Limits of Short-Time Evolution of Local Hamiltonians

Ali Hamed Moosavian, Seyed Sajad Kahani, and Salman Beigi

QuOne Lab, Phanous Research and Innovation Centre, Tehran, Iran

arXiv:2104.12808v5 [quant-ph] 24 Jun 2022

Evolutions of local Hamiltonians in short times are expected to remain local and thus limited. In this paper, we validate this intuition by proving some limitations on short-time evolutions of local time-dependent Hamiltonians. We show that the distribution of the measurement output of short-time (at most logarithmic) evolutions of local Hamiltonians are *concentrated* and satisfy an *isoperimetric inequality*. To showcase explicit applications of our results, we study the MaxCut problem and conclude that quantum annealing needs at least a run-time that scales logarithmically in the problem size to beat classical algorithms on MaxCut. To establish our results, we also prove a Lieb-Robinson bound that works for time-dependent Hamiltonians which might be of independent interest.

1 Introduction

The Quantum Annealing (QA) algorithm [1] was first introduced as a general purpose optimization algorithm for quantum computers in [2]. The QA algorithm is based on the adiabatic theorem in quantum mechanics [3]. The adiabatic theorem states that if we initialize a system in its ground state, and then evolve the Hamiltonian of the system slowly enough, as long as the spectral gap of the Hamiltonian does not close, the system remains close to the ground state of the system at all times [4, 5]. Because of its simple conceptual structure, QA has been implemented on many different platforms and applied to a variety of different problems, see *e.g.* [6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20].

In this manuscript, we focus on short-time QA algorithms. This setup is motivated by the current state of the art quantum technologies (see methods of [21]). On one hand, we still lack universal quantum computers capable of simulating general QA algorithms [22, 23], and on the other hand the analog QA devices lack fault-tolerance. Therefore, the total time of reliably performing a continuous computation on these analog devices is bounded by their decoherence time.

By considering short-time quantum annealing, we exit the scope of the adiabatic theorem. Hence, our results are not sensitive to the Hamiltonian gap. In this context, the algorithms considered in our work are sometimes called "diabatic quantum annealing" in the literature (E.g. see [24] and references therein).

Quantum Approximate Optimization Algorithm (QAOA) is another general purpose quantum algorithm that can yield approximate solutions to many optimization problems, particularly combinatorial ones [25]. Based on QAOA's local nature, recent developments have shown that low-depth QAOAs are unable to produce better approximate solutions to some optimization problems than the best known classical algorithms [26, 27, 28, 29, 30].

In particular, Bravyi et al [28] use the *isoperimetric inequality* of [31] that puts restrictions on the output distributions of all low-depth circuits, to show that low-depth (at most logarithmic) QAOA for MaxCut does not outperform best classical algorithms on *Ramanujan graphs*. Also, Farhi et al [26] show that low-depth QAOA for the problems of MaxCut and Maximum Independent Set, does not perform well on random regular bipartite graphs.

Thinking of the correspondence between short-time QA algorithms and low-depth quantum circuits, particularly low-depth QAOAs, it is natural to ask whether the above bounds hold for QA algorithms or not. In this paper, we show that most of the aforementioned results on limitations of QAOAs still hold for short-time QA algorithms as well. Throughout this article we use short-time to represent times that scale at most logarithmi-

Ali Hamed Moosavian: ali@phanous.ir
Seyed Sajad Kahani: sajad@phanous.ir
Salman Beigi: salman@phanous.ir

cally with the system size. We first show that the isoperimetric inequality of [31] remains to hold for output distributions of short-time QA computations. Then, using this inequality as a tool (besides others), we generalize the results of [28] and [26] to the QA case.

QA and QAOA can be thought of two instances of optimal control problems. In QA we look for an optimal path in the space of Hamiltonians from a starting Hamiltonian to a target one, while in QAOA we look for optimal rotation angles of some local circuits. Applying *Pontryagin's maximum principle*, some authors have argued that the optimal solution to the first control theory problem takes the so called *bang-bang* cond problem. However, it was shortly noted that in some common problem instances the control Hamiltonian of the system becomes *singular*, a condition that was presumed to be extremely rare, and then the optimal control protocol is no longer of the bang-bang form [32, 33, 34]. This means that, in general, the optimal control problem for QA does not reduce to that of the QAOA, and needs thoughtful considerations. In fact, the observations of [32, 33, 34] show that we cannot directly use the results on the obstructions of QAOAs, to prove the aforementioned limitations of short-time QA algorithms.

Another idea for relating results on low-depth quantum circuits (QAOAs) to QA algorithms is *trotterization* (see [35] and references therein), which approximates the unitary evolution of a local Hamiltonian with a quantum circuit. Although one can imagine this approach working in some very specific instances, in general the trotterized circuit must have an asymptotic depth that grows polynomially with the number of qubits. This prohibits the application of the aforementioned QAOA results, [26, 27, 28, 29, 30], as they only apply to logarithmic depths.

We use *Lieb-Robinson bounds* as a main tool for proving our results [36, 37, 38, 39]. To this end, we first prove a Lieb-Robinson bound that works for time-dependent Hamiltonians and is amenable to QA algorithms. To the best of our knowledge, this is the first time-dependent Lieb-Robinson bound.

As another technical tool, we use the notion of γ_2-*norm*, also called *factorization norm*, from matrix analysis to bound the derivative of certain functions on the space of operators. This enables us to bound certain errors induced by applying the Lieb-Robinson approximation.

Here is a summary of our results and the structure of the paper. In Sec. 2 we state our version of the Lieb-Robinson bound that is a fundamental tool in our work. Then in Sec. 3 we outline the proof of two different bounds on the output distributions of short-time evolutions of local time-dependent Hamiltonians. Our first bound is a generalization of the well-known *Chebyshev inequality* which says that the Hamming weight of the measurement outcome of a short-time local Hamiltonian evolution is *concentrated* around its mean. The next bound, that is independent of the first one, is a generalization of the isoperimetric inequality of [31] for short-time Hamiltonian evolutions. Then, in Sec. 4, we take the approach of [28], and show that the $\mathbb{Z}_2$-symmetry of the MaxCut Hamiltonian limits the performance of QA algorithms on Ramanujan graphs. To prove this result we use our isoperimetric inequality. Next, in Sec. 5 , we prove a similar limitation on random regular bipartite graphs using the Lieb-Robinson bound. This result is an extension of the result of [26] for QA algorithms. Final remarks come in Sec. 6, and many of the detailed proofs are left for Supplemental materials.

2 Lieb-Robinson Bound

One of our main tools in this manuscript is the Lieb-Robinson bound [40]. Roughly speaking, it states that the evolution of a local observable in the Heisenberg picture, remains almost local in short times. We use the following version of the Lieb-Robinson bound that notably works for *time-dependent Hamiltonians* as well.

To state our theorem we need a definition. Given a graph G and a subset of its vertices A, we denote the set of vertices on the L-*boundary* of A by $\partial_L(A)$. That is,

$$\partial_L(A) = \{x \in A : \exists y \in A^c, \mathrm{dist}_G(x,y) \leq L\} \quad (1)$$
$$\cup \{x \in A^c : \exists y \in A, \mathrm{dist}_G(x,y) \leq L\}, \quad (2)$$

where A^c is the subgraph consisting of vertices not in A and $\mathrm{dist}_G(x,y)$ is the graph distance between vertices x and y.

Theorem 1 (Lieb-Robinson Bound)**.** *Let G be a graph that may contain loops and parallel edges*

with maximum degree $\Delta > 1$. Let $\mathbf{H}(t)$ be a time-dependent Hamiltonian defined on vertices of G of the form

$$\mathbf{H}(t) = \sum_e u_e(t)\mathbf{h}_e, \tag{3}$$

where $\mathbf{h}_e$ is time-independent and acts non-trivially only on the two ends of edge e in G. Let $\mathbf{U}(t)$ be the unitary evolution associated with $\mathbf{H}(t)$. Suppose that $\|u_e(t)\mathbf{h}_e\| \leq g$ for any e and $0 \leq t \leq T$. Throughout this paper $\|.\|$ represents the spectral norm unless it is clearly defined otherwise. Let $\mathbf{O}_A$ be an observable acting on region A of the graph. For $L > 1$ let

$$\widetilde{\mathbf{H}}_A(t) := \sum_{e \subseteq A \cup \partial_L(A)} u_e(t)\mathbf{h}_e, \tag{4}$$

be the Hamiltonian consisting of terms in the region $A \cup \partial_L(A)$, and let $\mathbf{V}_A(t)$ be the unitary evolution associated to $\widetilde{\mathbf{H}}_A(t)$. Then we have

$$\left\|\mathbf{U}^\dagger(T)\mathbf{O}_A\mathbf{U}(T) - \mathbf{V}_A^\dagger(T)\mathbf{O}_A\mathbf{V}_A(T)\right\| \leq \sqrt{\frac{2}{\pi}}\,|A| \cdot \|\mathbf{O}_A\| \cdot \mathrm{e}^{-L\left(\log L - \log T - \log(4g(\Delta-1))\right) - \frac{1}{2}\log L}\,. \tag{5}$$

Intuitively speaking, this theorem says that, as $\mathbf{O}_A$ is local, it remains *almost* local under a short-time evolution of a local Hamiltonian. Indeed, since $\widetilde{\mathbf{H}}_A(t)$ is non-trivial only in the region $A \cup \partial_L(A)$, then $\mathbf{V}_A(T)$ also acts non-trivially only on this region. Therefore, by this theorem as it is schematically depicted in Fig. 1, $\mathbf{U}^\dagger(T)\mathbf{O}_A\mathbf{U}(T)$ is approximated by $\mathbf{V}_A^\dagger(T)\mathbf{O}_A\mathbf{V}_A(T)$ that acts only on $A \cup \partial_L(A)$.

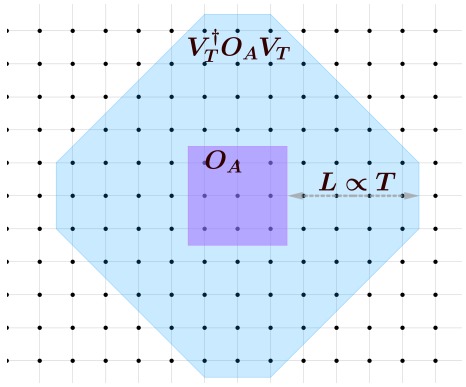


Figure 1: (Color online) Illustration of Theorem 1. Here, A is the region on which the operator $\mathbf{O}_A$ acts, and the larger shaded area represents $\partial_L(A)$. $\mathbf{V}_A^\dagger(T)\mathbf{O}_A\mathbf{V}_A(T)$ acts non-trivially on $A \cup \partial_L(A)$ and approximates $\mathbf{U}^\dagger(T)\mathbf{O}_A\mathbf{U}(T)$, the time evolution of $\mathbf{O}_A$ in the Heisenberg picture.

Some remarks are in line regarding this theorem. First, as the proof of this theorem given in Appendix A shows, the above Lieb-Robinson bound can easily be generalized for local Hamiltonians with k-body interactions, by replacing the term $\log(4g(\Delta-1))$ on the right hand side of Eq. (5) with $\log(2kg(\Delta-1))$. Second, this theorem in the case where $\mathbf{H}$ is time-independent is proven in [41, 37, 38]. Third, the type of time-dependent Hamiltonians (Equation (3)) considered by this theorem are motivated by some quantum annealing applications in mind. Nevertheless, note that in principle any time-dependent Hamiltonian can be written in this form; simply consider the expansion of each local term of the Hamiltonian in the Pauli basis.

3 Limits on the output distribution of short-time evaluations

In this section we present our main results, namely, how the output distribution of a short-time quantum annealing algorithm is constrained. Our first main result is a generalization of the well-known *Chebyshev inequality* in probability theory that is a concentration of measure inequality. Our second result is an *isoperimetric inequality*. A similar result for low-depth circuits has been proven in [31].

In the following we assume that $\mathbf{H}(t)$ is a time-dependent Hamiltonian defined on the vertices of a graph G as in Eq. (3). We assume that starting with a *product state*

$$|\psi_0\rangle = |s_1\rangle \otimes |s_2\rangle \otimes \cdots \otimes |s_n\rangle\,, \tag{6}$$

the system evolves under the Hamiltonian $\mathbf{H}(t)$. Let

$$|\psi_t\rangle = \mathbf{U}(t)\,|\psi_0\rangle\,, \tag{7}$$

where as before $\mathbf{U}(t)$ is the associated unitary evolution. We assume that at time T we make a measurement in the computational basis to a get a distribution $p(\mathbf{x})$ on $\{0,1\}^n$.

