal{U}$; Lemma 8 of Ref. [28])**.** *For any integers n and T such that $2 \leq n \leq T \leq 2^n$, the number of distinct Boolean functions $f : \{0,1\}^n \to \{0,1\}$ that can be computed by depth-T quantum circuits on n qubits that use geometrically local two-qubit gates on a constant-dimensional lattice from a finite get set is at least $2^{\tilde{\Omega}(Tn)}$.*

Even if we allow an unlimited number of ancilla qubits, we have an upper bound on the number of Boolean functions that such $\mathcal{U}$'s with G two-qubit gates can compute with high probability.

Lemma 6.5 (Upper bound on computing Boolean functions with high probability; Lemma 9 of Ref. [28])**.** *The number of Boolean functions $f : \{0,1\}^n \to \{0,1\}$ that can be computed with high probability by quantum circuits with unlimited ancilla qubits using G two-qubit gates from any gate set is at most $2^{\tilde{O}(G\log n)}$.*

We can now prove Thm. 6.1. In Lemma 6.3, we showed that time evolution of the state $|\psi\rangle$ under the clock Hamiltonian for a circuit $\mathcal{U}$ with G gates can evaluate $\mathcal{U}|0\rangle^{\otimes n}$ in time $t = \mathcal{O}(G^2)$ with high probability. Since $Tn \geq G$, Lemma 6.4 implies that the number of distinct Boolean functions that can be computed by time-evolving with the Hamiltonians for corresponding $\mathcal{U}$'s of G gates, and with high probability, is at least $2^{\tilde{\Omega}(G)}$. However, by Lemma 6.5, a circuit that uses G gates can compute at most $2^{\tilde{O}(G\log n)}$ Boolean functions with high probability. This implies that $\tilde{\Omega}(G)$ two-qubit gates are required to perform the time evolution with constant error in the worst case. Equivalently, replacing the state $|\psi\rangle$ with its low-energy projection $|\psi_\Delta\rangle \in \mathcal{S}_\Delta$, we conclude that

$$\tilde{\Omega}\left(t\sqrt{\lambda\Delta}\right) \tag{144}$$

two-qubit gates are required to simulate the time evolution of $|\psi_\Delta\rangle$ with constant error and in the worst case. Note that in Thm. 6.1, we report this gate complexity under the equivalent condition $|\langle\psi_G|\,\Pi U^\dagger \Pi e^{-itH}\,|\psi_G\rangle| \geq 2/3$, where $|\psi_G\rangle \equiv |\psi_\Delta\rangle$.

7 Conclusions

We described a quantum algorithm for simulating the time evolution of states supported in a low-energy subspace of a Hamiltonian. Under certain conditions, the quantum algorithm provides improved runtimes over alternative Hamiltonian simulation methods that apply

generally, like QSP. Specifically, the improvement arises when the Hamiltonian is "gap-amplifiable", i.e., it is PSD, can be expressed as $H = \lambda A^\dagger A$, and its "square root" A can be efficiently accessed. These Hamiltonians include frustration-free Hamiltonians that appear ubiquitously in physics and quantum computing. In applications of these examples, creating access to A (e.g., in the form of an LCU or a block-encoding) may incur a similar cost as creating access to H. Hence, our improved bounds on query complexity are anticipated to translate into improved bounds on gate complexities, provided that the initial state is mostly supported on the low-energy subspace of such Hamiltonians. In general, we observe that it is always possible to shift a Hamiltonian by a constant to become PSD. While this shift might not yield any improvement in the worst-case scenario, it is conceivable that some moderate improvement could still be achieved by employing our algorithm in this case; we provided an example of this using a nearly frustration-free Hamiltonian.

We also proved several important lower bounds for the query complexity of this Hamiltonian simulation problem. First, we showed that the low-energy condition of the initial state cannot be exploited to improve the runtime of Hamiltonian simulation methods that apply generally, like QSP, in the worst case (e.g., when we do not have efficient access to A). This answers a long-standing question in Hamiltonian simulation; for example Ref. [44] also required the Hamiltonians to be PSD and noted that it was not clear how to improve the runtime of Hamiltonian simulation in the low-energy subspace in general. Our results represent strengthened lower bounds compared to previous no-fast-forwarding theorems for generic Hamiltonians.

Second, we proved matching lower bounds in query complexity to show that our algorithm has optimal dependence on time and the size of the low-energy subspace. In the asymptotic (time-dominated) regime that is most interesting for applications, our quantum algorithm for gap-amplifiable Hamiltonians is a strict improvement over generic methods like QSP. Beyond the oracle model, we also proved a matching lower bound on the gate complexity in this regime. We note that in the asymptotic (intermediate) regime where time-dependence in sublinear, the cost of our algorithm also depends on the precision; we proved a lower bound in terms of time in this regime but did not prove a lower bound in in terms the precision, however, we note that our upper bound only has sublogarithmic dependence on $1/\epsilon$.

Our quantum algorithm builds on a polynomial approximation to e^{-itx^2} as a function of x. Since we are interested in a domain where $|x| \ll 1$, the degree of the polynomial is smaller than that needed for approximating the function for all $x \in [-1, 1]$. For gap-amplifiable Hamiltonians, x is associated with A (or $A^\dagger$), and the degree of the polynomial with the query complexity. Related results and improvements are thus expected for implementing other functions of gap-amplifiable Hamiltonians when the initial state is low-energy.

Since low-energy states appear ubiquitously in physics simulation, quantum chemistry, optimization and beyond, we expect our results to be broadly useful.

8 Acknowledgements

We thank Robin Kothari for discussions, feedback on the presentation, and for pointing out Ref. [28], and Thomas E. O'Brien for reviewing an earlier draft of this article and comments. We also thank Ryan Babbush, William J. Huggins, Stephen Jordan, Robbie King, and Ramis Movassagh at Google's Quantum AI, and Isaac Chuang and Aram Harrow at MIT, for discussions.

References

[1] Tameem Albash and Daniel A Lidar. Adiabatic quantum computation. *Reviews of Modern Physics*, 90(1):015002, 2018. DOI: 10.1103/RevModPhys.90.015002.

[2] Dong An, Jin-Peng Liu, and Lin Lin. Linear combination of hamiltonian simulation for nonunitary dynamics with optimal state preparation cost. *Phys. Rev. Lett.*, 131: 150603, Oct 2023. DOI: 10.1103/PhysRevLett.131.150603.

[3] 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. DOI: 10.1126/science.1113479.

[4] Yosi Atia and Dorit Aharonov. Fast-forwarding of hamiltonians and exponentially precise measurements. *Nature communications*, 8(1):1572, 2017. DOI: 10.1038/s41467-017-01637-7.

[5] Ryan Babbush, Dominic W Berry, Ian D Kivlichan, Annie Y Wei, Peter J Love, and Alán Aspuru-Guzik. Exponentially more precise quantum simulation of fermions in second quantization. *New Journal of Physics*, 18(3):033032, mar 2016. DOI: 10.1088/1367-2630/18/3/033032.

[6] Ryan Babbush et al. Exponentially more precise quantum simulation of fermions in the configuration interaction representation. *Quantum Science and Technology*, 3(1): 015006, (2017). DOI: 10.1088/2058-9565/aa9463.

[7] Dominic W Berry and Andrew M. Childs. Black-box hamiltonian simulation and unitary implementation. *Quantum Information & Computation*, 12(1-2):29–62, (2012). DOI: 10.26421/QIC12.1-2-4.

[8] Dominic W Berry, Graeme Ahokas, Richard Cleve, and Barry C. Sanders. Efficient quantum algorithms for simulating sparse hamiltonians. *Communications in Mathematical Physics*, 270(2):359–371, (2007). DOI: 10.1007/s00220-006-0150-x.

[9] Dominic W Berry, Andrew M Childs, Richard Cleve, Robin Kothari, and Rolando D Somma. Simulating hamiltonian dynamics with a truncated taylor series. *Physical review letters*, 114(9):090502, (2015). DOI: 10.1103/PhysRevLett.114.090502.

[10] Dominic W Berry, Andrew M Childs, Richard Cleve, Robin Kothari, and Rolando D Somma. Exponential improvement in precision for simulating sparse hamiltonians. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 283–292, 2014. DOI: 10.1145/2591796.2591854.

[11] Sergio Boixo, Emanuel Knill, and Rolando D Somma. Eigenpath traversal by phase randomization. *Quantum Inf. Comput.*, 9(9&10):833–855, 2009. DOI: 10.26421/QIC9.9-10-7.

[12] Peter Borwein and Tamás Erdélyi. *Polynomials and polynomial inequalities*, volume 161. Springer Science & Business Media, 2012. DOI: 10.1007/978-1-4612-0793-1.

[13] Michel Boyer, Gilles Brassard, Peter Høyer, and Alain Tapp. Tight bounds on quantum searching. *Fortschritte der Physik: Progress of Physics*, 46(4-5):493–505, 1998. DOI: 10.1002/3527603093.ch10.