3.1 A Chebyshev bound

Theorem 2 (Quantum Chebyshev bound)**.** *Let* $\mathbf{H}(t)$ *be a Hamiltonian on a graph with maximum degree* $\Delta > 1$ *as in Eq.* (3) *with* $\|u_e(t)\mathbf{h}_e\| \leq g$ *for any edge* e*. Let*

$$T \leq \frac{\kappa_1}{8g\Delta^{\frac{2-\kappa_1}{\kappa_1}}\log\Delta}\log n, \tag{8}$$

where $0 < \kappa_1 < 1$ *is an arbitrary constant. Let* $p(\mathbf{x})$ *be the output distribution of measuring* $|\psi_T\rangle = \mathbf{U}(T)|\psi_0\rangle$ *in the computational basis. We assume* $|\psi_0\rangle$ *to be a product state as in Eq.* (6)*. Then for any* $c > 0$ *and* $\kappa_1/2 < \kappa_2 < 1/2$ *we have*

$$\begin{aligned}&\Pr\left[\left|w_H(\mathbf{x}) - \left(\frac{n}{2} - \frac{m}{2}\right)\right| \geq cn^{\frac{1}{2}+\kappa_2}\right] \\ &\leq \frac{3}{2c^2}n^{-(2\kappa_2-\kappa_1)},\end{aligned} \tag{9}$$

where $w_H(\mathbf{x})$ *is the hamming weight of* $\mathbf{x} \in \{0,1\}^n$*, i.e. number of non-zero coordinates of* $\mathbf{x}$*, and* $m = \sum_i \langle\psi_T|\mathbf{Z}_i|\psi_T\rangle$ *with* $\mathbf{Z}$ *being the Pauli-*z *operator.*

This theorem says that, when T is bounded, with high probability the hamming weight of the measurement output is close to $\frac{1}{2}(n-m)$ which is the average of $w_H(\mathbf{x})$. That is, the Hamming weight of the measurement outcome is *concentrated*.

The proof idea of this theorem is similar to that of the standard Chebyshev inequality; we need to bound the variance of $w_H(\mathbf{x})$. In the standard Chebyshev inequality the coordinates of $\mathbf{x}$ are independent while here they are correlated. Nevertheless, since T is bounded, by the Lieb-Robinson bound (Theorem 1) pairs of coordinates whose associated Hamming weights are far from each other are almost independent. This allows us to bound the variance of $w_H(\mathbf{x})$. A detailed proof of this theorem is given in Appendix B.

As a result of this theorem, as long as the initial state is a product state, no short-time annealing computation can produce the n-qubit GHZ state $\frac{1}{\sqrt{2}}(|0\cdots0\rangle + |1\cdots1\rangle)$ in sub-logarithmic times. This is because measuring this state in the computational basis gives two far apart peaks in the hamming weight w_H, Fig. 2. We note that this statement holds for any annealing computation even with gapless Hamiltonians. Intuitively speaking, this is a result of the local nature of the Hamiltonian in Eq. (3).

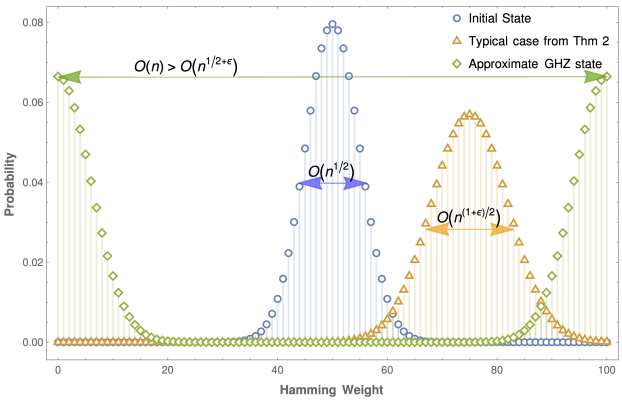


Figure 2: (Color online) Illustration of Theorem 2 depicting the probability distribution of the Hamming weight of three different n-qubit states for $n = 100$. The centered distribution in the middle (blue) is the typical distribution for a product initial state (e.g., $|+\rangle^{\otimes n}$). The distribution that is peaked just to the right (orange) represents the distribution after a short time evolution of the initial state described by Theorem 2. The distribution that has two peaks on either side of the spectrum (green) is an approximation for the GHZ state. This shows that the GHZ state cannot be generated even approximately by a short-time quantum annealing.

3.2 An isoperimetric inequality

We now prove that the output distribution of a short-time quantum annealing algorithm satisfies an *isoperimetric inequality*, see Figs. 3 and 4. Such a result for low-depth circuits was first proved in [31].

Theorem 3. *Let* $\mathbf{H}(t)$ *be a Hamiltonian on a graph with maximum degree* $\Delta > 1$ *as in Eq.* (3) *with* $\|u_e(t)\mathbf{h}_e\| \leq g$ *for any edge* e*. Let* $|\psi_0\rangle = |s_1\rangle \otimes |s_2\rangle \otimes \cdots \otimes |s_n\rangle$ *be an arbitrary product state and let* $|\psi_T\rangle = \mathbf{U}(T)|\psi_0\rangle$ *be the state of the system at time* $T > 0$*. Let* $p(\mathbf{x})$ *be the distribution on* $\{0,1\}^{n_0}$ *resulting from measuring the first* n_0 *qubits of* $|\psi_T\rangle$ *in the computational basis,* i.e.*,*

$$p(\mathbf{x}) = \sum_{\mathbf{y}\in\{0,1\}^{n-n_0}} |\langle\mathbf{x},\mathbf{y}|\psi_T\rangle|^2. \tag{10}$$

Then for any $F \subset \{0,1\}^{n_0}$ *with* $p(F) \leq \frac{1}{2}$ *and any* $0 \leq \theta \leq 1/2$ *the following results hold:*

(i) *Assume that*

$$T \leq \frac{\kappa_1}{4g\log(\Delta)\Delta^{\frac{1+2\kappa_1+\kappa_2}{\kappa_1}}}\log n_0\,, \tag{11}$$

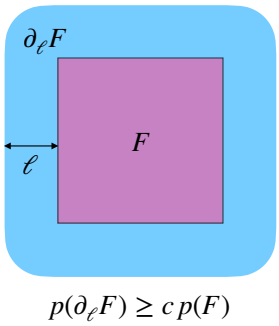


Figure 3: (Color online) A schematic presentation of the isoperimetric inequality, stating that the boundary of a region is large comparing to the region itself.

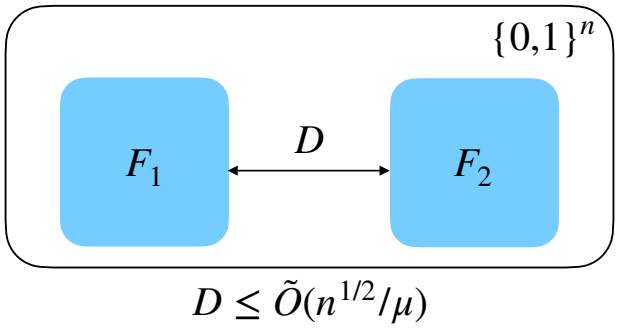


Figure 4: (Color online) Theorem 4 says that measuring in terms of the output distribution of a short-time QA algorithm, if two sets have size at least μ, then their distance (ignoring some logarithmic factors) is at most $n^{1/2}/\mu$.

for some constants $\kappa_1, \kappa_2 > 0$. Then we have

$$p(\partial_\ell(F)) \geq \frac{1}{8}(2n_0^{1+\kappa_1})^{-2\theta} p(F) - 2n_0^{-\kappa_2}. \tag{12}$$

where

$$\ell = \frac{\kappa_1}{\sqrt{2}\log(\Delta)} \log(n_0) n_0^{\left(\frac{1}{2}-\theta\right)(1+\kappa_1)}. \tag{13}$$

(ii) *Assume that we measure all the qubits,* i.e. $n_0 = n$. *Moreover, assume that*

$$T \leq \frac{\kappa_1}{4g(\Delta-1)} \log n\,, \tag{14}$$

and

$$\ell = \frac{\kappa_1(1+\kappa_2)}{2} \log(n)\sqrt{n} \tag{15}$$

for some constants $\kappa_1, \kappa_2 > 0$. Then we have

$$p(\partial_\ell(F)) \geq \frac{1}{8}p(F) - \frac{5}{4}n^{-\left(\kappa_1(1+\kappa_2)\log(1+\kappa_2)-1\right)}. \tag{16}$$

This theorem says that, measuring in terms of the outcome probabilities of a short-time quantum annealing computation, the size of a boundary of a set is large comparing to the size of a set itself, see Fig. 3. The proof of this theorem is given in Appendix D.

Theorem 4. *Let $\mathbf{H}(t)$ be a Hamiltonian on a graph with maximum degree $\Delta > 1$ as in Eq. (3) with $\|u_e(t)\mathbf{h}_e\| \leq g$ for any edge e. Let $|\psi_0\rangle = |s_1\rangle \otimes |s_2\rangle \otimes \cdots \otimes |s_n\rangle$ be an arbitrary product state and let $|\psi_T\rangle = \mathbf{U}(T)|\psi_0\rangle$ be the state of the system at time $T > 0$. Let $p(\mathbf{x})$ be the distribution on $\{0,1\}^n$ resulting from measuring qubits of $|\psi_T\rangle$ in the computational basis. Moreover, assume that*

$$\begin{aligned} T &\leq \frac{\kappa_1}{4g(\Delta-1)} \log n, \\ \ell &= \frac{\kappa_1(1+\kappa_2)}{2} \log(n)\sqrt{n}, \end{aligned} \tag{17}$$

for some constants $\kappa_1, \kappa_2 > 0$. Then for two arbitrary disjoint sets $F_1, F_2 \subset \{0,1\}^n$ with Hamming distance $D = \mathrm{dist}_H(F_1, F_2)$ and $\mu := \min\{p(F_1), p(F_2)\}$, we have

$$D < \frac{16\ell}{\mu} + \frac{10D}{\mu} n^{-\left(\kappa_1(1+\kappa_2)\log(1+\kappa_2)-1\right)} \tag{18}$$

$$\leq \frac{16\ell}{\mu} + \frac{10}{\mu} n^{-\left(\kappa_1(1+\kappa_2)\log(1+\kappa_2)-2\right)}. \tag{19}$$

The contra-positive version of this theorem says that if two far apart subsets of $\{0,1\}^n$ have high probability under the distribution produced by the annealing algorithm, then the run-time of the annealing algorithm is necessarily long.

Here is a high level perspectice of how the proofs are laid out in the appendix. In order to prove Theorem 3, we need to bound the derivative of a certain Chebyshev polynomial of some quantum operators. Because of the approximate nature of the Lieb-Robinson bound, proving this requires an accurate mathematical analysis of the polynomials using the γ_2-norm. More naive approaches would have disproportionately large errors. The proof is laid out in Appendix C. Then, by utilizing these results we prove Theorem 3 in Appendix D. Theorem 4 is basically a corollary of Theorem 3 and it is proven in Appendix E.

4 MaxCut on Ramanujan Graphs

In this section and the following one, we present limits of short-time quantum annealing algorithms for the MaxCut problem. Here, inspired by [28] our strategy is to use the $\mathbb{Z}_2$-symmetry of the MaxCut Hamiltonian, and show that short-time quantum annealing algorithm cannot outperform best classical algorithms for MaxCut on Ramanujan graphs [42].

We first describe the quantum annealing algorithm for MaxCut.