[14] Sergey Bravyi and Barbara Terhal. Complexity of stoquastic frustration-free hamiltonians. *Siam journal on computing*, 39(4):1462–1485, (2010). DOI: 10.1137/08072689X.

[15] Earl Campbell. A random compiler for fast hamiltonian simulation. *Physical review letters*, 123:070503, (2019). DOI: 10.1103/PhysRevLett.123.070503.

[16] Jianxin Chen, Zhengfeng Ji, David Kribs, Zhaohui Wei, and Bei Zeng. Ground-state spaces of frustration-free hamiltonians. *Journal of mathematical physics*, 53(10), 2012. DOI: 10.1063/1.4748527.

[17] Andrew M Childs and Jeffrey Goldstone. Spatial search by quantum walk. *Physical Review A*, 70(2):022314, 2004. DOI: 10.1103/PhysRevA.70.022314.

[18] Andrew M Childs, Robin Kothari, and Rolando D Somma. Quantum algorithm for systems of linear equations with exponentially improved dependence on precision. *SIAM Journal on Computing*, 46(6):1920–1950, (2017). DOI: 10.1137/16M1087072.

[19] Andrew M Childs, Yuan Su, Minh C Tran, Nathan Wiebe, and Shuchen Zhu. Theory of trotter error with commutator scaling. *Physical Review X*, 11(1):011020, (2021). DOI: 10.1103/PhysRevX.11.011020.

[20] Anirban Narayan Chowdhury and Rolando D Somma. Quantum algorithms for gibbs sampling and hitting-time estimation. *Quant. Inf. Comp.*, 17:0041, 2017. DOI: 10.1145/3313276.3316366.

[21] Niel de Beaudrap, Matthias Ohliger, Tobias J Osborne, and Jens Eisert. Solving frustration-free spin systems. *Physical review letters*, 105:060504, (2010). DOI: 10.1103/PhysRevLett.105.060504.

[22] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, and Michael Sipser. Quantum computation by adiabatic evolution. *Preprint at https://arxiv.org/abs/quant-ph/0001106*, 2000. DOI: 10.48550/arXiv.\protect \vrule width0pt\protect \href {http://arxiv.org/abs/quant-ph/0001106}{quant-ph/0001106}.

[23] Richard P Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6–7):467–488, 1982. DOI: 10.1007/BF02650179.

[24] András Gilyén. personal communication, 2023.

[25] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 193–204, 2019. DOI: 10.1145/3313276.3316366.

[26] Weiyuan Gong, Shuo Zhou, and Tongyang Li. Complexity of Digital Quantum Simulation in the Low-Energy Subspace: Applications and a Lower Bound. *Quantum*, 8: 1409, July 2024. ISSN 2521-327X. DOI: 10.22331/q-2024-07-15-1409.

[27] Shouzhen Gu, Rolando D Somma, and Burak Şahinoğlu. Fast-forwarding quantum evolution. *Quantum*, 5:577, November 2021. DOI: 10.22331/q-2021-11-15-577.

[28] Jeongwan Haah, Matthew Hastings, Robin Kothari, and Guang Hao Low. Quantum algorithm for simulating real time evolution of lattice hamiltonians. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 350–360. IEEE, (2018). DOI: 10.1137/18M1231511.

[29] Zoe Holmes, Gopikrishnan Muraleedharan, Rolando D Somma, Yigit Subasi, and Burak Şahinoğlu. Quantum algorithms from fluctuation theorems: Thermal-state preparation. *Quantum*, 6:825, 2022. DOI: 10.22331/q-2022-10-06-825.

[30] Peter Hoyer, Troy Lee, and Robert Spalek. Negative weights make adversaries stronger. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 526–535, 2007. DOI: 10.1145/1250790.1250867.

[31] Stephen P. Jordan. Fast quantum computation at arbitrarily low energy. *Phys. Rev. A*, 95:032305, Mar 2017. DOI: 10.1103/PhysRevA.95.032305.

[32] Stephen P Jordan, Keith S.M. Lee, and John Preskill. Quantum algorithms for quantum field theories. *Science*, 336:1130, (2012). DOI: 10.1126/science.1217069.

[33] Benjamin P Lanyon, James D Whitfield, Geoff G Gillett, Michael E Goggin, Marcelo P Almeida, Ivan Kassal, Jacob D Biamonte, Masoud Mohseni, Ben J Powell, Marco Barbieri, et al. Towards quantum chemistry on a quantum computer. *Nature chemistry*, 2(2):106–111, 2010. DOI: 10.1038/nchem.483.

[34] Seth Lloyd. Universal quantum simulators. *Science*, 273(5278):1073–1078, (1996). DOI: 10.1126/science.273.5278.1073.

[35] Guang Hao Low. Hamiltonian simulation with nearly optimal dependence on spectral norm. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, page 491–502, New York, NY, USA, 2019. DOI: 10.1145/3313276.3316386.

[36] Guang Hao Low and Isaac L. Chuang. Optimal hamiltonian simulation by quantum signal processing. *Physical review letters*, 118(1):010501, (2017). DOI: 10.1103/PhysRevLett.118.010501.

[37] Guang Hao Low and Isaac L Chuang. Hamiltonian simulation by uniform spectral amplification. *Preprint at https://arxiv.org/abs/1707.05391*, 2017. DOI: 10.48550/arXiv.1707.05391.

[38] Guang Hao Low and Isaac L Chuang. Hamiltonian simulation by qubitization. *Quantum*, 3:163, 2019. DOI: 10.22331/q-2019-07-12-163.

[39] Nikhil S Mande and Ronald de Wolf. Tight bounds for quantum phase estimation and related problems. *Preprint at https://arxiv.org/abs/2305.04908*, 2023. DOI: 10.48550/arXiv.2305.04908.

[40] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010. DOI: 10.1119/1.1463744.

[41] Davide Orsucci and Vedran Dunjko. On solving classes of positive-definite quantum linear systems with quadratically improved runtime in the condition number. *Quantum*, 5:573, November 2021. ISSN 2521-327X. DOI: 10.22331/q-2021-11-08-573.

[42] Ben W Reichardt. Span programs are equivalent to quantum query algorithms. *SIAM Journal on Computing*, 43(3):1206–1219, 2014. DOI: 10.1137/100792640.

[43] Subir Sachdev. Quantum phase transitions. *Physics world*, 12(4):33, 1999. DOI: 10.1002/9780470022184.hmm108.

[44] Burak Şahinoğlu and Rolando D Somma. Hamiltonian simulation in the low-energy subspace. *npj Quantum Information*, 7(1):119, 2021. DOI: 10.1038/s41534-021-00451-w.

[45] Norbert Schuch, Ignacio Cirac, and Frank Verstraete. Computational difficulty of finding matrix product ground states. *Physical review letters*, 100(25):250501, 2008. DOI: 10.1103/PhysRevLett.100.250501.

[46] Rolando Somma, Gerardo Ortiz, James E Gubernatis, Emanuel Knill, and Raymond Laflamme. Simulating physical phenomena by quantum networks. *Physical Review A*, 65(4):042323, (2002). DOI: 10.1103/PhysRevA.65.042323.

[47] Rolando D Somma. A trotter-suzuki approximation for lie groups with applications to hamiltonian simulation. *Journal of Mathematical Physics*, 57(6):062202, (2016). DOI: 10.1063/1.4952761.

[48] Rolando D. Somma. Quantum simulations of one dimensional quantum systems. *Quantum Info. Comput.*, 16(13–14):1125–1168, oct 2016. ISSN 1533-7146.

[49] Rolando D Somma and Sergio Boixo. Spectral gap amplification. *SIAM Journal on Computing*, 42(2):593–610, 2013. DOI: 10.1137/120871997.

[50] Matthew Thibodeau and Bryan K. Clark. Nearly-frustration-free ground state preparation. *Quantum*, 7:1084, August 2023. ISSN 2521-327X. DOI: 10.22331/q-2023-08-16-1084.

[51] Minh C Tran, Yuan Su, Daniel Carney, and Jacob M Taylor. Faster digital quantum simulation by symmetry protection. *PRX Quantum*, 2(1):010323, 2021. DOI: 10.1103/PRXQuantum.2.010323.

[52] Nathan Wiebe, Dominic Berry, Peter Høyer, and Barry C. Sanders. Higher order decompositions of ordered operator exponentials. *Journal of Physics A: Mathematical and Theoretical*, 43(6):065203, (2010). DOI: 10.1088/1751-8113/43/6/065203.

A Proof of Lemma 2.1

We prove Lemma 2.1, reproduced here for convenience. It is a simplified form of Corollary 66 of Ref. [25], with a minor correction in the bounds of ϵ from $\epsilon \in (0, 1/2B)$ to $\epsilon \in (0, 3B/2)$ [24]. Consistent with our notation throughout this work, we will ignore log log factors associated with the Jacobi-Anger expansion in Lemma A.2.