Given a graph $G = (V, E)$, the MaxCut problem asks for a *cut* with maximum size in G; meaning what is the maximum number of edges between vertices in V_+ and V_- where $V_+ \cup V_- = V$ form a partition of vertices? Labeling vertices in V_+ with $+1$ and vertices in V_- with -1, this problem is equivalent to minimizing the energy of the following Hamiltonian:

$$\mathbf{C} := -\frac{1}{2}\sum_{(i,j)\in E}(\mathbf{I} - \mathbf{Z}_i\mathbf{Z}_j), \tag{20}$$

Where $\mathbf{Z}_i$'s are the Pauli-z operators. Indeed, letting Cut* be the maximum cut of G we have

$$\mathrm{Cut}^* = -\min_{|\phi\rangle} \langle\phi|\,\mathbf{C}\,|\phi\rangle\,. \tag{21}$$

We may use quantum annealing to find the ground state and minimum energy of $\mathbf{C}$. To this end, a common choice for the *driving Hamiltonian* is the summation of Pauli-x operators:

$$\mathbf{B} := -\sum_{k\in V}\mathbf{X}_k. \tag{22}$$

Using a control parameter $u : [0, T] \to [0, 1]$ with $u(0) = 0$ and $u(T) = 1$, the resulting time-dependent Hamiltonian is given by

$$\mathbf{H}(t) := (1 - u(t))\,\mathbf{B} + u(t)\mathbf{C} \tag{23}$$

At time $t = 0$, we start with the ground state of $\mathbf{H}(0) = \mathbf{B}$ that is equal to

$$|\psi_0\rangle = |+\rangle^{\otimes n}\,, \tag{24}$$

where $n = |V|$ is the number of vertices. In the following we show that if T is bounded, then

$$\langle\psi_T|\,\mathbf{C}\,|\psi_T\rangle\,, \tag{25}$$

is bounded away from the ground energy, i.e., $-\mathrm{Cut}^*$, for a certain family of graphs to be defined.

For a graph $G = (V, E)$ on n vertices, its *Cheeger constant* is defined by [43]:

$$h(G) = \min_{S\subset V, 0<|S|\le\frac{n}{2}} \frac{\left|\partial^E(S)\right|}{|S|}\,, \tag{26}$$

where $\partial^E(S)$ is the set of *edges* with one end in S and one end outside of S. *Ramanujan graphs* are certain families of graphs that have a large Cheeger constant; for a Δ-regular Ramanujan graph G we have

$$h(G) \ge \frac{1}{2}(\Delta - 2\sqrt{\Delta - 1}). \tag{27}$$

It is known that [44, 45, 46] for any $\Delta \ge 3$ there exists an infinite family of Δ-regular *bipartite* Ramanujan graphs. We note that for such a graph, as a bipartite graph, we have

$$\mathrm{Cut}^* = |E| = \frac{1}{2}\Delta n. \tag{28}$$

Theorem 5. *Let G be a Δ-regular bipartite Ramanujan graph on n vertices. Consider the quantum annealing algorithm given by Eqs. (23) and (24). Assume that T satisfies*

$$T \le \frac{\kappa_1}{4\Delta}\log n\,, \tag{29}$$

for some $\kappa_1 > 0$. Then, for any $0 < \alpha < 1/2$, $0 < \epsilon < 1$ and sufficiently large n we have

$$\frac{-\langle\psi_T|\,\mathbf{C}\,|\psi_T\rangle}{\mathrm{Cut}^*} < 1 - \alpha(1-\epsilon) + 2\alpha(1-\epsilon)\frac{\sqrt{\Delta-1}}{\Delta}. \tag{30}$$

We note that for appropriate choices of α, ϵ and for sufficiently large (but constant) $\Delta \ge 6$, the right hand side of Eq. (30) can be made smaller than 0.87856. This means that the quantum annealing algorithm with $T \le O(\log n)$ does not improve on the classical algorithm of Goemans and Williamson for MaxCut [47].

As the proof of this theorem in Appendix F shows, the smallest n for which the bound Eq. (30) holds can be explicitly estimated in terms of $\alpha, \epsilon, \kappa_1$ and Δ.

We remark that, comparing to Eq. (14), in Eq. (29) the parameter Δ is replaced with $\Delta + 1$. The point is that in the MaxCut Hamiltonian Eq. (23), the number of terms acting on a vertex is $\Delta + 1$ due to the extra term coming from $\mathbf{B}$.

5 MAXCUT on Random Bipartite Graphs

Inspired by the recent work of Farhi, Gamarnik and Gutmann [26], in this section we claim that short-time QA algorithm for MaxCut does not perform well on random bipartite graphs. The idea is that such a short-time QA algorithm *does not see the whole graph*, so it cannot distinguish between a bipartite random graph and an arbitrary random graph. Thus, the output cut-size of these algorithms on random bipartite graphs is bounded by the MaxCut of arbitrary random graphs.

Theorem 6. *Consider the QA algorithm using the Hamiltonian (23) with the initial state* $|\psi_0\rangle = |+\rangle^{\otimes n}$*, and time*

$$T \leq \frac{\kappa}{16 \log(\Delta) \Delta^{1+4/\kappa}} \log n \,, \tag{31}$$

for some $0 < \kappa < 1$*. Then, for any* $\nu > 0$ *and sufficiently large* n*, the expected cut size from applying the algorithm on a random* Δ*-regular bipartite graphs is at most*

$$\left(\frac{\Delta}{4} + O(\sqrt{\Delta}) + 2\nu \right) n \,. \tag{32}$$

Note that, by the bound in Eq. (32), the performance of the QA algorithm on random Δ-regular bipartite graphs is not much better than outputting a random cut.

The proof of this theorem is given in Appendix G.

Farhi et al also show how the low-depth QAOA algorithm cannot outperform classical algorithms for the maximum independent set problem on random bipartite graphs [26]. Using similar arguments to what we provided in this section, that result can also be generalized to short-time QA algorithms.

6 Conclusions

Motivated by short-time quantum annealing, in this manuscript we proved limits on the output distribution of local time-dependent Hamiltonian evolutions. In particular, we showed that the distribution of the measurement outcome of a local Hamiltonian evolution is concentrated and satisfies an isoperimetric inequality if the evolution time is short. Using our framework, we generalized results about limitations of low-depth quantum circuits, to short-time QA algorithms. To this end, an essential tool that we developed, is a Lieb-Robinson bound that works for time-dependent Hamiltonians and is amenable to QA algorithms.

Our results are general in the sense that they work for the circuit model as well; they imply limitations on the output distribution of low-depth circuits with a locality structure. The point is that we can assume that the control parameters $u_e(t)$ in Eq. (3) have a so-called bang-bang form: they are either zero or take a fixed value, and for each t the values of $u_{e_1}(t), u_{e_2}(t)$ for intersecting edges e_1, e_2 cannot both be non-zero. With such bang-bang control parameters, low-depth local circuits can be seen as special cases of short-time evolutions of local Hamiltonians for which our results hold.

The short-time condition is strictly essential for the bounds we have proven, and our results do not hold without this assumption. Indeed, there are several known examples where quantum annealing with Poly(n) run-time can provide superpolynomial speedups. For example, see [48, 49, 50].

We believe that the techniques presented here can be applied to generalize other results about the limitation of low-depth circuits too, see *e.g.* [29, 27], for short-time QA algorithm. We leave this for future research.

References

[1] T. Kadowaki and H. Nishimori. Quantum annealing in the transverse Ising model. Physical Review E **58**, 5355–5363 (1998).

[2] E. Farhi, J. Goldstone, S. Gutmann and M. Sipser. Quantum Computation by Adiabatic Evolution. arXiv:0001106 [quant-ph] (2000).

[3] T. Kato. On the adiabatic theorem of quantum mechanics. Journal of the Physical Society of Japan **5**, 435–439 (1950).

[4] M. Born and V. Fock. Beweis des adiabatensatzes. Zeitschrift für Physik **51**, 165–180 (1928).

[5] T. Albash and D. A. Lidar. Adiabatic quantum computation. Reviews of Modern Physics **90**, 015002 (2018).

[6] I. Hen and F. M. Spedalieri. Quantum Annealing for Constrained Optimization. Physical Review Applied **5**, 034007 (2016).

[7] S. Puri, C. K. Andersen, A. L. Grimsmo, and A. Blais. Quantum annealing with all-to-all connected nonlinear oscillators. Nature Communications **8**, 15785 (2017).

[8] W. Lechner, P. Hauke, and P. Zoller. A quantum annealing architecture with all-to-all connectivity from local interactions. Science Advances **1**, e1500838 (2015).

[9] S. Jiang, K. A. Britt, A. J. McCaskey, T. S. Humble, and S. Kais. Quantum Annealing for Prime Factorization. Scientific Reports **8**, 17667 (2018).

[10] R. Y. Li, R. Di Felice, R. Rohs, and D. A. Lidar. Quantum annealing versus classical machine learning applied to a simplified computational biology problem. NPJ quantum information **4**, 1–10 (2018).

[11] L. Stella, G. E. Santoro, and E. Tosatti. Optimization by quantum annealing: Lessons from simple cases. Physical Review B **72**, 014303 (2005).

[12] O. Titiloye and A. Crispin. Quantum annealing of the graph coloring problem. Discrete Optimization **8**, 376–384 (2011).

[13] A. Mott, J. Job, J.-R. Vlimant, D. Lidar, and M. Spiropulu. Solving a Higgs optimization problem with quantum annealing for machine learning. Nature **550**, 375–379 (2017).

[14] K. L. Pudenz, T. Albash, and D. A Lidar. Quantum annealing correction for random Ising problems. Physical Review A **91**, 042302 (2015).

[15] A. Perdomo-Ortiz, N. Dickson, M. Drew-Brook, G. Rose, and A. Aspuru-Guzik. Finding low-energy conformations of lattice protein models by quantum annealing. Scientific reports **2**, 571 (2012).

[16] K. L. Pudenz, T. Albash, and D. A Lidar. Error-corrected quantum annealing with hundreds of qubits. Nature communications **5**, 1–10 (2014).

[17] R. Martoňák, G. E. Santoro, and E. Tosatti. Quantum annealing of the traveling-salesman problem. Physical Review E **70**, 057701 (2004).

[18] S. H. Adachi and M. P. Henderson. Application of quantum annealing to training of deep neural networks. arXiv:1510.06356 [quant-ph] (2015).

[19] M. W Johnson, et al. Quantum annealing with manufactured spins. Nature **473**, 194–198 (2011).

[20] S. Boixo, T. Albash, F. M. Spedalieri, N. Chancellor, and D. A. Lidar. Experimental signature of programmable quantum annealing. Nature communications **4**, 1–8 (2013).

[21] A. D. King, et al. Coherent quantum annealing in a programmable 2000-qubit Ising chain. arXiv:2202.05847 [quant-ph] (2022).

[22] B. Foxen, et al. Demonstrating a Continuous Set of Two-qubit Gates for Near-term Quantum Algorithms. Physical Review Letters **125**, 120504 (2020).

[23] K. Wright, et al. Benchmarking an 11-qubit quantum computer. Nature communications **10**, 1–6 (2019).

[24] E. J. Crosson and D. A. Lidar. Prospects for quantum enhancement with diabatic quantum annealing. Nature Reviews Physics **3**, 466–489 (2021).

[25] E. Farhi, J. Goldstone, and S. Gutmann. A Quantum Approximate Optimization Algorithm. arXiv:1411.4028 [quant-ph] (2014).

[26] E. Farhi, D. Gamarnik, and S. Gutmann. The Quantum Approximate Optimization Algorithm Needs to See the Whole Graph: Worst Case Examples. arXiv:2005.08747 [quant-ph] (2020)

[27] E. Farhi, D. Gamarnik, and S. Gutmann. The Quantum Approximate Optimization Algorithm Needs to See the Whole Graph: A Typical Case. arXiv:2004.09002 [quant-ph] (2020).

[28] S. Bravyi, A. Kliesch, R. Koenig, and E. Tang. Obstacles to Variational Quantum Optimization from Symmetry Protection. Physical Review Letters **125**, 260505 (2020).

[29] S. Bravyi, D. Gosset, and R. Movassagh. Classical algorithms for quantum mean values. Nature Physics **17**, 337–341 (2021).

[30] S. Bravyi, A. Kliesch, R. Koenig, and E. Tang. Hybrid quantum-classical algorithms for approximate graph coloring. Quantum **6**, 678 (2022).

[31] L. Eldar and A. W. Harrow. Local Hamiltonians Whose Ground States are Hard to Approximate. In 2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS), 427–438 (2017).

[32] L. T. Brady, C. L. Baldwin, A. Bapat, Y. Kharkov, and A. V. Gorshkov. Optimal Protocols in Quantum Annealing and Quantum Approximate Optimization Algorithm Problems. Physical Review Letters **126**, 070505 (2021).

[33] L. T. Brady, L. Kocia, P. Bienias, A. Bapat, Y. Kharkov, and A. V. Gorshkov. Behavior of Analog Quantum Algorithms. arXiv:2107.01218 [quant-ph] (2021).

[34] L. C. Venuti, D. D'Alessandro, and D. A. Lidar. Optimal Control for Quantum Optimization of Closed and Open Systems. Physical Review Applied **16**, 054023 (2021).

[35] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu. Theory of Trotter Error with Commutator Scaling. Physical Review X **11**, 011020 (2021).

[36] B. Nachtergaele, Y. Ogata, and R. Sims. Propagation of correlations in quantum lattice systems. Journal of Statistical Physics **124**, 1–13 (2006).

[37] B. Nachtergaele and R. Sims. Lieb-Robinson bounds in quantum many-body physics. Contemporary Mathematics **529**, 141–176 (2010).

[38] S. Bravyi, M. B. Hastings, and F. Verstraete. Lieb-robinson bounds and the generation of correlations and topological quantum order. Physical Review Letters **97**, 050401 (2006).

[39] C.-F. Chen and A. Lucas. Operator growth bounds from graph theory. Communications in Mathematical Physics **385**, pages1273–1323 (2021).

[40] E.H. Lieb and D.W. Robinson. The finite group velocity of quantum spin systems. Communications in Mathematical Physics **28**, 251–257 (1972).

[41] J. Haah, M. B. Hastings, R. Kothari, and G. H. Low. Quantum algorithm for simulating real time evolution of lattice Hamiltonians. 2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS), 350–360 (2018).

[42] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. Combinatorica **8**, 261–277 (1988).

[43] B. Mohar. Isoperimetric numbers of graphs. Journal of Combinatorial Theory, Series B **47**, 274–291 (1989).

[44] A. W. Marcus, D. A. Spielman, and N. Srivastava. Interlacing Families IV: Bipartite Ramanujan Graphs of All Sizes. In 2015 IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS), 1358–1377 (2015).

[45] A. W. Marcus, D. A. Spielman, and N. Srivastava. Interlacing Families IV: Bipartite Ramanujan Graphs of All Sizes. SIAM Journal on Computing **47**, 2488–2509 (2018).

[46] C. Hall, D. Puder, and W. F. Sawin. Ramanujan coverings of graphs. Advances in Mathematics **323**, 367–410 (2018).

[47] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. Journal of the ACM **42**, 1115–1145 (1995).

[48] R. D. Somma, D. Nagaj, and M. Kieferová. Quantum Speedup by Quantum Annealing. Physical Review Letters **109**, 050501 (2012).

[49] M. B. Hastings. The Power of Adiabatic Quantum Computation with No Sign Problem. Quantum **5**, 597 (2021).

[50] A. Gilyén, M. B. Hastings, and U. Vazirani. (Sub)Exponential advantage of adiabatic Quantum computation with no sign problem. In Proceedings of the Annual ACM Symposium on Theory of Computing (STOC), 1357–1369 (2021).

[51] R. Bhatia. Matrix analysis. Graduate Texts in Mathematics. Springer New York (1996).

[52] R. Bhatia. Positive definite matrices. Princeton University Press, (2007).

[53] B.D. McKay, N.C. Wormald, and B. Wysocka. Short Cycles in Random Regular Graphs. The Electronic Journal of Combinatorics **11**, 1–12 (2004).

[54] F. Kardoš, D. Král, and J. Volec. Maximum edge-cuts in cubic graphs with large girth and in random cubic graphs. Random Structures & Algorithms **41**, 506–520 (2012).

[55] D. Coppersmith, D. Gamarnik, M.T. Hajiaghayi, and G.B. Sorkin. Random MAX SAT, random MAX CUT, and their phase transitions. Random Structures and Algorithms **24**, 502–545 (2004).

[56] A. Dembo, A. Montanari, and S. Sen. Extremal cuts of sparse random graphs. Annals of Probability **45**, 1190–1217 (2017).

A Proof of the Lieb-Robinson Bound

In this appendix we give a proof of Theorem 1. As mentioned before, the proof ideas are essentially the same as those of the original Lieb-Robinson bound. Here, for simplicity of presentation of the proof we first give two lemmas.

Lemma 1. *([37]) Let $\mathbf{X} = \mathbf{X}(t)$ and $\mathbf{Y} = \mathbf{Y}(t)$ be time-dependent bounded operators with $\mathbf{X}$ being Hermitian. Let $\mathbf{F} = \mathbf{F}(t)$ be given by*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{F} = i[\mathbf{X}, \mathbf{F}] + \mathbf{Y}. \tag{33}$$

Then we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{F}(t)\| \leq \|\mathbf{Y}(t)\|. \tag{34}$$

Proof. Writing the definition of the derivative we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{F}\| &= \lim_{h\to 0} \frac{\|\mathbf{F}(t+h)\| - \|\mathbf{F}(t)\|}{h} \\
&= \lim_{h\to 0} \frac{\|\mathbf{F}(t) + ih[\mathbf{X}(t), \mathbf{F}(t)] + h\mathbf{Y}(t)\| + O(h^2) - \|\mathbf{F}(t)\|}{h} \\
&= \lim_{h\to 0} \frac{\|(I + ih\mathbf{X})\mathbf{F}(I - ih\mathbf{X}) + ih^2\mathbf{X}\mathbf{F}\mathbf{X} + h\mathbf{Y}\| - \|\mathbf{F}\|}{h} \\
&= \lim_{h\to 0} \frac{\|\,\mathrm{e}^{ih\mathbf{X}}\,\mathbf{F}\,\mathrm{e}^{-ih\mathbf{X}} + ih^2\mathbf{X}\mathbf{F}\mathbf{X} + h\mathbf{Y}\| + O(h^2) - \|\mathbf{F}\|}{h} \\
&\leq \lim_{h\to 0} \frac{\|\,\mathrm{e}^{ih\mathbf{X}}\,\mathbf{F}\,\mathrm{e}^{-ih\mathbf{X}}\,\| - \|\mathbf{F}\|}{h} + h\|\mathbf{X}\mathbf{F}\mathbf{X}\| + \|\mathbf{Y}\| \\
&= \|\mathbf{Y}\|\,,
\end{aligned} \tag{35}$$

where in the fourth line we use $\mathrm{e}^{\pm ih\mathbf{X}} = I \pm ih\mathbf{X} + O(h^2)$, and in the last line we use the fact that $\mathbf{X}$ is Hermitian and $e^{\pm ih\mathbf{X}}$ is unitary. □

Here is our second building block of the proof.

Lemma 2. *Using the notation of Theorem 1 and assuming that $L > 1$ we have*

$$\|\mathbf{U}(t)^\dagger\mathbf{O}\mathbf{U}(t) - \mathbf{V}(t)^\dagger\mathbf{O}\mathbf{V}(t)\| \leq 2\|\mathbf{O}_A\| \sum_{e\sim A}\int_0^t |u_e(\tau)| \cdot \|\mathbf{U}^\dagger(\tau)\mathbf{h}_e\mathbf{U}(\tau) - \mathbf{V}(\tau)^\dagger\mathbf{h}_e\mathbf{V}(\tau)\|\mathrm{d}\tau\,, \tag{36}$$

where $\widetilde{\mathbf{H}}_A(t) := \sum_{e\subseteq A\cup\partial_L(A)} u_e(t)\mathbf{h}_e$ is the Hamiltonian consisting of terms in the region $A\cup\partial_L(A)$, and $\mathbf{V}(t)$ is the unitary evolution associated to $\widetilde{\mathbf{H}}_A(t)$. Also, by $e\sim A$ we mean e intersects A. Moreover, if $A = e_0$ for some edge e_0 and $\mathbf{O}_A = \mathbf{h}_{e_0}$, then the sum in Eq. (36) can be further restricted to $e \neq e_0$.

Proof. Let

$$\mathbf{F}(t) := \mathbf{U}(t)^\dagger\mathbf{O}\mathbf{U}(t) - \mathbf{V}(t)^\dagger\mathbf{O}\mathbf{V}(t). \tag{37}$$

Then from Schrodinger's equation we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{F}(t) &= i\mathbf{U}(t)^\dagger[\mathbf{H}(t),\mathbf{O}]\mathbf{U}(t) - i\mathbf{V}(t)^\dagger[\widetilde{\mathbf{H}}(t),\mathbf{O}]\mathbf{V}(t)\\ &= \sum_{e\sim A} i\mathbf{U}(t)^\dagger[\mathbf{h}_e(t),\mathbf{O}]\mathbf{U}(t) - i\mathbf{V}(t)^\dagger[\mathbf{h}_e(t),\mathbf{O}]\mathbf{V}(t)\\ &= \sum_{e\sim A} i[\mathbf{U}(t)^\dagger\mathbf{h}_e(t)\mathbf{U}(t), \mathbf{U}(t)^\dagger\mathbf{O}\mathbf{U}(t)^\dagger] - i[\mathbf{V}(t)^\dagger\mathbf{h}_e(t)\mathbf{V}(t), \mathbf{V}(t)^\dagger\mathbf{O}\mathbf{V}(t)]\\ &= \sum_{e\sim A} i[\mathbf{U}(t)^\dagger\mathbf{h}_e(t)\mathbf{U}(t), \mathbf{F}(t) + \mathbf{V}(t)^\dagger\mathbf{O}\mathbf{V}(t)^\dagger] - i[\mathbf{V}(t)^\dagger\mathbf{h}_e(t)\mathbf{V}(t), \mathbf{V}(t)^\dagger\mathbf{O}\mathbf{V}(t)]\\ &= i\Big[\sum_{e\sim A}\mathbf{U}(t)^\dagger\mathbf{h}_e(t)\mathbf{U}(t), \mathbf{F}(t)\Big] + i\sum_{e\sim A}\left[\mathbf{U}(t)^\dagger\mathbf{h}_e(t)\mathbf{U}(t) - \mathbf{V}(t)^\dagger\mathbf{h}_e(t)\mathbf{V}(t), \mathbf{V}(t)^\dagger\mathbf{O}\mathbf{V}(t)\right].\end{aligned} \tag{38}$$

Therefore, using $\mathbf{F}(0) = 0$ and Lemma 1 we can bound

$$\begin{aligned}\|\mathbf{F}(t)\| &= \Big\|\int_0^t \frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{F}(\tau)\mathrm{d}\tau\Big\|\\ &\le \int_0^t \Big\|\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{F}(\tau)\Big\|\mathrm{d}\tau\\ &\le \int_0^t \sum_{e\sim A} \|[\mathbf{U}(\tau)^\dagger\mathbf{h}_e(\tau)\mathbf{U}(\tau) - \mathbf{V}(\tau)^\dagger\mathbf{h}_e(\tau)\mathbf{V}(\tau), \mathbf{V}(\tau)^\dagger\mathbf{O}\mathbf{V}(\tau)]\|\mathrm{d}\tau\\ &\le 2\|\mathbf{O}\|\sum_{e\sim A}\int_0^t |u_e(\tau)|\cdot\|\mathbf{U}(\tau)^\dagger\mathbf{h}_e\mathbf{U}(\tau) - \mathbf{V}(\tau)^\dagger\mathbf{h}_e\mathbf{V}(\tau)\|\mathrm{d}\tau\,,\end{aligned} \tag{39}$$

where we use $\|[\mathbf{X},\mathbf{Y}]\| \le 2\|\mathbf{X}\|\|\mathbf{Y}\|$ in the last line. Note that if $\mathbf{O}_A = \mathbf{h}_{e_0}$, in the second line of Eq. (38) and thereafter, we can further restrict the sum to $e \neq e_0$.