Lemma 2.1 (Piecewise polynomial approximations based on a local Taylor series)**.** *Let $r \in (0, 1]$, $\delta \in (0, r]$, and let $f : [-r-\delta, r+\delta]$ be such that $f(x) = \sum_{k=0}^{\infty} a_k x^k$ for all $x \in [-r-\delta, r+\delta]$. Choose any $B > 0$ such that $B \geq \sum_{k=0}^{\infty} |a_k|(r+\delta)^k$. For $\epsilon \in (0, 3B/2)$, there is an efficiently computable polynomial P of degree $\mathcal{O}(\frac{1}{\delta} \log \frac{B}{\epsilon})$ such that*

$$|f(x) - P(x)|_{[-r,r]} \leq \epsilon \quad \text{and} \quad |P(x)|_{[-1,1]} \leq \epsilon + |f(x)|_{[-r-\frac{\delta}{2}, r+\frac{\delta}{2}]} \,. \tag{145}$$

Proof. Our proof largely follows that of Ref. [25]. Let $L(x) = \frac{x}{r+\delta}$ be the linear transformation from $[-r-\delta, r+\delta]$ to $[-1, 1]$. Let $g(y) = f(L^{-1}(y))$; this looks like $f(x)$ on $[-r, r]$ expanded onto $[-1, 1]$. Since $y = x/(r+\delta)$, we have $g(y) = \sum_{k=0}^{\infty}(a_k(r+\delta)^k)y^k = \sum_{k=0}^{\infty} b_k y^k$ for $b_k := a_k(r+\delta)^k$. For any $K \geq 1$ and $x : |x| \leq r$, we obtain

$$\left| g(y) - \sum_{k=0}^{K-1} b_k y^k \right| = \left| \sum_{k=K}^{\infty} b_k y^k \right| \leq \sum_{k=K}^{\infty} \left| b_k \left(1 - \frac{\delta}{2(r+\delta)}\right)^k \right| \,, \tag{146}$$

since we only need to bound the error for $|x| \leq r$, so that

$$|y| \leq \frac{r}{r+\delta} = 1 - \frac{\delta}{r+\delta} \leq 1 - \frac{\delta}{2(r+\delta)} \,. \tag{147}$$

In particular, we will take

$$K = \left\lceil \frac{2(r+\delta)}{\delta} \log\left(\frac{12B}{\epsilon}\right) \right\rceil , \tag{148}$$

where $\frac{12B}{\epsilon} > 1$, i.e., $\epsilon < 12B$. Continuing, we have

$$\sum_{k=K}^{\infty} \left| b_k \left(1 - \frac{\delta}{2(r+\delta)}\right)^k \right| \leq \left(1 - \frac{\delta}{2(r+\delta)}\right)^K \sum_{k=K}^{\infty} |b_k| \tag{149}$$

$$\leq \exp\left[-K \frac{\delta}{2(r+\delta)}\right] \sum_{k=K}^{\infty} |a_k|(r+\delta)^k \tag{150}$$

$$\leq \exp\left[-K \frac{\delta}{2(r+\delta)}\right] \sum_{k=0}^{\infty} |a_k|(r+\delta)^k \tag{151}$$

$$\leq \exp\left[-K \frac{\delta}{2(r+\delta)}\right] B \,, \tag{152}$$

since $\left(1 - \frac{\delta}{2(r+\delta)}\right)^K \leq \exp\left[-K \frac{\delta}{2(r+\delta)}\right]$. Finally, we plug in K to obtain

$$\left| g(y) - \sum_{k=0}^{K-1} b_k y^k \right| \leq \exp\left[- \left\lceil \frac{2(r+\delta)}{\delta} \log\left(\frac{12B}{\epsilon}\right) \right\rceil \frac{\delta}{2(r+\delta)} \right] B \leq \frac{\epsilon}{12} \,. \tag{153}$$

We now require the following result (adapted from Lemma 65 of Ref. [25]).

Lemma A.1 (Low-weight approximation by Fourier series)**.** *Let* $\delta' \in (0,1)$*,* $\epsilon \in (0,1)$*, and* $g : \mathbb{R} \to \mathbb{C}$ *be such that* $\left|g(y) - \sum_{k=0}^{K-1} b_k y^k\right| \leq \epsilon/12$ *for all* $y \in [-1+\delta', 1-\delta']$*. Then, there exists* $\mathbf{c} := (c_{-M}, \ldots, c_M) \in \mathbb{C}^{2M+1}$ *such that*

$$\left|g(y) - \sum_{m=-M}^{M} c_m e^{i\pi m y/2}\right| \leq \epsilon/3 \tag{154}$$

for all $y \in [-1+\delta', 1-\delta']$*, where*

$$M = \max\left(2\left\lceil \frac{1}{\delta'} \log \frac{12\|\mathbf{b}\|_1}{\epsilon} \right\rceil, 0\right) , \tag{155}$$

$\mathbf{b} := (b_0, \ldots, b_{K-1}) \in \mathbb{C}^K$*, and* $\|\mathbf{c}\|_1 \leq \|\mathbf{b}\|_1$*. Moreover* $\mathbf{c}$ *can be efficiently calculated on a classical computer in time* $\mathrm{poly}(K, M, \log 1/\epsilon)$*.*

We apply this result directly to the function $g(y) = \sum_{k=0}^{\infty} b_k y^k = \sum_{k=0}^{\infty} a_k (r+\delta)^k y^k$ of Lemma 2.1 to obtain

$$\left|\sum_{k=0}^{K-1} b_k y^k - \sum_{m=-M}^{M} c_m e^{i\pi m y/2}\right| \leq \frac{\epsilon}{3} , \tag{156}$$

where $y \in \left[-1+\frac{\delta}{2(r+\delta)}, 1-\frac{\delta}{2(r+\delta)}\right]$ and $\|\mathbf{c}\|_1 \leq \|\mathbf{b}\|_1 \leq \sum_{k=0}^{\infty} |b_k| \leq B$. Let $\delta' := \frac{\delta}{2(r+\delta)}$. Then, Lemma A.1 gives

$$M = \max\left(2\left\lceil \frac{2(r+\delta)}{\delta} \log \frac{12\|\mathbf{b}\|_1}{\epsilon} \right\rceil, 0\right) = \mathcal{O}\left(\frac{r}{\delta} \log\left(\frac{B}{\epsilon}\right)\right) . \tag{157}$$

The restriction on ϵ here is that $\epsilon/3 \in (0,1)$. Changing variables back to x, the Fourier approximation of $f(x) = \sum_{k=0}^{\infty} a_k x^k$ can be expressed as

$$\tilde{f}(x) = \sum_{m=-M}^{M} c_m e^{i\pi m x/2(r+\delta)} . \tag{158}$$

This is an $\epsilon/3$-approximation to f on $[-r-\delta/2, r+\delta/2]$, i.e., $|f(x) - \tilde{f}(x)| \leq \epsilon/3$ in the domain.

To return to polynomials, we require the following two results (Lemmas 57 and 59 of Ref. [25]).

Lemma A.2 (Polynomial approximations of trigonometric functions)**.** *Let* $t \in \mathbb{R} \setminus \{0\}, \epsilon \in (0, 1/e)$*, and*

$$R = \left\lfloor \frac{1}{2} r\left(\frac{e|t|}{2}, \frac{5}{4}\epsilon\right) \right\rfloor , \tag{159}$$

where $r(u, \xi)$ *is defined as the solution to*

$$\xi = \left(\frac{u}{r}\right)^r : r \geq u , \tag{160}$$

which satisfies

$$r(u,\xi) = \Theta\left(u + \frac{\log(1/\xi)}{\log(e + \log(1/\xi)/u)}\right) = \mathcal{O}\left(u + \log\frac{1}{\xi}\right) \tag{161}$$

for all $u > 0$ and $\xi \in (0,1)$. Then, for all $x \in [-1,1]$,

$$\left|\cos(tx) - J_0(t) + 2\sum_{k=1}^{R}(-1)^k J_{2k}(t)T_{2k}(x)\right|_{[-1,1]} \leq \epsilon\ , \tag{162}$$

$$\left|\sin(tx) - 2\sum_{k=0}^{R}(-1)^k J_{2k+1}(t)T_{2k+1}(x)\right|_{[-1,1]} \leq \epsilon\ , \tag{163}$$

where $J_m(t)$ are the Bessel functions of the first kind and $T_m(x)$ are the Chebyshev polynomials of the first kind.

Hence, this result shows that $\cos(tx)$ and $\sin(tx)$ can be approximated by polynomials of even degree up to $2R$ and odd degree up to $2R+1$, respectively.