□

We can now give the proof of Theorem 1. As before, let $\mathbf{F}(t) := \mathbf{U}(t)^\dagger\mathbf{O}\mathbf{U}(t) - \mathbf{V}(t)^\dagger\mathbf{O}\mathbf{V}(t)$. By Lemma 2 we have

$$\|\mathbf{F}(T)\| \le 2\|\mathbf{O}\|\sum_{e_1\sim A}\int_0^T |u_{e_1}(t_1)|\cdot\|\mathbf{U}(t_1)^\dagger\mathbf{h}_{e_1}\mathbf{U}(t_1) - \mathbf{V}(t_1)^\dagger\mathbf{h}_{e_1}\mathbf{V}(t_1)\|\mathrm{d}t_1\,. \tag{40}$$

Assuming that $L > 2$, we can apply Lemma 2 on each of the terms in the above sum, in which $A = e_1$ and $\mathbf{O}_A$ is replaced by $\mathbf{h}_{e_1}$. Therefore,

$$\begin{aligned}\|\mathbf{F}(T)\| &\le 2\|\mathbf{O}_A\|\sum_{e_1\sim A}\int_0^T 2\|u_{e_1}(t)\mathbf{h}_{e_1}\|\int_0^{t_1}\sum_{\substack{e_2\sim e_1\\ e_2\neq e_1}}|u_{e_2}(t_2)|\cdot\|\mathbf{U}(t_2)^\dagger\mathbf{h}_{e_2}\mathbf{U}(t_2) - \mathbf{V}(t_2)^\dagger\mathbf{h}_{e_2}\mathbf{V}(t_2)\|\mathrm{d}t_2\mathrm{d}t_1\\ &\le 2\|\mathbf{O}_A\|\,2g\sum_{e_1\sim A}\sum_{\substack{e_2\sim e_1\\ e_2\neq e_1}}\int_0^T\int_0^{t_1}|u_{e_2}(t_2)|\cdot\|\mathbf{U}(t_2)^\dagger\mathbf{h}_{e_2}\mathbf{U}(t_2) - \mathbf{V}(t_2)^\dagger\mathbf{h}_{e_2}\mathbf{V}(t_2)\|\mathrm{d}t_2\mathrm{d}t_1\,.\end{aligned} \tag{41}$$

We can repeat the same process for L times to obtain

$$\begin{aligned}
&\frac{\|\mathbf{F}(T)\|}{2\|\mathbf{O}_A\|(2g)^{L-1}} \\
&\leq \sum_{\substack{(e_1,e_2,\ldots,e_L):\\ \text{walk starting at } A}} \int_0^T \int_0^{t_1} \cdots \int_0^{t_{L-1}} |u_{e_L}(t_L)| \cdot \|\mathbf{U}(t_L)^\dagger \mathbf{h}_{e_L} \mathbf{U}(t_L) - \mathbf{V}(t_L)^\dagger \mathbf{h}_{e_2} \mathbf{V}(t_L)\| \mathrm{d}t_L \ldots \mathrm{d}t_2 \mathrm{d}t_1 \\
&\leq \sum_{\substack{(e_1,e_2,\ldots,e_L):\\ \text{walk starting at } A}} \int_0^T \int_0^{t_1} \cdots \int_0^{t_{L-1}} 2g \ \mathrm{d}t_L \ldots \mathrm{d}t_2 \mathrm{d}t_1 \\
&= 2g \frac{T^L}{L!} N_L \,,
\end{aligned} \tag{42}$$

where N_L denotes the *number of walks* of length L starting at A, i.e., N_L is the number of sequences of edges $(e_1, \ldots, e_L)$ such that $e_1 \sim A$, $e_i \sim e_{i+1}$ and $e_i \neq e_{i+1}$ for $i = 1, \ldots, L-1$. We note that as the maximum degree of the graph G is Δ, the number of choices for e_1 is at most $\Delta|A| \leq 2(\Delta-1)|A|$. Moreover, the number of choices for e_{i+1} given e_i is at most $2(\Delta-1)$. Thus $N_L \leq (2(\Delta-1))^L |A|$. Therefore, using the Stirling's approximation we have

$$\begin{aligned}
\|\mathbf{F}(T)\| &\leq 2|A| \ \|\mathbf{O}_A\| \frac{\big(4g(\Delta-1)T\big)^L}{L!} \\
&\leq 2|A| \cdot \|\mathbf{O}_A\| \big(4g(\Delta-1)T\big)^L \frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{-L\log L + L - \frac{1}{2}\log L} \\
&\leq \sqrt{\frac{2}{\pi}} |A| \cdot \|\mathbf{O}_A\| \, \mathrm{e}^{-L\left(\log L - \log T - \log\left(4g(\Delta-1)\right)\right) - \frac{1}{2}\log L} \,.
\end{aligned} \tag{43}$$

B Proof of Theorem 2

Let $\boldsymbol{\zeta} = \sum_i \mathbf{Z}_i$ and use Markov's inequality to obtain

$$\Pr\left[\left| w_H(\mathbf{x}) - \left(\frac{n}{2} - \frac{m}{2}\right)\right| \geq c n^{\frac{1}{2}+\kappa_2}\right] = \Pr\left[\left|\langle \mathbf{x}| \boldsymbol{\zeta} |\mathbf{x}\rangle - m\right| \geq 2c n^{\frac{1}{2}+\kappa_2}\right] \tag{44}$$

$$\leq \frac{1}{\left(2c n^{\frac{1}{2}+\kappa_2}\right)^2} \mathbb{E}\left[\left(\langle \mathbf{x}| \boldsymbol{\zeta} |\mathbf{x}\rangle - m\right)^2\right] . \tag{45}$$

Thus to prove the theorem it suffices to bound $\mathbb{E}\big[\big(\langle \mathbf{x}| \boldsymbol{\zeta} |\mathbf{x}\rangle - m\big)^2\big]$.

Let $m_i = \langle \psi_T| \mathbf{Z}_i |\psi_T\rangle$. Then we have

$$\begin{aligned}
&\mathbb{E}\left[\left(\langle \mathbf{x}| \boldsymbol{\zeta} |\mathbf{x}\rangle - m\right)^2\right] \\
&= \mathbb{E}\left[\left(\sum_i \langle \mathbf{x}| \mathbf{Z}_i |\mathbf{x}\rangle - m_i\right)^2\right] \\
&= \sum_{i,j} \mathbb{E}\big[\big(\langle \mathbf{x}| \mathbf{Z}_i |\mathbf{x}\rangle - m_i\big) \cdot \big(\langle \mathbf{x}| \mathbf{Z}_j |\mathbf{x}\rangle - m_j\big)\big] \\
&= \sum_{i,j} \langle \psi_T| (\mathbf{Z}_i - m_i)(\mathbf{Z}_j - m_j) |\psi_T\rangle \\
&= \sum_{\substack{i,j \\ \mathrm{dist}_G(i,j)\leq 2L}} \langle \psi_T| (\mathbf{Z}_i - m_i)(\mathbf{Z}_j - m_j) |\psi_T\rangle + \sum_{\substack{i,j \\ \mathrm{dist}_G(i,j)>2L}} \langle \psi_T| (\mathbf{Z}_i - m_i)(\mathbf{Z}_j - m_j) |\psi_T\rangle \,,
\end{aligned} \tag{46}$$

where in the third line we use the fact that $\mathbf{Z}_i$ is diagonal in the computational basis, $\mathrm{dist}_G(i,j)$ denotes the distance of vertices i,j on the graph G, and L is given by

$$L = \frac{\kappa_1}{2\log\Delta}\log n \geq 4g\Delta^{\frac{2-\kappa_1}{\kappa_1}}T. \tag{47}$$

We analyze terms in the above sum separately. If $\mathrm{dist}_G(i,j) \leq 2L$, then by the Cauchy-Schwarz inequality we have

$$\begin{aligned}
\langle\psi_T|\left(\mathbf{Z}_i - m_i\right)\left(\mathbf{Z}_j - m_j\right)|\psi_T\rangle &\leq \sqrt{\langle\psi_T|\left(\mathbf{Z}_i - m_i\right)^2|\psi_T\rangle \cdot \langle\psi_T|\left(\mathbf{Z}_j - m_j\right)^2|\psi_T\rangle} && (48)\\
&= \sqrt{(1-m_i^2)(1-m_j^2)} && (49)\\
&\leq 1\,. && (50)
\end{aligned}$$

Next, by the Lieb-Robinson bound (Theorem 1), for any i there is an operator $\tilde{\mathbf{Z}}_i$ that acts only on $\partial_L(\{i\})$ and

$$\|\mathbf{U}(T)^\dagger(\mathbf{Z}_i - m_i)\mathbf{U}(T) - \tilde{\mathbf{Z}}_i\| \leq \mathrm{e}^{-L\left(\log L - \log T - \log 4g\Delta\right)} \leq \mathrm{e}^{-L\left(2\frac{1-\kappa_1}{\kappa_1}\log\Delta\right)} = n^{-(1-\kappa_1)}\,, \tag{51}$$

where the second inequality follows from (47). Therefore, letting $\boldsymbol{\Upsilon}_i = \mathbf{U}(T)^\dagger(\mathbf{Z}_i - m_i)\mathbf{U}(T) - \tilde{\mathbf{Z}}_i$, if $\mathrm{dist}_G(i,j) > 2L$ we have

$$\begin{aligned}
&\langle\psi_T|\left(\mathbf{Z}_i - m_i\right)\left(\mathbf{Z}_j - m_j\right)|\psi_T\rangle\\
&= \langle\psi_0|\,\mathbf{U}(T)^\dagger\left(\mathbf{Z}_i - m_i\right)\left(\mathbf{Z}_j - m_j\right)\mathbf{U}(T)\,|\psi_0\rangle\\
&= \langle\psi_0|\left(\tilde{\mathbf{Z}}_i + \boldsymbol{\Upsilon}_i\right)\left(\tilde{\mathbf{Z}}_j + \boldsymbol{\Upsilon}_j\right)|\psi_0\rangle\\
&= \langle\psi_0|\,\tilde{\mathbf{Z}}_i\tilde{\mathbf{Z}}_j\,|\psi_0\rangle + \langle\psi_0|\,\boldsymbol{\Upsilon}_i\tilde{\mathbf{Z}}_j\,|\psi_0\rangle + \langle\psi_0|\,\tilde{\mathbf{Z}}_i\boldsymbol{\Upsilon}_j\,|\psi_0\rangle + \langle\psi_0|\,\boldsymbol{\Upsilon}_i\boldsymbol{\Upsilon}_j\,|\psi_0\rangle\\
&= \langle\psi_0|\,\tilde{\mathbf{Z}}_i\,|\psi_0\rangle\,\langle\psi_0|\,\tilde{\mathbf{Z}}_j\,|\psi_0\rangle + \langle\psi_0|\,\boldsymbol{\Upsilon}_i\tilde{\mathbf{Z}}_j\,|\psi_0\rangle + \langle\psi_0|\,\tilde{\mathbf{Z}}_i\boldsymbol{\Upsilon}_j\,|\psi_0\rangle + \langle\psi_0|\,\boldsymbol{\Upsilon}_i\boldsymbol{\Upsilon}_j\,|\psi_0\rangle\,,
\end{aligned} \tag{52}$$

where in the last line we use the facts that $|\psi_0\rangle$ is a product state and that the supports of $\tilde{\mathbf{Z}}_i, \tilde{\mathbf{Z}}_j$ do not intersect. Now we have

$$\begin{aligned}
|\,\langle\psi_0|\,\tilde{\mathbf{Z}}_i\,|\psi_0\rangle\,| &= |\,\langle\psi_T|\left(\mathbf{Z}_i - m_i\right)|\psi_T\rangle - \langle\psi_0|\,\boldsymbol{\Upsilon}_i\,|\psi_0\rangle\,| && (53)\\
&= |\,\langle\psi_0|\,\boldsymbol{\Upsilon}_i\,|\psi_0\rangle\,| && (54)\\
&\leq \|\boldsymbol{\Upsilon}_i\| && (55)\\
&\leq n^{-(1-\kappa_1)}. && (56)
\end{aligned}$$

Using this, and similar inequalities, we find that

$$\langle\psi_T|\left(\mathbf{Z}_i - m_i\right)\left(\mathbf{Z}_j - m_j\right)|\psi_T\rangle \leq 2\left(n^{-(1-\kappa_1)}\right)^2 + 2n^{-(1-\kappa_1)} \leq 4n^{-(1-\kappa_1)}. \tag{57}$$

Putting these together yields

$$\begin{aligned}
\mathbb{E}\left[\left(\langle\mathbf{x}|\,\boldsymbol{\zeta}\,|\mathbf{x}\rangle - m\right)^2\right] &\leq \sum_{\substack{i,j\\ \mathrm{dist}_G(i,j)\leq 2L}} 1 + \sum_{\substack{i,j\\ \mathrm{dist}_G(i,j)>2L}} 4n^{-(1-\kappa_1)} && (58)\\
&\leq 2n\Delta^{2L} + n^2(4n^{-(1-\kappa_1)}) && (59)\\
&= 6n^{1+\kappa_1}, && (60)
\end{aligned}$$

where in the second line we use

$$1 + \Delta\left(1 + (\Delta-1) + (\Delta-1)^2 + \cdots(\Delta-1)^{2L-1}\right) \leq 2\Delta^{2L}. \tag{61}$$

Using this in (45) the desired inequality follows.

C Derivative of a matrix function

In this appendix we compute a bound on the derivative of a certain polynomial function of matrices. This bound is an important ingredient in the proof of Theorem 3. Indeed, in the proof, applying the Lieb-Robinson bound, a desired operator is approximated by a local one. Then, a polynomial function is applied on this operator, and the blow-up of the error of the Lieb-Robinson approximation induced by this polynomial should be estimated. We note that a naive approach to estimate this error is ineffective since the degree of this polynomial is large and some of its coefficients are exponentially large. Therefore, we need a more careful analysis of this error which itself demands a more careful analysis of the derivative of that polynomial as a map on the space of matrices.

We first need to fix some notations. Given two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same dimensions, we denote their *entry-wise product* by $\mathbf{A} \circ \mathbf{B}$. In other words, $\mathbf{A} \circ \mathbf{B}$ is a matrix whose (i,j)-th entry equals the product of (i,j)-th entries of $\mathbf{A}$ and $\mathbf{B}$, i.e., $(\mathbf{A} \circ \mathbf{B})_{ij} = \mathbf{A}_{ij} \cdot \mathbf{B}_{ij}$. The matrix $\mathbf{A} \circ \mathbf{B}$ is called the Schur or Hadamard product of $\mathbf{A}$ and $\mathbf{B}$.

For a continuously differentiable function $f : (a,b) \to \mathbb{R}$ and a sequence $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_d)$ of numbers in (a,b) we let $\mathcal{D}_{f,\boldsymbol{\lambda}}$ be the $d \times d$ matrix whose (i,j)-th entry equals

$$\left(\mathcal{D}_{f,\boldsymbol{\lambda}}\right)_{ij} = \begin{cases} \frac{f(\lambda_i) - f(\lambda_j)}{\lambda_i - \lambda_j} & \lambda_i \neq \lambda_j, \\ f'(\lambda_i) & \lambda_i = \lambda_j, \end{cases} \tag{62}$$

where $f'(\lambda_i)$ is the derivative of f at λ_i. For a diagonal matrix $\boldsymbol{\Lambda} = \mathrm{diag}(\boldsymbol{\lambda})$ with diagonal elements in $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_d)$ we let

$$\mathcal{D}_{f,\boldsymbol{\Lambda}} := \mathcal{D}_{f,\boldsymbol{\lambda}}. \tag{63}$$

Also, for a self-adjoint matrix $\mathbf{A}$, that can be diagonalized as $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^\dagger$ where $\mathbf{U}$ is unitary and $\boldsymbol{\Lambda} = \mathrm{diag}(\boldsymbol{\lambda})$ is a diagonal matrix containing eigenvalues of $\mathbf{A}$, we let

$$\mathcal{D}_{f,\mathbf{A}} = \mathbf{U}\mathcal{D}_{f,\boldsymbol{\Lambda}}\mathbf{U}^\dagger. \tag{64}$$

Lemma 3. *[51, pp. 124] Let $\mathbf{A}, \mathbf{E}$ be two Hermitian matrices and f be a real continuously differentiable function on an interval containing the eigenvalues of $\mathbf{A} + t\mathbf{E}$ for all $t \in (-1,1)$. Then we have*

$$\frac{\mathrm{d}}{\mathrm{d}t} f\left(\mathbf{A} + t\mathbf{E}\right)\Big|_{t=0} = \mathcal{D}_{f,\mathbf{A}} \circ \mathbf{E}\,. \tag{65}$$

By this lemma, to bound the derivative of a matrix function as above, we need to bound the norm of the Schur product as a super-operator on the space of matrices. This super-operator norm is given by the so called γ_2-norm.

Lemma 4. *[52, pp. 79] For a $d \times d$ matrix $\mathbf{M}$ let*

$$\gamma_2(\mathbf{M}) := \inf \max\{\| \left|v_i\right\rangle \|^2 : 1 \leq i \leq d\} \cup \{\| \left|w_j\right\rangle \|^2 : 1 \leq j \leq d\}, \tag{66}$$

where the infimum is taken over all vectors $\left|v_i\right\rangle, \left|w_j\right\rangle$ satisfying $\mathbf{M}_{ij} = \langle v_i \left| w_j \right\rangle$ for all i,j. Then for any $d \times d$ matrix $\mathbf{E}$ we have

$$\|\mathbf{M} \circ \mathbf{E}\| \leq \gamma_2(\mathbf{M}) \|\mathbf{E}\|. \tag{67}$$

It is not hard to verify that γ_2-norm as defined in Eq. (66) satisfies the triangle inequality and is a norm. We also note that if $\mathbf{M}_{ij} = \langle v_i \left| w_j \right\rangle$ and $\mathbf{M}'_{ij} = \langle v'_i \left| w'_j \right\rangle$ for two matrices $\mathbf{M}$ and $\mathbf{M}'$, then we have $(\mathbf{M} \circ \mathbf{M}')_{ij} = \langle \tilde{v}_i \left| \tilde{w}_j \right\rangle$ where $\left|\tilde{v}_i\right\rangle = \left|v_i\right\rangle \otimes \left|v'_i\right\rangle$ and $\left|\tilde{w}_j\right\rangle = \left|w_j\right\rangle \otimes \left|w'_j\right\rangle$. Therefore,

$$\gamma_2(\mathbf{M} \circ \mathbf{M}') \leq \gamma_2(\mathbf{M}) \cdot \gamma_2(\mathbf{M}'). \tag{68}$$

The following lemma is a simple consequence of the lemmas above.

Lemma 5. *For any two Hermitian matrices* $\mathbf{A}, \mathbf{E}$ *and polynomial* P*, we have*

$$\|P(\mathbf{A}+\mathbf{E}) - P(\mathbf{A})\| \leq \|\mathbf{E}\| \cdot \int_0^1 \gamma_2(\mathcal{D}_{P,\mathbf{A}+t\mathbf{E}})\mathrm{d}t\,. \tag{69}$$

Proof. By Lemma 3 we have

$$\|P(\mathbf{A}+\mathbf{E}) - P(\mathbf{A})\| = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} P(\mathbf{A}+t\mathbf{E})\mathrm{d}t = \int_0^1 \left(\mathcal{D}_{P,\mathbf{A}+t\mathbf{E}} \circ \mathbf{E}\right) \mathrm{d}t \tag{70}$$

Then the result follows using Lemma 5 and the triangle inequality. □

In the proof of Theorem 3 we will use this lemma for a polynomial that is defined in terms of the Chebyshev's polynomials. Recall that the Chebyshev's polynomial $T_n(x)$ is defined by

$$T_n(\cos\theta) = \cos(n\theta). \tag{71}$$

This equation shows that $|T_n(x)| \leq 1 = T_n(1)$ if $|x| \leq 1$. Moreover, $T_n(x)$ as a degree n polynomial, has n roots in the interval $[-1, 1]$. Thus, it is monotone in the interval $[1, \infty)$ and we have $|T_n(x)| \leq T_n(1+\delta)$ for $x \in [-1, 1+\delta]$. The Chebyshev polynomials can also be defined by recursive equations: we have $T_0(x) = 1$, $T_1(x) = x$ and

$$T_{2n+1}(x) = 2T_{n+1}(x)T_n(x) - x, \tag{72}$$

$$T_{2n}(x) = 2T_n^2(x) - 1\,. \tag{73}$$

Lemma 6. *For any* $0 \leq \delta < 1$ *and any sequence* $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_d)$ *with* $-1 \leq \lambda_i \leq 1+\delta$ *we have*

$$\gamma_2(\mathcal{D}_{T_n,\boldsymbol{\lambda}}) \leq (2n^2-1)T_n(1+\delta), \qquad \forall n \geq 1\,. \tag{74}$$

Proof. We prove the lemma by induction on n. For $n = 1$ it can be easily shown by setting $T_1(x) = x$.

For an even number $2n$ we can write:

$$\begin{aligned}
(\mathcal{D}_{T_{2n},\boldsymbol{\lambda}})_{ij} &= \frac{T_{2n}(\lambda_i) - T_{2n}(\lambda_j)}{\lambda_i - \lambda_j} && (75)\\
&= 2\frac{T_n^2(\lambda_i) - T_n^2(\lambda_j)}{\lambda_i - \lambda_j} && (76)\\
&= 2\frac{(T_n(\lambda_i) - T_n(\lambda_j))\,(T_n(\lambda_i) + T_n(\lambda_j))}{\lambda_i - \lambda_j} && (77)\\
&= 2\,(\mathcal{D}_{T_n,\boldsymbol{\lambda}})_{ij}\,(T_n(\lambda_i) + T_n(\lambda_j)) && (78)\\
&= 2\,(\mathcal{D}_{T_n,\boldsymbol{\lambda}})_{ij}\left(R_{T_n,\boldsymbol{\lambda}} + R^{\intercal}_{T_n,\boldsymbol{\lambda}}\right)_{ij}\,, && (79)
\end{aligned}$$

where $\left(\mathcal{R}_{T_n,\boldsymbol{\lambda}}\right)_{ij} = T_n(\lambda_i)$, and $R^{\intercal}_{T_n,\boldsymbol{\lambda}}$ is its transpose. Now we can write:

$$\begin{aligned}
\gamma_2(\mathcal{D}_{T_{2n},\boldsymbol{\lambda}}) &\leq 2\gamma_2(\mathcal{D}_{T_n,\boldsymbol{\lambda}})\gamma_2(R_{T_n,\boldsymbol{\lambda}} + R^{\intercal}_{T_n,\boldsymbol{\lambda}})\,,\\
&\leq 2\left((2n^2-1)T_n(1+\delta)\right)\left(\gamma_2(R_{T_n,\boldsymbol{\lambda}}) + \gamma_2(R^{\intercal}_{T_n,\boldsymbol{\lambda}})\right)\,.
\end{aligned} \tag{80}$$

Using $|T_n(x)| \leq T_n(1+\delta)$ for $x \in [-1, 1+\delta]$, we can find that

$$\gamma_2(R_{T_n,\boldsymbol{\lambda}}), \gamma_2(R^{\intercal}_{T_n,\boldsymbol{\lambda}}) \leq T_n(1+\delta)\,. \tag{81}$$

As a result

$$\gamma_2(\mathcal{D}_{T_{2n},\boldsymbol{\lambda}}) \leq 4(2n^2-1)T_n^2(1+\delta)\,. \tag{82}$$

Now recall that $T_n^2(1+\delta) = \frac{T_{2n}(1+\delta)+1}{2}$ and knowing that $T_{2n}(1+\delta) \geq 1$

$$\begin{aligned}\gamma_2(\mathcal{D}_{T_{2n},\boldsymbol{\lambda}}) &\leq 2(2n^2-1)\left(T_{2n}(1+\delta)+1\right), \\ &\leq 4(2n^2-1)T_{2n}(1+\delta), \\ &\leq \left(2(2n)^2-1\right)T_{2n}(1+\delta).\end{aligned} \tag{83}$$

Similarly for an odd number $2n+1$, we can write:

$$\left(\mathcal{D}_{T_{2n+1},\boldsymbol{\lambda}}\right)_{ij} = 2\frac{T_n(\lambda_i)T_{n+1}(\lambda_i) - T_n(\lambda_j)T_{n+1}(\lambda_j)}{\lambda_i - \lambda_j} - 1 \tag{84}$$

$$= 2\frac{T_n(\lambda_i)T_{n+1}(\lambda_i) - T_n(\lambda_j)T_{n+1}(\lambda_i) + T_n(\lambda_j)T_{n+1}(\lambda_i) - T_n(\lambda_j)T_{n+1}(\lambda_j)}{\lambda_i - \lambda_j} - 1 \tag{85}$$

$$= 2\left(\left(\mathcal{D}_{T_n,\boldsymbol{\lambda}}\right)_{ij} T_{n+1}(\lambda_i) + T_n(\lambda_j)\left(\mathcal{D}_{T_{n+1},\boldsymbol{\lambda}}\right)_{ij}\right) - 1 \tag{86}$$

$$= 2\left(\left(\mathcal{D}_{T_n,\boldsymbol{\lambda}}\right)_{ij}\left(\left(R_{T_{n+1},\boldsymbol{\lambda}}\right)_{ij} + \left(R^{\intercal}_{T_n,\boldsymbol{\lambda}}\right)\left(\mathcal{D}_{T_{n+1},\boldsymbol{\lambda}}\right)_{ij}\right) - 1\right. . \tag{87}$$

Now by bounding $\gamma_2(R_{T_n,\boldsymbol{\lambda}})$ and $\gamma_2(R_{T_{n+1},\boldsymbol{\lambda}})$

$$\begin{aligned}\gamma_2(\mathcal{D}_{T_{2n+1},\boldsymbol{\lambda}}) &\leq 2\gamma_2(\mathcal{D}_{T_n,\boldsymbol{\lambda}})\gamma_2\left(\left(R_{T_{n+1},\boldsymbol{\lambda}}\right) + 2\gamma_2(\mathcal{D}_{T_{n+1},\boldsymbol{\lambda}})\gamma_2\left(R^{\intercal}_{T_n,\boldsymbol{\lambda}}\right) + 1\right., \\ &\leq 2(2n^2-1)T_n(1+\delta)\gamma_2(R_{T_{n+1},\boldsymbol{\lambda}}) + \\ &\quad 2\left(2(n+1)^2-1\right)T_{n+1}(1+\delta)\gamma_2(R^{\intercal}_{T_n,\boldsymbol{\lambda}}) + 1, \\ &\leq 2(2n^2-1)T_n(1+\delta)T_{n+1}(1+\delta) + \\ &\quad 2\left(2(n+1)^2-1\right)T_{n+1}(1+\delta)T_n(1+\delta) + 1, \\ &\leq 2\left((2n+1)^2-1\right)T_n(1+\delta)T_{n+1}(1+\delta) + 1.\end{aligned} \tag{88}$$

Next using $2T_n(1+\delta)T_{n+1}(1+\delta) = T_{2n+1}(1+\delta) + 1 + \delta$ and the fact that $T_{2n+1}(1+\delta) \geq 1+\delta$ we obtain

$$\begin{aligned}\gamma_2(\mathcal{D}_{T_{2n+1},\boldsymbol{\lambda}}) &\leq \left((2n+1)^2-1\right)\left(T_{2n+1}(1+\delta)+1+\delta\right)+1, \\ &\leq \left(2(2n+1)^2-1\right)T_{2n+1}(1+\delta).\end{aligned} \tag{89}$$

□

We can now state the main lemma that will be used in the proof of Theorem 3.