Corollary A.3. *There exist polynomials $P_{\sin}, P_{\cos}$ of degree*

$$\mathcal{O}\left(t + \log\frac{1}{\epsilon}\right) \tag{164}$$

such that for all $x \in [-1,1]$,

$$|\cos(tx) - P_{\cos}(x)|_{[-1,1]} \leq \epsilon, \quad |\sin(tx) - P_{\sin}(x)|_{[-1,1]} \leq \epsilon \tag{165}$$

and

$$|P_{\cos}(x)|_{[-1,1]} \leq 1, \quad |P_{\sin}(x)|_{[-1,1]} \leq 1 \tag{166}$$

for $\epsilon \in (0, 2/e)$.

Proof. We rescale the polynomial approximation by $\frac{1}{1+\epsilon}$ to ensure it is confined to magnitude 1. By the triangle inequality, for any function $h(x)$ that is ϵ-approximated by $P(x)$,

$$\left|\frac{1}{1+\epsilon}P(x) - h(x)\right|_{[-1,1]} \leq \frac{1}{1+\epsilon}|P(x) - h(x)|_{[-1,1]} + \frac{\epsilon}{1+\epsilon}|h(x)|_{[-1,1]} \leq \frac{2\epsilon}{1+\epsilon} < 2\epsilon\ . \tag{167}$$

Hence, this rescaling only takes ϵ to 2ϵ. Accordingly, we rescale the bounds of ϵ to $(0, 2/e)$. □

We need to approximate $\sin(\pi m x/2(r+\delta))$ and $\cos(\pi m x/2(r+\delta))$, taking

$$t = \frac{\pi m}{2(r+\delta)} = \mathcal{O}\left(\frac{1}{\delta}\log\frac{B}{\epsilon}\right) \tag{168}$$

and applying the above lemma. Hence, a normalized polynomial approximation $P_{\tilde{f}}(x)$ to $\tilde{f}(x)$ with error $\|P_{\tilde{f}}(x) - \tilde{f}(x)\|_{[-1,1]} \leq \xi$ has degree

$$\mathcal{O}\left(\frac{1}{\delta}\log\frac{B}{\epsilon} + \log\frac{1}{\xi}\right)\ . \tag{169}$$

We choose $\xi = \epsilon/3B \leq 2/e$ to give a degree of $\mathcal{O}\left(\frac{1}{\delta}\log\frac{B}{\epsilon}\right)$ and $\epsilon \leq 6B/e$. Summarizing the properties of $P_{\tilde{f}}$, we have

$$|P_{\tilde{f}}(x)\|_{[-1,1]} \leq B, \quad |P_{\tilde{f}}(x) - \tilde{f}(x)|_{[-1,1]} \leq \frac{\epsilon}{3}, \quad |P_{\tilde{f}}(x) - f(x)|_{[-r-\delta/2, r+\delta/2]} \leq \frac{2\epsilon}{3} \tag{170}$$

where the first property follows from $\|c\|_1 \leq B$; the second property follows from the $\epsilon/3B$-approximation of each trigonometric function and the fact $\|c\|_1 \leq B$; and the third property adds the $\epsilon/3$ error of the original Fourier approximation.

Since B may exceed 1, the first condition does not guarantee that $P_{\tilde{f}}$ is properly normalized to use as a QSVT polynomial. To obtain the final polynomial, we need a polynomial approximation of the rectangle function (Lemma 29 of Ref. [25]).

Lemma A.4 (Polynomial approximation of the rectangle function)**.** *Let $\delta, \epsilon \in (0, 1/2)$ and $t \in [-1, 1]$. There exists an even polynomial P of degree $\mathcal{O}\left(\frac{1}{\delta}\log\frac{1}{\epsilon}\right)$ such that*

$$|P(x)|_{[-1,1]\setminus[-t-\delta,t+\delta]} \leq \epsilon, \qquad |P(x)-1|_{[-t+\delta,t-\delta]} \leq \epsilon, \qquad |P(x)|_{[-1,1]} \leq 1\ . \tag{171}$$

We define polynomial Q as the product of $P_{\tilde{f}}$ and a polynomial approximation P_{rect} of a rectangle function that satisfies

$$|P_{\text{rect}}(x)-1|_{[-r,r]} \leq \frac{\epsilon}{3B}, \quad |P_{\text{rect}}(x)|_{[-1,1]\setminus[-r-\delta/2,r+\delta/2]} \leq \frac{\epsilon}{3B}, \quad |P_{\text{rect}}(x)|_{[-1,1]} \leq 1\ . \tag{172}$$

The rectangle function must have $\epsilon/3B < 1/2$, i.e., $\epsilon < 3B/2$. This is the tightest condition on ϵ, providing the condition in the final theorem statement. Since the rectangle function is $\epsilon/3B$-close to 1 on $[-r, r]$, the total error in $[-r, r]$ is $(\epsilon/3B + \epsilon/3B)(B) + \epsilon/3 = \epsilon$. The three terms come from the rectangle function error, the polynomial approximation of the Fourier expansion expansion, and the Fourier expansion respectively. Hence, Q satisfies

$$|Q(x) - f(x)|_{[-r,r]} \leq \epsilon\ . \tag{173}$$

Due to the rectangle function approximating 0 outside $[-r-\delta/2, r+\delta/2]$, we similarly have

$$|Q(x)|_{[-1,1]\setminus[-r-\delta/2,r+\delta/2]} \leq \epsilon\ . \tag{174}$$

Because the rectangle function is at most 1, and $P_{\tilde{f}}$ approximates the original function f over $[-r-\delta/2, r+\delta/2]$, we have

$$|Q(x)|_{[-1,1]} \leq \epsilon + |f(x)|_{[-r-\delta/2,r+\delta/2]}\ . \tag{175}$$

Finally, the degree of Q is $\mathcal{O}\left(\frac{1}{\delta}\log\frac{B}{\epsilon}\right)$ for $\epsilon \in (0, 3B/2)$, completing the Lemma.

□

B Quantum search task for lower bounds

To show lower bounds for both gap-amplifiable Hamiltonians and generic Hamiltonians, we shall reduce to lower bounds for a particular search task. We define the following search problem; heuristically, it consists of K instances of a typical Grover search, with one marked element per instance.

Problem 1 (K-partition quantum search)**.** *We denote the K-partition quantum search over a length-KN bitstring by* $\mathrm{SEARCH}_{K,N}$*, where $N = 2^n$, and $K \geq 1$, $n \geq 1$ are integer. Each partition $k \in [K]$ consists of N bits, where one bit is marked $f_k(x_k) = 1$ for some $x_k \in \{0, \ldots, N-1\}$, and all other bits $x \neq x_k$ are marked $f_k(x) = 0$. Access to the functions $f_1, \ldots, f_K : \{0,1\}^N \to \{0,1\}$ is provided through a phase oracle O_f that implements $|k\rangle |x\rangle \mapsto (-1)^{f_k(x)} |k\rangle |x\rangle$. The K-partition quantum search problem is to return the marked element in each partition, i.e., $x_1, \ldots, x_K$, with probability at least $2/3$.*

Briefly, we comment on the connection between quantum search and Hamiltonian simulation, and note that the use of multiple repetitions of quantum search is critical to obtaining a meaningful lower bound. For a single instance of Grover's quantum search in dimension N, there exists an example of an n-qubit Hamiltonian with spectral gap $\Delta = \Theta(1/\sqrt{N})$ such that a quantum walk for time $t \sim 1/\Delta$ can find the marked item; see Sec. 4.3 and Ref. [40]. In our case, we want to vary different parameters independently, and the only independent parameter in that example is N. Showing a lower bound for times independent of, and larger than, $1/\Delta$ will become important when we must separate time lower bounds from error lower bounds that also depend on $1/\Delta$. By including K copies of that quantum system, each associated with an instance of quantum search of dimension N, the energy scale associated with the low-energy subspace increases with K to include multiple energy levels, while the evolution time depends solely on N. A lower bound on $\mathrm{SEARCH}_{K,N}$ is then carried to a lower bound on simulating such quantum systems.

In practice, we consider the related decision problem $(\mathrm{PARITY} \circ \mathrm{OR})_{K,N}$, which can be solved by $\mathrm{SEARCH}_{K,N}$.

Problem 2 (K-partition PARITY $\circ$ OR decision problem)**.** *We denote the K-partition decision problem over a length-KN bitstring by* $(\mathrm{PARITY} \circ \mathrm{OR})_{K,N}$*, where $N = 2^n$, and $K \geq 1$, $n \geq 1$ are integer. Each partition $k \in [K]$ consists of N bits, where one bit is marked $f_k(x_k) = 1$ for some $x_k \in \{0, \ldots, N-1\}$, and all other bits $x \neq x_k$ are marked $f_k(x) = 0$. Access to the functions $f_1, \ldots, f_K : \{0,1\}^N \to \{0,1\}$ is provided through a phase oracle O_f that implements $|k\rangle |x\rangle \mapsto (-1)^{f_k(x)} |k\rangle |x\rangle$. Let $g : \{0,1\}^N \to \{0,1\}$ be such that $g(x)$, $x \in \{0,1\}^N$, is the OR of the first $N/2$ bits of x. The K-partition* PARITY $\circ$ OR *decision problem is to return* YES *if* $\mathrm{PARITY}(g(x_1), \ldots, g(x_K)) = 1$ *and* NO *otherwise, with probability at least $2/3$.*

Known lower bounds for PARITY and for OR are sufficient to place a lower bound on $(\mathrm{PARITY} \circ \mathrm{OR})_{K,N}$ using the adversary method.

Lemma B.1 (Query complexity of K-partition PARITY $\circ$ OR)**.** *Deciding* $(\mathrm{PARITY} \circ \mathrm{OR})_{K,N}$ *requires $\Omega(K\sqrt{N})$ queries to O_f.*

Proof. Consider the Boolean function $\mathrm{PARITY}(z)$, where $z \in \{0,1\}^K$, and define the Boolean function $g(x) : \{0,1\}^N \to \{0,1\}$ to be the OR of the first $N/2$ bits of x. The function composition $h = \mathrm{PARITY} \circ g$ evaluates $\mathrm{PARITY}(g(x_1), \ldots, g(x_K))$, which is precisely $(\mathrm{PARITY} \circ \mathrm{OR})_{K,N}$. The quantum adversary method provides lower bounds on the query complexity of PARITY and g as $\mathrm{ADV}(\mathrm{PARITY}) = c_1 K$ and $\mathrm{ADV}(g) = c_2\sqrt{N}$,

for constants $c_1 > 0$ and $c_2 > 0$. References [30, 42] imply a lower bound on the query complexity of the composition of Boolean functions as the product of corresponding lower bounds. Hence,

$$\mathrm{ADV}(\mathrm{PARITY} \circ g) = cK\sqrt{N} \tag{176}$$

for $c = c_1 c_2$. Since our phase oracle is equivalent to a bit-flip oracle, the implication is that solving $(\mathrm{PARITY} \circ \mathrm{OR})_{K,N}$ requires $\Omega(K\sqrt{N})$ queries to O_f. □

Next, we use the lower bound on the query complexity of $(\mathrm{PARITY} \circ \mathrm{OR})_{K,N}$ to show a lower bound on preparing a particular state of K quantum systems that emerges naturally from time evolution. Since the search within each partition is over N elements, we rewrite the oracle to implement $f_1, \ldots, f_K : \{0,1\}^n \to \{0,1\}$ for $N = 2^n$. Although one-hot encoding each marked element was convenient when writing the problem in terms of Boolean functions above, this formulation is more useful in practice as it reduces the overhead in the second register presented to the oracle.

Lemma B.2 (Query complexity of state preparation)**.** *Let $f_1, \ldots, f_K : \{0,1\}^n \to \{0,1\}$ be functions, where $N = 2^n$, $K \geq 1$, and $n \geq 1$ are integer. For each partition $k \in [K]$, assume that $f_k(x_k) = 1$ for some marked item $x_k \in \{0,1\}^n$, and $f_k(x) = 0$ otherwise. Then, at least $\Omega(K\sqrt{N})$ queries to the K-partition Grover oracle O_f that implements $|k\rangle|x\rangle \mapsto (-1)^{f_k(x)}|k\rangle|x\rangle$ are required to prepare a Kn-qubit state*

$$|\phi\rangle = \bigotimes_{k=1}^{K} |\phi_k\rangle \ , \tag{177}$$

satisfying $|\langle\phi_k|x_k\rangle|^2 \geq \alpha^2$ for constant $\alpha \in (0,1)$, where $|x_k\rangle$ is the n-qubit state corresponding to the marked item x_k in the k^{th} partition. Here, $|\phi_k\rangle \in \mathbb{C}^N$ and $|\phi\rangle \in \mathbb{C}^{N^k}$.

Proof. Observe that with probability at least $2/3$ and with $\mathcal{O}(K)$ queries to the oracle O_f, concentration implies that a constant number of copies of the state $|\phi\rangle$ are sufficient to successfully find βK elements $S_{\beta K} = \{x_{i_1}, \ldots, x_{i_{\beta K}}\}$ for $i_1, \ldots, i_{\beta K} \in [K]$ and any constant β. This follows from measuring the copies of $|\phi\rangle$ in the computational basis and verifying each measurement outcome with O_f for the identification of successful search outcomes. We denote the number of queries to O_f required to obtain $S_{\beta K}$ by Q.

Consider the remaining unsuccessful $(1-\beta)K$ search problems. By Ref. [13], it suffices to perform $(1-\beta)K$ independent Grover searches using a total of $\frac{\pi}{4}(1-\beta)K\sqrt{N}$ queries to the oracle to solve these instances with probability at least $2/3$. Hence, preparing βK copies of $|\phi\rangle$ in addition to the Grover searches would allow us to solve Problem 2. By Lemma B.1, there exists some constant c such that at least $cK\sqrt{N}$ queries are required to decide $(\mathrm{PARITY} \circ \mathrm{OR})_{K,N}$, implying that

$$Q + \frac{\pi}{4}(1-\beta)K\sqrt{N} \geq cK\sqrt{N} \ . \tag{178}$$

Choosing constant β such that $1 > \beta > 1 - \frac{4c}{\pi}$ is always possible since $c > 0$ (and we can assume $4c < \pi$). This implies that $Q = \Omega(K\sqrt{N})$ is necessary to satisfy the lower bound of Lemma B.1; i.e., $\Omega(K\sqrt{N})$ queries to O_f are required to prepare $|\phi\rangle$. □

Lemma B.2 directly implies Lemma 4.1, which is the main result used to prove the lower bounds in Secs. 4 and 5.

C Proofs of Lemmas 5.4 and 5.5

We begin with the proof of Lemma 5.4. Following Appendix A of Ref. [49], let $\tilde{H}_x$ be a Hamiltonian acting on $\mathbb{C}^N$ defined over an expander graph $G = (N, d, \eta)$ with $N = 2^n$ vertices, constant degree $d = \mathcal{O}(1)$, and second-largest eigenvalue ηd of the adjacency matrix, taking $\eta \leq 1/2$. Here, $x \in \{0,1\}^n$ refers to a marked item in Grover's search. The matrix elements of $\tilde{H}_x$ are given by

$$\langle y| \tilde{H}_x |z\rangle = \begin{cases} \frac{1}{N-1} & y = z = x \\ -\frac{1}{d\sqrt{N-1}} & \{y,z\} \in E \text{ and } y = x \text{ or } y = z \\ -\frac{1}{d} & \{y,z\} \in E \text{ and } x \neq y \neq z \neq x \\ 1 & y = z \neq x \\ 0 & \text{otherwise} \end{cases} , \tag{179}$$

for edge set E. By construction, this Hamiltonian is $(d+1)$-sparse; we must convert it into a Hamiltonian that is gap-amplifiable. For each ordered edge (y, z) with $y \leq z$, we observe that

$$\tilde{H}_x = \sum_{(y,z)\in E} (c_y |y\rangle - c_z |z\rangle)(c_y \langle y| - c_z \langle z|) \tag{180}$$

for

$$c_v = \begin{cases} \frac{1}{\sqrt{d(N-1)}} & v = x \\ \frac{1}{\sqrt{d}} & \text{else} \end{cases} . \tag{181}$$

To rewrite the Hamiltonian as a sum of projectors or PSD terms, we rescale the above terms to obtain

$$H_x = \sum_{(y,z)\in E} \frac{(c_y |y\rangle - c_z |z\rangle)(c_y \langle y| - c_z \langle z|)}{c_y^2 + c_z^2} = \sum_{(y,z)\in E} \Pi_{(y,z)} . \tag{182}$$

Note that $(\Pi_{(y,z)})^2 = \Pi_{(y,z)}$. Hence, H_x remains $(d+1)$-sparse; moreover, as shown in Ref. [49], H_x has a spectral gap of at least $\frac{1}{4(N-1)}$ and $\|H_x\|_{\max} = \mathcal{O}(\|H_x\|) = \mathcal{O}(1)$.

In this decomposition over projectors $\Pi_{(y,z)}$, the number of terms in the frustration-free basis scales extensively with N. To ensure efficient access, we color the graph with $\mathcal{O}(d) = \mathcal{O}(1)$ colors, grouping together projectors that do not share a vertex using their orthogonality. This decomposition produces $H_x = \sum_{l=1}^{L} \Pi_{l,x}$ where the number of terms $L = \mathcal{O}(1)$ is constant, and each $\Pi_{l,x}$ is a projector that depends on x. Note that this coloring also implies an upper bound on $\|H_x\|$: since there are at most $d+1$ colors and hence $d+1$ projectors, the largest eigenvalue is at most $d+1 = \mathcal{O}(1)$. □

Next we provide the proof of Lemma 5.5.