Lemma 7. *For any n define the polynomial $C_n(x)$ by*

$$C_n(x) := 1 - \frac{T_n(f(x))}{T_n(f(0))}, \tag{90}$$

where

$$f(x) = \frac{1+\epsilon-2x}{1-\epsilon}, \tag{91}$$

and $0 \leq \epsilon \leq 1/3$, then for any sequence $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_d)$ with $0 \leq \lambda_i \leq 1$ we have

$$\gamma_2(\mathcal{D}_{C_n,\boldsymbol{\lambda}}) \leq 6n^2 - 3 \tag{92}$$

Proof. One can easily verify that

$$\begin{aligned}\mathcal{D}_{C_n,\boldsymbol{\lambda}} &= -\frac{1}{T_n(f(0))}\mathcal{D}_{T_n\circ f,\boldsymbol{\lambda}} \\ &= -\frac{1}{T_n(f(0))}\mathcal{D}_{T_n,f(\boldsymbol{\lambda})} \circ \mathcal{D}_{f,\boldsymbol{\lambda}} \\ &= -\frac{1}{T_n(f(0))}\mathcal{D}_{T_n,f(\boldsymbol{\lambda})} \circ \left(\frac{-2}{1-\epsilon}\mathbf{J}\right)\end{aligned} \tag{93}$$

Where $\mathbf{J}$ is an all-one matrix.

Then we introduce $\delta := \frac{2\epsilon}{1-\epsilon}$ and we have $f(0) = 1 + \delta$

$$\begin{aligned}\gamma_2(\mathcal{D}_{C_n,\boldsymbol{\lambda}}) &= -\frac{1}{T_n(1+\delta)}\gamma_2(\mathcal{D}_{T_n,f(\boldsymbol{\lambda})})(\frac{-2}{1-\epsilon}) \\ &\leq \frac{1}{T_n(1+\delta)}\gamma_2(\mathcal{D}_{T_n,f(\boldsymbol{\lambda})})(\frac{2}{1-\frac{1}{3}})\end{aligned} \tag{94}$$

Knowing $f(\lambda_i) \leq 1 + \delta$ and $0 \leq \delta \leq 1$, by applying Lemma 6

$$\begin{aligned}\gamma_2(\mathcal{D}_{C_n,\boldsymbol{\lambda}}) &\leq \frac{3}{T_n(1+\delta)}(2n^2-1)T_n(1+\delta) \\ \gamma_2(\mathcal{D}_{C_n,\boldsymbol{\lambda}}) &\leq 6n^2 - 3\end{aligned} \tag{95}$$

□

D A general isoperimetric inequality

In this appendix we prove Theorem 3. This is a general isoperimetric inequality that will be used in the proof of Theorem 4.

Proof. Let N_1 be the set of qubits within distance L from qubits $N_0 = \{1, \ldots, n_0\}$ in the graph, i.e.,

$$N_1 = N_0 \cup \partial_L(N_0). \tag{96}$$

Observe that using (61) we have $n_1 = |N_1| \leq \min\{2n_0\Delta^L, n\}$. For simplicity of presentation we assume that $N_1 = \{1, \ldots, n_1\}$. Define

$$\mathbf{P}_0 := |s_1\rangle\langle s_1| \otimes \cdots |s_{n_1}\rangle\langle s_{n_1}| \otimes \mathbf{I}^{\otimes(n-n_1)}, \tag{97}$$

and

$$\boldsymbol{\Gamma}_0 := \frac{1}{n_1}\sum_{i=1}^{n_1} |\bar{s}_i\rangle\langle\bar{s}_i|_i\,, \tag{98}$$

where $|\bar{s}_i\rangle$ is the qubit state orthogonal to $|s_i\rangle$ and $|\bar{s}_i\rangle\langle\bar{s}_i|_i$ is the operator acting on the i-qubit by projecting on $|\bar{s}_i\rangle$.

Let $T_m(x)$ for $m = \frac{1}{2}n_1^{\frac{1}{2}-\theta}$ be the Chebyshev polynomial of degree m given by Eq. (71). Let

$$f(x) = \frac{1 + \frac{1}{n_1} - 2x}{1 - \frac{1}{n_1}}, \tag{99}$$

and define

$$C_m(x) := 1 - \frac{T_m(f(x))}{T_m(f(0))}. \tag{100}$$

Next, define

$$\mathbf{K} := \mathbf{U}C_m(\boldsymbol{\Gamma}_0)\mathbf{U}^\dagger = C_m(\mathbf{U}\boldsymbol{\Gamma}_0\mathbf{U}^\dagger)\,, \tag{101}$$

where $\mathbf{U} = \mathbf{U}(T)$. We note that $C_m(0) = 0$, and for $0 \leq x \leq 1$ we have

$$0 \leq C_m(x) \leq 1 + \frac{1}{T_m(f(0))} \leq 2. \tag{102}$$

Moreover, as shown in [31] we have

$$C_m(x) \geq 1 - \frac{1}{1 + 2\frac{m^2}{n_1}} \geq \frac{m^2}{n_1} = \frac{1}{4}n_1^{-2\theta}, \qquad \frac{1}{n_1} \leq x \leq 1. \tag{103}$$

Note that the spectrum of $\mathbf{\Gamma}_0$, regardless of multiplicities, equals $\{0, \frac{1}{n_1}, \frac{2}{n_1} \ldots, 1\}$. Thus, by the above properties of $C_m(x)$, the spectrum of $\mathbf{K}$ belongs to $\{0\} \cup [n_1^{-2\theta}/4, 2]$. Moreover, the projection on the eigenspace of $\mathbf{K}$ associated to the 0-eigenvalue equals $\mathbf{P} = \mathbf{U}\mathbf{P}_0\mathbf{U}^\dagger$. Therefore,

$$\frac{1}{4} n_1^{-2\theta} (\mathbf{I} - \mathbf{P}) \preceq \mathbf{K} \preceq 2(\mathbf{I} - \mathbf{P}). \tag{104}$$

The next step is to approximate $\mathbf{K}$ with a sum of local operators. Based on Theorem 1, for $1 \le i \le n_1$ let $\mathbf{Q}_i$ be an L-local operator such that

$$\left\| \mathbf{Q}_i - \mathbf{U} \left| \bar{s}_i \right\rangle\!\left\langle \bar{s}_i \right|_i \mathbf{U}^\dagger \right\| \le \epsilon(L), \tag{105}$$

where

$$\epsilon(L) = \sqrt{\frac{2}{\pi}}\, \mathrm{e}^{-L\left(\log L - \log T - \log(4g(\Delta-1))\right) - \frac{1}{2}\log L}. \tag{106}$$

Indeed, $\mathbf{Q}_i$ acts on qubits within distance L from i in the graph. Then letting

$$\tilde{\mathbf{\Gamma}} := \frac{1}{n_1} \sum_{i=1}^{n_1} \mathbf{Q}_i, \tag{107}$$

we have

$$\left\| \tilde{\mathbf{\Gamma}} - \mathbf{U}\mathbf{\Gamma}_0\mathbf{U}^\dagger \right\| \le \epsilon(L). \tag{108}$$

We claim that $\tilde{\mathbf{K}} = C_m(\tilde{\mathbf{\Gamma}})$ is close to $\mathbf{K} = C_m(\mathbf{U}\mathbf{\Gamma}_0\mathbf{U}^\dagger)$. To prove this we use Lemma 5 and Lemma 7. To apply the latter lemma we need to have a bound on the spectrum of $(1-r)\mathbf{U}\mathbf{\Gamma}_0\mathbf{U}^\dagger + r\tilde{\mathbf{\Gamma}}$. Here, a crucial observation is that by construction of $\mathbf{Q}_i$ in the proof of Theorem 1 we have $0 \preceq \mathbf{Q}_i \preceq \mathbf{I}$ which gives $0 \preceq \tilde{\mathbf{\Gamma}} \preceq 1$. On the other hand, by definition $0 \preceq \mathbf{\Gamma}_0 \preceq \mathbf{I}$ which yields $0 \preceq (1-r)\mathbf{U}\mathbf{\Gamma}_0\mathbf{U}^\dagger + r\tilde{\mathbf{\Gamma}} \preceq \mathbf{I}$ for $0 \le r \le 1$. As a result, by Lemma 5 and Lemma 7 we have

$$\left\| \widetilde{\mathbf{K}} - \mathbf{K} \right\| \le (6m^2 - 3)\epsilon(L) \le 6m^2\epsilon(L). \tag{109}$$

Now recall that $\tilde{\mathbf{\Gamma}}$ is a sum of L-local operators, and $C_m(x)$ is a degree m polynomial. As a result $\tilde{\mathbf{K}} = C_m(\tilde{\mathbf{\Gamma}})$ is a sum of ℓ'-local operators where

$$\ell' = mL.$$

Next, as in [31] we partition $\{0,1\}^n$ into four sets:

$$F_1 := \left(F \setminus \partial_{\ell'}(F)\right) \times \{0,1\}^{n-n_0}, \tag{110}$$
$$F_2 := \left(F \cap \partial_{\ell'}(F)\right) \times \{0,1\}^{n-n_0}, \tag{111}$$
$$F_3 := \left(F^c \cap \partial_{\ell'}(F)\right) \times \{0,1\}^{n-n_0}, \tag{112}$$
$$F_4 := \left(F^c \setminus \partial_{\ell'}(F)\right) \times \{0,1\}^{n-n_0}. \tag{113}$$

Here, $\partial_{\ell'}(F)$ is defined with respect to the Hamming distance on $\{0,1\}^{n_0}$. Then, expanding $|\psi\rangle = |\psi_T\rangle$ in the computational basis, we can write

$$|\psi\rangle = |\phi_1\rangle + |\phi_2\rangle + |\phi_3\rangle + |\phi_4\rangle\,, \tag{114}$$

where $|\phi_j\rangle$ is a linear combination of $|\mathbf{x}\rangle$'s with $\mathbf{x} \in F_j$. The point of this decomposition is that $p(\partial_{\ell'}(F)) = \|\phi_2\|^2 + \|\phi_3\|^2$. On the other hand, since $\widetilde{\mathbf{K}}$ is ℓ'-local we have

$$\langle\phi_1|\widetilde{\mathbf{K}}|\phi_3\rangle = \langle\phi_1|\widetilde{\mathbf{K}}|\phi_4\rangle = \langle\phi_2|\widetilde{\mathbf{K}}|\phi_4\rangle = 0. \tag{115}$$

Therefore, because $\mathbf{\Gamma}_0\,|\psi_0\rangle = 0$, we have

$$\langle\psi|\mathbf{K}|\psi\rangle = \langle\psi|\,\mathbf{U}C_m(\mathbf{\Gamma}_0)\mathbf{U}^\dagger\,|\psi\rangle = \langle\psi_0|\,C_m(\mathbf{\Gamma}_0)\,|\psi_0\rangle = 0. \tag{116}$$