Let $|\perp\rangle = \frac{1}{\sqrt{N-1}} \sum_{y\neq x} |y\rangle$. The ground state of H_x has zero energy and is given by

$$|\phi_0\rangle = \frac{1}{\sqrt{2}} |x\rangle + \frac{1}{\sqrt{2}} |\perp\rangle . \tag{183}$$

Hence, $|\langle\phi_0|s\rangle| > 1/\sqrt{2}$. To show the uniform superposition $|s\rangle$ is low-energy in H_x, we evaluate powers of $\langle s| H_x^k |s\rangle$. First, we write H_x explicitly, using $(a,b) \in E$ to denote ordered edges (a,b) in the edge set and $\mathcal{N}(x)$ to denote the edges with x at one vertex:

$$H_x = \sum_{y\in\mathcal{N}(x)} \left(\frac{1}{d} + \frac{1}{d(N-1)}\right)^{-1} \left(\frac{1}{\sqrt{d(N-1)}}|x\rangle - \frac{1}{\sqrt{d}}|y\rangle\right)\left(\frac{1}{\sqrt{d(N-1)}}\langle x| - \frac{1}{\sqrt{d}}\langle y|\right) \tag{184}$$

$$+ \sum_{\substack{(y,z)\in E\\ y\neq x, z\neq x}} \left(\frac{1}{d}+\frac{1}{d}\right)^{-1}\left(\frac{1}{\sqrt{d}}|y\rangle - \frac{1}{\sqrt{d}}|z\rangle\right)\left(\frac{1}{\sqrt{d}}\langle y| - \frac{1}{\sqrt{d}}\langle z|\right) \tag{185}$$

$$= \frac{1}{N}\sum_{y\in\mathcal{N}(x)} \left(|x\rangle\langle x| + (N-1)|y\rangle\langle y| - \sqrt{N-1}(|x\rangle\langle y| + |y\rangle\langle x|)\right) \tag{186}$$

$$+ \frac{1}{2}\sum_{\substack{(y,z)\in E\\ y\neq x, z\neq x}} (|y\rangle\langle y| + |z\rangle\langle z| - |y\rangle\langle z| - |z\rangle\langle y|)\ . \tag{187}$$

In general, we find that states of the form $\alpha|x\rangle + \beta|\perp\rangle$ have energy $d(\alpha-\beta)^2/N$. We introduce the states

$$|\psi_1\rangle = \sqrt{\frac{1+\sqrt{\frac{1}{N}}}{2}}|x\rangle + \sqrt{\frac{1-\sqrt{\frac{1}{N}}}{2}}|\perp\rangle, \quad |\psi_2\rangle = \frac{1}{\sqrt{2\left(1+\frac{1}{\sqrt{N}-1}\right)}}|x\rangle - \sqrt{\frac{1+\sqrt{\frac{1}{N}}}{2}}|\perp\rangle\ , \tag{188}$$

which produce the convenient decomposition

$$|s\rangle = \sqrt{\frac{1+\frac{1}{\sqrt{N}}}{2}}|\psi_1\rangle - \sqrt{\frac{1-\frac{1}{\sqrt{N}}}{2}}|\psi_2\rangle \tag{189}$$

and have energies

$$\langle\psi_1| H_x |\psi_1\rangle = \mathcal{O}(1/N^2) \tag{190}$$

$$\langle\psi_2| H_x |\psi_2\rangle = \mathcal{O}(1/N)\ . \tag{191}$$

By explicit computation, $|\langle\psi_1|\phi_0\rangle|^2 = 1 - \mathcal{O}(1/N)$ and $|\langle\psi_2|\phi_0\rangle| = \mathcal{O}(1/N)$. By Markov's inequality, for decomposition

$$|\psi_2\rangle = \sum_j d_j |\phi_j\rangle\ , \tag{192}$$

we have that

$$\sum_{j\,:\,\lambda_j \geq \frac{c}{N}} |d_j|^2 = \mathcal{O}\left(\frac{1}{c}\right) \tag{193}$$

for any $c > 0$. Consequently, for any choice of (arbitrarily small) constant $\delta > 0$, there exists a choice of $\Delta = \mathcal{O}(1/N)$ such that for projector Π_Δ onto the low-energy subspace, the state $|\psi\rangle = \Pi_\Delta|s\rangle / \|\Pi_\Delta|\psi\rangle\|$ has overlap $|\langle\psi|s\rangle| \geq 1-\delta$. Moreover, due to the support of $|s\rangle$ on $|\psi_1\rangle$, we can choose δ such that $|\psi\rangle$ satisfies $|\langle\psi|\phi_0\rangle| > 1/\sqrt{2} - \delta > 1/2$. □

D Proof of Lemma 6.3

For convenience, we reproduce the definitions and lemma statement before proving the result.

The clock Hamiltonian H encodes a quantum circuit $\mathcal{U} = U_G U_{G-1} \cdots U_1$ acting on n qubits. We pad the circuit with identity gates, introducing unitaries V_x given by

$$V_x = \begin{cases} \mathbb{1}_N & \text{if } \ 1 \leq x \leq c_1 G \\ U_{x-c_1 G} & \text{if } \ c_1 G + 1 \leq x \leq (c_1+1)G \\ \mathbb{1}_N & \text{if } \ (c_1+1)G + 1 \leq x \leq (c_1+c_2)G \, . \end{cases} \tag{194}$$

There are $G' = (c_1 + c_2)G$ unitaries V_x. The clock Hamiltonian is

$$H = \sum_{x=1}^{G'} \left(-V_x \otimes |x\rangle\langle x-1| - V_x^\dagger \otimes |x-1\rangle\langle x| + \mathbb{1}_N \otimes |x\rangle\langle x| + \mathbb{1}_N \otimes |x-1\rangle\langle x-1| \right) \, , \tag{195}$$

where $G' = (c_1 + c_2)G$. We assume periodic boundary conditions, i.e., $x = 0$ is equivalent to $x = G'$. The initial state is supported only on basis states corresponding to initial padding with the identities $\mathbb{1}_N$, that is,

$$|\psi\rangle = |0\rangle^{\otimes n} \otimes \sum_{x=0}^{c_1 G} \psi(x) \, |x\rangle \ . \tag{196}$$

The amplitudes are those of a Gaussian wavepacket centered at site $x_0 = (c_1/2)G$ and with initial rightwards momentum $p_0 > 0$, such that

$$\psi(x) = \eta \exp\left[-\frac{(x-x_0)^2}{2\sigma^2} + i p_0 x \right], \quad \eta = \left(\sum_{x=0}^{c_1 G} \exp\left[-\frac{(x-x_0)^2}{\sigma^2} \right] \right)^{-1/2} \, . \tag{197}$$

Without loss of generality, we choose to start the computation by $\mathcal{U}$ in state $|0\rangle^{\otimes n}$. We seek to prove the following result from Sec. 6.2.

Lemma 6.3 (Time evolution of initial Gaussian-like state under the clock Hamiltonian)**.** *Let H be the clock Hamiltonian for the G-gate unitary $\mathcal{U} = U_G U_{G-1} \ldots U_1$ acting on n qubits as defined in Eq. (195); it contains $c_1 G$ initial identity gates and $(c_2 - c_1 - 1)G$ final identity gates, for even integers c_1, c_2. Let $|\psi\rangle$ be the state defined in Eq. (196), with amplitudes determined by Eqs. (197), and centered at site $x_0 = (c_1/2)G$. Fix the width of the wavepacket to $\sigma = \hat{\sigma} x_0$ and the momentum to $p_0 = \hat{p}_0 / x_0$ for positive constants $\hat{\sigma}$, and $\hat{p}_0$. Consider the time-evolved state $e^{-itH} |\psi\rangle$ for time $t = \hat{t} x_0^2 = \mathcal{O}(G^2)$ for constant $\hat{t} > 0$. Then, for any $G \geq 1$, there exists a choice of constants c_1, c_2, $\hat{\sigma}$, $\hat{p}_0$, and $\hat{t}$, such that*

$$\frac{1}{2} \left\| \mathcal{U} |0\rangle\langle 0|^{\otimes n} \mathcal{U}^\dagger - \Pi_n e^{-itH} |\psi\rangle\langle\psi| e^{itH} \Pi_n \right\|_1 \leq \frac{1}{4} \, , \tag{198}$$

where Π_n is the projector onto the space of the n qubits where the computation occurs, and $\frac{1}{2}\|\cdot\|_1$ indicates the trace distance.

Proof. On the first register of n qubits where the computation occurs, we introduce an initial state $|h_0\rangle := |0\rangle^{\otimes n}$ and subsequent states $|h_1\rangle$ through $|h_{G'}\rangle$ given by

$$|h_i\rangle = V_i V_{i-1} \cdots V_1 \, |h_0\rangle \ . \tag{199}$$

Here, $G' = (c_1 + c_2)G$, for some even constants $c_1 > 0$ and $c_2 > 0$. Due to its connection to a "tight-binding model", the clock Hamiltonian admits a convenient diagonalization through the Fourier transform. In the subspace fixed by $|h_0\rangle$, the eigenstates are

$$|\varphi_k\rangle = \frac{1}{\sqrt{G'}} \sum_{j=0}^{G'-1} e^{i2\pi jk/G'} |h_j\rangle |j\rangle \tag{200}$$

where $0 \leq k \leq G' - 1$. The corresponding eigenvalues are $\gamma_k = 2(1 - \cos(2\pi k/G'))$. These satisfy $0 \leq \gamma_k \leq 2$, showing that H is PSD. The other eigenspaces are degenerate and fixed by other choices of $|h_0\rangle$.

Since $c_1 G = 2x_0$, we can express the initial state $|\psi(0)\rangle := |\psi\rangle$ of Eq. (196) in terms of the eigenstates as

$$|\psi(0)\rangle = \frac{\eta}{\sqrt{G'}} \sum_{x=0}^{2x_0} \sum_{k=0}^{G'-1} \exp\left[-\frac{(x-x_0)^2}{2(\hat{\sigma}x_0)^2} + i\frac{\hat{p}_0}{x_0}x\right] e^{-i2\pi xk/G'} |\varphi_k\rangle \ . \tag{201}$$

Time evolution under H for time t produces

$$|\psi(t)\rangle = \frac{\eta}{\sqrt{G'}} \sum_{x=0}^{2x_0} \sum_{k=0}^{G'-1} \exp\left[-\frac{(x-x_0)^2}{2(\hat{\sigma}x_0)^2} + i\frac{\hat{p}_0}{x_0}x\right] e^{-it\gamma_k} e^{-i2\pi xk/G'} |\varphi_k\rangle \tag{202}$$

$$= \frac{\eta}{G'} \sum_{x=0}^{2x_0} \sum_{k=0}^{G'-1} \sum_{j=0}^{G'-1} \exp\left[-\frac{(x-x_0)^2}{2(\hat{\sigma}x_0)^2} + i\frac{\hat{p}_0}{x_0}x\right] e^{-it2(1-\cos(2\pi k/G'))} e^{i2\pi(j-x)k/G'} |h_j\rangle |j\rangle \ . \tag{203}$$

We will ultimately use a state $|\psi_\Delta\rangle$ that is entirely supported in a low-energy subspace Δ of H; all eigenvalues γ_k in this subspace will satisfy

$$\gamma_k \leq \Delta = \mathcal{O}(\hat{p}_0^2/G^2) \ , \tag{204}$$

implying that all $k = \mathcal{O}((c_1 + c_2)\hat{p}_0)$ supported by $|\psi_\Delta\rangle$ are also bounded by a constant. In this low-energy subspace, the eigenvalues of H can be approximated as

$$\gamma_k = \left(\frac{2\pi k}{(c_1 + c_2)G}\right)^2 + \mathcal{O}(G^{-4}) \ . \tag{205}$$

We choose evolution time $t = \hat{t}x_0^2$ for constant $\hat{t}$ and evaluate the support of the time-evolved state on a particular basis state $|h_j\rangle |j\rangle$. We replace γ_k in the time evolution with the small-k approximation in Eq. (205) and keep track of the additional error introduced by the approximation. Due to the choice of simulation time $t = \Theta(G^2)$, the $\mathcal{O}(G^{-4})$ error in the approximation of the eigenvalue produces $\mathcal{O}(G^{-2})$ error in the exponent, i.e.,

$$\langle h_j| \langle j|\psi(t)\rangle = \frac{\eta}{(c_1 + c_2)G} \sum_{x=0}^{c_1 G} \sum_{k=0}^{G'-1} \exp\Bigg[-\frac{2}{(\hat{\sigma}c_1)^2} \frac{(x - c_1 G/2)^2}{G^2} + i\left(\frac{2\hat{p}_0}{c_1}\frac{x}{G} - \hat{t}\left(\frac{\pi k}{1 + c_2/c_1}\right)^2 + \frac{2\pi k}{c_1 + c_2}\frac{j-x}{G} + \mathcal{O}(G^{-2})\right)\Bigg] \ . \tag{206}$$

Here, we used $G' = (c_1 + c_2)G$ and $x_0 = c_1 G/2$. Equivalently, we can offset the sum to be

$$\langle h_j | \langle j|\psi(t)\rangle = \frac{1}{c_1+c_2}\frac{\eta}{G}\sum_{x=-c_1G/2}^{c_1G/2-1}\sum_{k=0}^{G'-1}\exp\left[-\frac{2}{(\hat{\sigma}c_1)^2}\frac{x^2}{G^2}+\mathcal{O}(G^{-2})\right] \times \exp\left[i\left(\hat{p}_0\left(\frac{2}{c_1}\frac{x}{G}+1\right)-\hat{t}\left(\frac{\pi k}{1+c_2/c_1}\right)^2+\frac{2\pi k}{c_1+c_2}\left(\frac{j-x}{G}-\frac{c_1}{2}\right)\right)\right] \tag{207}$$

$$= \frac{1}{c_1+c_2}\frac{\eta}{G}\sum_{k=0}^{G'-1}\exp\left[-\frac{\hat{\sigma}^2}{2}(\hat{p}_0-\alpha k)^2+i\left(\hat{p}_0-(\alpha k)^2\hat{t}+\alpha k\left(\frac{2}{c_1}\frac{j}{G}-1\right)\right)\right] \times \sum_{x=-c_1G/2}^{c_1G/2-1}\exp\left[-\frac{1}{2}\left(\frac{2}{\hat{\sigma}c_1G}x-i\hat{\sigma}\left(\hat{p}_0-\alpha k\right)\right)^2+\mathcal{O}(G^{-2})\right], \tag{208}$$

where we defined for convenience a constant

$$\alpha := \frac{\pi}{1+c_2/c_1}\,. \tag{209}$$

For fixed k, the last sum is of the form

$$\sum_{x=-S/2}^{S/2-1}\exp\left[-\frac{1}{2}\left(\frac{x}{S\hat{\sigma}/2}-i\beta\right)^2\right] \tag{210}$$

for $\beta = \hat{\sigma}(\hat{p}_0 - \alpha k)$ and $S = c_1 G$. We will expand its range at the cost of additional error ϵ_{trunc}, i.e.,

$$\sum_{x=-S/2}^{S/2-1}\exp\left[-\frac{1}{2}\left(\frac{x}{S\hat{\sigma}/2}-i\beta\right)^2\right] = \sum_{x=-\infty}^{\infty}\exp\left[-\frac{1}{2}\left(\frac{x}{S\hat{\sigma}/2}-i\beta\right)^2\right]+2\epsilon_{\text{trunc}}\,. \tag{211}$$

For small $\hat{\sigma}$, the truncation error is bounded by

$$\epsilon_{\text{trunc}} = \left|\sum_{x=S/2}^{\infty}\exp\left[-\frac{1}{2}\left(\frac{x}{S\hat{\sigma}/2}-i\beta\right)^2\right]\right| \tag{212}$$

$$\leq \sum_{x=S/2}^{\infty}\exp\left[-\frac{1}{2}\left(\frac{x^2}{(S\hat{\sigma}/2)^2}-\beta^2\right)\right] \tag{213}$$

$$\leq \exp\left[\frac{\beta^2}{2}\right]\int_{S/2-1}^{\infty}\exp\left[-\frac{1}{2}\frac{x^2}{(S\hat{\sigma}/2)^2}\right]dx \tag{214}$$

$$\leq \frac{\sqrt{2\pi}}{4}S\hat{\sigma}\,\text{erfc}\left(\frac{S-2}{\sqrt{2}S\hat{\sigma}}\right)\exp\left[\frac{\beta^2}{2}\right] \tag{215}$$

$$= \mathcal{O}\left(S\hat{\sigma}^2 e^{-1/8\hat{\sigma}^2}\right)\,, \tag{216}$$

where after evaluating the integral we used the fact that β is proportional to $\hat{\sigma}$. Applying Poisson summation, we have

$$\sum_{x=-\infty}^{\infty}\exp\left[-\frac{1}{2}\left(\frac{x}{S\hat{\sigma}/2}-i\beta\right)^2\right] = \sqrt{\frac{\pi}{2}}S\hat{\sigma}e^{-\beta^2/2}\sum_{\omega=-\infty}^{\infty}\exp\left[-\frac{1}{2}\left(\pi\omega S\hat{\sigma}-\beta\right)^2\right] \tag{217}$$

$$= \sqrt{\frac{\pi}{2}}c_1 G\hat{\sigma}+\mathcal{O}(\epsilon_{|\omega|>0})\,, \tag{218}$$

where

$$\epsilon_{|\omega|>0} \leq \sqrt{\frac{\pi}{2}}\hat{\sigma}Se^{-\beta^2/2}\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}\left(\pi\omega\hat{\sigma}S-\beta\right)^2\right]d\omega \tag{219}$$

$$\leq e^{-\beta^2/2} \tag{220}$$

$$\leq 1\,. \tag{221}$$

Note that above, we took β to be constant compared to G, despite its dependence on k. For the purposes of our proof, this holds because the state $|\psi_\Delta\rangle$ has arbitrarily large overlap with $|\psi(0)\rangle$, and $|\psi_\Delta\rangle$ is confined to a low-energy subspace where k is at most constant-sized with respect to G. The time-evolved state thus satisfies

$$\langle h_j|\langle j|\psi(t)\rangle = \sqrt{\frac{\pi}{2}}\frac{\hat{\sigma}\eta}{1+c_2/c_1}\left(1+\mathcal{O}\left(\hat{\sigma}e^{-1/8\hat{\sigma}^2}\right)\right)$$
$$\times\sum_{k=0}^{G'-1}\exp\left[-\frac{\hat{\sigma}^2}{2}\left(\hat{p}_0-\alpha k\right)^2+i\left(\hat{p}_0-(\alpha k)^2\hat{t}+\alpha k\left(\frac{2}{c_1}\frac{j}{G}-1\right)+\mathcal{O}(G^{-2})\right)\right] \tag{222}$$

$$=\sqrt{\frac{1}{2\pi}}\alpha\hat{\sigma}\eta\left(1+\mathcal{O}\left(\hat{\sigma}e^{-1/8\hat{\sigma}^2}\right)\right)\exp\left[\frac{1}{2}\left(\hat{p}_0(2i-\hat{p}_0\hat{\sigma}^2)-\left(\frac{\alpha(\hat{j}-i\hat{p}_0\hat{\sigma}^2)}{\gamma}\right)^2\right)\right]$$
$$\times\sum_{k=-G'/2}^{G'/2-1}\exp\left[-\frac{\gamma^2}{2}\left(k-\left(\frac{\alpha(\hat{p}_0\hat{\sigma}^2+i\hat{j})}{\gamma^2}-\frac{G'}{2}\right)\right)^2+\mathcal{O}(G^{-2})\right] \tag{223}$$

for

$$\gamma=\alpha\left(\hat{\sigma}^2+2i\hat{t}\right)^{1/2},\quad \hat{j}=\frac{2}{c_1}\frac{j}{G}-1\,. \tag{224}$$

To evaluate the last sum, we introduce constant-size variables

$$\gamma_1=\hat{\sigma}^2\alpha^2,\quad \gamma_2=2\hat{t}\alpha^2 \tag{225}$$

as well as

$$\mu_1=\frac{\alpha^3(\hat{p}_0\hat{\sigma}^4+2\hat{j}\hat{t})}{\gamma_1^2+\gamma_2^2}-\frac{G'}{2},\quad \mu_2=\frac{\alpha^3\hat{\sigma}^2(\hat{j}-2\hat{p}_0\hat{t})}{\gamma_1^2+\gamma_2^2},\quad \mu=\mu_1-\frac{\gamma_2}{\gamma_1}\mu_2\,. \tag{226}$$

The last sum can now be rewritten as

$$\sum_{k=-G'/2}^{G'/2-1}\exp\left[-\frac{\gamma_1+i\gamma_2}{2}\left(k-(\mu_1+i\mu_2)\right)^2\right]\,. \tag{227}$$

Similarly to before, we expand the limits of the sum at the cost of error

$$\epsilon'_{\text{trunc}}=\left|\sum_{x=G'/2}^{\infty}\exp\left[-\frac{\gamma_1+i\gamma_2}{2}\left(k-(\mu_1+i\mu_2)\right)^2\right]\right| \tag{228}$$

$$\leq\exp\left[\frac{\mu_2^2(\gamma_1^2+\gamma_2^2)}{2\gamma_1}\right]\int_{G'/2-1}^{\infty}\exp\left[-\frac{\gamma_1}{2}\left(k-\mu\right)^2\right]dk \tag{229}$$

$$\leq\sqrt{\frac{\pi}{2\gamma_1}}\exp\left[\frac{\mu_2^2(\gamma_1^2+\gamma_2^2)}{2\gamma_1}\right]\operatorname{erfc}\left(\sqrt{\frac{\gamma_1}{2}}\left(\frac{G'}{2}-(\mu+1)\right)\right) \tag{230}$$

$$=\mathcal{O}\left(\frac{1}{G}e^{-(\alpha\hat{\sigma}G)^2/2}\right)\,. \tag{231}$$

Up to additional error ϵ'_{trunc}, our sum is now approximated by

$$\sum_{k=-\infty}^{\infty} \exp\left[-\frac{\gamma_1+i\gamma_2}{2}\left(k-(\mu_1+i\mu_2)\right)^2\right] = \vartheta_3(0, e^{-(\gamma_1+i\gamma_2)/2}) \tag{232}$$

for Jacobi elliptic theta function ϑ_3. Note that this is independent of the site j. Moreover, since the error from ϵ'_{trunc} is dominated by the error from ϵ_{trunc}, we have

$$\begin{aligned} |\langle h_j|\langle j|\psi(t)\rangle|^2 &\propto \left|\exp\left[\frac{1}{2}\left(\hat{p}_0(2i-\hat{p}_0\hat{\sigma}^2)-\left(\frac{\alpha(\hat{j}-i\hat{p}_0\hat{\sigma}^2)}{\gamma}\right)^2\right)+\mathcal{O}(G^{-2})\right]\right|^2 \\ &\quad\times\left(1+\mathcal{O}\left(\hat{\sigma}e^{-1/8\hat{\sigma}^2}\right)\right) \qquad (233) \\ &\propto \exp\left[-\frac{1}{2}\frac{\hat{\sigma}^2}{\hat{\sigma}^4+4\hat{t}^2}\left(\frac{2}{c_1}\frac{j}{G}-1-2\hat{p}_0\hat{t}\right)^2+\mathcal{O}(G^{-2})\right]\left(1+\mathcal{O}\left(\hat{\sigma}e^{-1/8\hat{\sigma}^2}\right)\right) \qquad (234) \end{aligned}$$

for a constant of proportionality independent of j. We wish to show that the wavepacket is localized on the identity gates at the end of the circuit. To achieve this behavior, we must appropriately choose the identity padding, wavepacket width, wavepacket momentum, and evolution time. One such choice is as follows, where R will ultimately be a large constant:

$$c_2 = R^2, \quad c_1 = R, \quad \hat{\sigma}^2 = 1/R, \quad \hat{p}_0 = R^2, \quad \hat{t} = 1/R\,, \tag{235}$$

which implies

$$|\langle h_j|\langle j|\psi(t)\rangle|^2 \propto \exp\left[-\frac{R}{10}\left(\frac{2}{R}\frac{j}{G}-1-2R\right)^2+\mathcal{O}(G^{-2})\right)\right]\left(1+e^{-\mathcal{O}(R)}\right)\,. \tag{236}$$

To show that the wavepacket is localized on the identity padding at the end of the Hamiltonian, we show that the support on the first $(c_1+1)G$ sites (i.e., before $\mathcal{U}$ is computed) is vanishingly small compared to the support on the final identity padding (i.e., after $\mathcal{U}$ is computed). The LHS below defines this ratio ζ after substituting in the R parameter. For our choice of time and momentum, it suffices to only evaluate the support on the last G sites of the Hamiltonian, i.e., we can upper-bound the ratio ζ by the quantity

$$\zeta = \frac{\sum_{j=0}^{(R+1)G}|\langle h_j|\langle j|\psi(t)\rangle|^2}{\sum_{j'=(R+1)G}^{(R^2+R+1)G}|\langle h_{j'}|\langle j'|\psi(t)\rangle|^2} \leq \frac{\sum_{j=0}^{(R+1)G}|\langle h_j|\langle j|\psi(t)\rangle|^2}{\sum_{j'=(R^2+R)G}^{(R^2+R+1)G}|\langle h_{j'}|\langle j'|\psi(t)\rangle|^2}\,. \tag{237}$$

We then bound the numerator and denominator by the maximal and minimal terms in each sum respectively to obtain

$$\begin{aligned} \zeta &\leq \left(1+e^{-\mathcal{O}(R)}\right)\frac{((R+1)G+1)\exp\left[-\frac{R}{10}\left(\frac{2(R+1)}{R}-1-2R\right)^2+\mathcal{O}(G^{-2})\right)\right]}{G\exp\left[-\frac{R}{10}\left(\frac{2(R^2+R)}{R}-1-2R\right)^2+\mathcal{O}(G^{-2})\right)\right]} \qquad (238) \\ &= \exp\left[-\Omega(R^3\log R)\right]\,. \qquad (239) \end{aligned}$$

Hence, we can always choose a sufficiently large constant R such that the support on the sites where the computation is unfinished becomes arbitrarily small, compared to the support on the sites where the computation is finished. This completes the proof. □