Thus by Eq. (109) we have

$$|\langle\psi|\tilde{\mathbf{K}}|\psi\rangle| \leq 6m^2\epsilon(L). \tag{117}$$

Let $|\psi'\rangle := -|\phi_1\rangle - |\phi_2\rangle + |\phi_3\rangle + |\phi_4\rangle$. Then using Eq. (115) we have

$$\langle\psi'|\tilde{\mathbf{K}}|\psi'\rangle = \sum_{(i,j)\in\{1,2\}^2\cup\{3,4\}^2} \langle\phi_i|\tilde{\mathbf{K}}|\phi_j\rangle - \sum_{(i,j)\in\{1,2\}\times\{3,4\}} \langle\phi_i|\tilde{\mathbf{K}}|\phi_j\rangle \tag{118}$$

$$= \||\phi_1\|^2 + \||\phi_2\|^2 + \||\phi_3\|^2 + \||\phi_4\|^2 + 2\,\mathrm{Re}\,\langle\phi_1|\tilde{\mathbf{K}}|\phi_2\rangle + 2\,\mathrm{Re}\,\langle\phi_3|\tilde{\mathbf{K}}|\phi_4\rangle - 2\,\mathrm{Re}\,\langle\phi_2|\tilde{\mathbf{K}}|\phi_3\rangle\,. \tag{119}$$

Considering the same expansion for $\langle\psi|\tilde{\mathbf{K}}|\psi\rangle$ and using Eq. (117), we obtain

$$\begin{aligned}\langle\psi'|\tilde{\mathbf{K}}|\psi'\rangle &\leq 4|\langle\phi_2|\tilde{\mathbf{K}}|\phi_3\rangle| + 6m^2\epsilon(L)\\ &\leq 4\|\tilde{\mathbf{K}}\|\cdot\||\phi_2\rangle\|\cdot\||\phi_3\rangle\| + 6m^2\epsilon(L)\\ &\leq 4\left(\||\phi_2\rangle\|^2 + \||\phi_3\rangle\|^2\right) + 6m^2\epsilon(L)\\ &= 4p(\partial_{\ell'}(F)) + 6m^2\epsilon(L),\end{aligned} \tag{120}$$

where in the third line we use $0 \preceq \tilde{\mathbf{\Gamma}} \preceq \mathbf{I}$ which gives $\|\tilde{\mathbf{K}}\| = \|C_m(\tilde{\mathbf{\Gamma}})\| \leq 2$.

We now prove a lower bound on $\langle\psi'|\tilde{\mathbf{K}}|\psi'\rangle$ in terms of $p(F)$. To this end we will use Eq. (104), so we first prove an upper bound on $\langle\psi'|\mathbf{P}|\psi'\rangle$. Define

$$\mathbf{R}_0 := \sum_{\mathbf{x}\in F^c} |\mathbf{x}\rangle\langle\mathbf{x}| \otimes \mathbf{I} - \sum_{\mathbf{x}\in F} |\mathbf{x}\rangle\langle\mathbf{x}| \otimes \mathbf{I}. \tag{121}$$

Note that $\mathbf{R}_0|\psi\rangle = |\psi'\rangle$ and that $\mathbf{R}_0$ is an n_0-local operator acting on the first n_0 qubits. Then by Theorem 1 there is $\tilde{\mathbf{R}}$ acting on qubits in N_1 such that

$$\|\tilde{\mathbf{R}} - \mathbf{R}_0\| \leq n_0\epsilon(L). \tag{122}$$

Then we have

$$\mathbf{P}_0\mathbf{U}^\dagger|\psi'\rangle = \mathbf{P}_0\mathbf{U}^\dagger\mathbf{R}_0|\psi\rangle \tag{123}$$

$$= \mathbf{P}_0\mathbf{U}^\dagger\mathbf{R}_0\mathbf{U}|\psi_0\rangle \tag{124}$$

$$= \mathbf{P}_0\tilde{\mathbf{R}}|\psi_0\rangle + |\delta\rangle\,, \tag{125}$$

where $|\delta\rangle$ is a vector with $\||\delta\rangle\| \leq n_0\epsilon(L)$. On the other hand, since $\tilde{\mathbf{R}}$ acts only on the first n_1 qubits and $|\psi_0\rangle = |s_0\rangle\cdots|s_n\rangle$, the vector $\tilde{\mathbf{R}}|\psi_0\rangle$ is a linear combination of vectors of the form $|\mathbf{x}\rangle\otimes|s_{n_1+1}\rangle\cdots|s_n\rangle$ for $\mathbf{x}\in\{0,1\}^{n_1}$. Therefore, applying the projection $\mathbf{P}_0$ on $\tilde{\mathbf{R}}|\psi_0\rangle$ we get a vector parallel to $|\psi_0\rangle$, i.e., $\mathbf{P}_0\tilde{\mathbf{R}}|\psi_0\rangle = \alpha|\psi_0\rangle$ for some $\alpha\in\mathbb{C}$. This means that

$$\mathbf{P}_0\mathbf{U}^\dagger|\psi'\rangle = \alpha|\psi_0\rangle + |\delta\rangle\,. \tag{126}$$

To compute α we write

$$\alpha = \langle\psi_0|\mathbf{P}_0\mathbf{U}^\dagger|\psi'\rangle - \langle\psi_0|\delta\rangle \tag{127}$$

$$= \langle\psi_0|\mathbf{P}_0\mathbf{U}^\dagger\mathbf{R}\mathbf{U}|\psi_0\rangle - \langle\psi_0|\delta\rangle \tag{128}$$

$$= \langle\psi_0|\mathbf{U}^\dagger\mathbf{R}\mathbf{U}|\psi_0\rangle - \langle\psi_0|\delta\rangle \tag{129}$$

$$= \langle\psi|\mathbf{R}|\psi\rangle - \langle\psi_0|\delta\rangle \tag{130}$$

$$= \langle\psi|\psi'\rangle - \langle\psi_0|\delta\rangle \tag{131}$$

$$= -\||\phi_1\rangle\|^2 - \||\phi_2\rangle\|^2 + \||\phi_3\rangle\|^2 + \||\phi_4\rangle\|^2 - \langle\psi_0|\delta\rangle \tag{132}$$

$$= (1-2p(F)) - \langle\psi_0|\delta\rangle, \tag{133}$$

where in the last line we have used $p(F) = \| \, |\phi_1\rangle \, \|^2 + \| \, |\phi_2\rangle \, \|^2$. Putting these together we arrive at

$$\mathbf{P}_0\mathbf{U}^\dagger \, |\psi'\rangle = \big((1-2p(F)) - \langle\psi_0|\delta\rangle\big) \, |\psi_0\rangle + |\delta\rangle \, . \tag{134}$$

Therefore,

$$\begin{aligned}
\langle\psi'| \, \mathbf{P} \, |\psi'\rangle &= \langle\psi'| \, \mathbf{U}\mathbf{P}_0\mathbf{U}^\dagger \, |\psi'\rangle && (135)\\
&= \|\mathbf{P}_0\mathbf{U}^\dagger \, |\psi'\rangle \, \|^2 && (136)\\
&\leq \big|(1-2p(F)) - \langle\psi_0|\delta\rangle\big|^2 + \| \, |\delta\rangle \, \|^2 + 2\big|(1-2p(F)) - \langle\psi_0|\delta\rangle\big| \cdot \| \, |\delta\rangle \, \| && (137)\\
&\leq (1-2p(F))^2 + 8\| \, |\delta\rangle \, \| && (138)\\
&\leq (1-2p(F))^2 + 8n_0\epsilon(L). && (139)
\end{aligned}$$

Next, using Eq. (104) we find that

$$\begin{aligned}
\langle\psi'| \, \mathbf{K} \, |\psi'\rangle &\geq \frac{1}{4}n_1^{-2\theta}\big(1 - \langle\psi'| \, \mathbf{P} \, |\psi'\rangle \, \big) && (140)\\
&\geq \frac{1}{4}n_1^{-2\theta}\Big(1 - (1-2p(F))^2 - 8n_0\epsilon(L)\Big) && (141)\\
&\geq \frac{1}{2}n_1^{-2\theta}p(F) - 2n_1^{-2\theta}n_0\epsilon(L), && (142)
\end{aligned}$$

where in the last line we used $p(F) \leq 1/2$ to conclude that $1-(1-2p(F))^2 \geq 2p(F)$. Then by Eq. (109) we have

$$\big| \, \langle\psi'| \, \tilde{\mathbf{K}} \, |\psi'\rangle \, \big| \geq \frac{1}{2}n_1^{-2\theta}p(F) - \big(2n_1^{-2\theta}n_0 + 6m^2\big)\epsilon(L). \tag{143}$$

Finally, comparing this inequality with Eq. (120) yields

$$4p(\partial_{\ell'}(F)) + 6m^2\epsilon(L) \geq \frac{1}{2}n_1^{-2\theta}p(F) - \big(2n_1^{-2\theta}n_0 + 6m^2\big)\epsilon(L), \tag{144}$$

or equivalently

$$p(\partial_{\ell'}(F)) \geq \frac{1}{8}n_1^{-2\theta}p(F) - \Big(\frac{1}{2}n_1^{-2\theta}n_0 + 3m^2\Big)\epsilon(L). \tag{145}$$

Now for the first part of the theorem let $L = \frac{k_1}{\log\Delta}\log n_0 = \beta 4g\Delta T$ with

$$\beta = \frac{\kappa_1}{4g\Delta\log(\Delta)T}\log n_0 \geq \Delta^{\frac{1+\kappa_1+\kappa_2}{\kappa_1}}, \tag{146}$$

where the inequality follows from the assumption

$$T \leq \frac{\kappa_1}{4g\Delta\log(\Delta)\Delta^{\frac{1+2\kappa_1+\kappa_2}{\kappa_1}}}\log n_0. \tag{147}$$

Then we have $n_1 \leq 2n_0\Delta^L = 2n_0^{1+\kappa_1}$ and $m = \frac{1}{2}n_1^{\frac{1}{2}-\theta} \leq \frac{1}{\sqrt{2}}n_0^{\frac{1}{2}(1+\kappa_1)}$. We also have

$$\begin{aligned}
\epsilon(L) &= \sqrt{\frac{2}{\pi}}\,\mathrm{e}^{-L\left(\log L - \log T - \log(4g(\Delta-1))\right) - \frac{1}{2}\log L} && (148)\\
&\leq \mathrm{e}^{-L\log\beta} && (149)\\
&= n_0^{-\frac{\kappa_1}{\log(\Delta)}\log\beta} && (150)\\
&\leq n_0^{-(1+\kappa_1+\kappa_2)}. && (151)
\end{aligned}$$

Therefore,

$$\Big(\frac{1}{2}n_1^{-2\theta}n_0 + 3m^2\Big)\epsilon(L) \le 2\, n_0^{1+\kappa_1} n_0^{-(1+\kappa_1+\kappa_2)} = 2n_0^{-\kappa_2}. \tag{152}$$

On the other hand,

$$\ell' = mL \le \frac{1}{2}(2n_0^{1+\kappa_1})^{\frac{1}{2}-\theta}\frac{\kappa_1}{\log(\Delta)}\log n_0 \le \ell. \tag{153}$$

We conclude that

$$p(\partial_\ell(F)) \ge p(\partial_{\ell'}(F)) \ge \frac{1}{8}n_1^{-2\theta}p(F) - 2n_0^{-\kappa_2} \ge \frac{1}{8}(2n_0^{1+\kappa_1})^{-2\theta}p(F) - 2n_0^{-\kappa_2}. \tag{154}$$

For the second part of the theorem, note that when $n_0 = n$, then $n_1 = n$. Using this in (145) for $\theta = 0$, we obtain

$$p(\partial_{\ell'}) \ge \frac{1}{8}p(F) - \frac{5}{4}n\epsilon(L). \tag{155}$$

Let $L = \kappa_1(1+\kappa_2)\log n$ and note that $\ell' = mL = \frac{1}{2}\sqrt{n}L = \ell$. On the other hand, using the given bound on T we find that $\epsilon(L) \le n^{-\kappa_1(1+\kappa_2)\log(1+\kappa_2)}$. Putting this in the above inequality yields the desired result.

□

E Proof Theorem 4

For an integer $0 \le d < D/\ell$ define $K_d \subseteq \{0,1\}^n$ by

$$K_d = \{\mathbf{x} : (d-1)\ell < \mathrm{dist}_H(\mathbf{x}, F_1) \le d\ell\}, \tag{156}$$

where $\mathrm{dist}_H(\mathbf{x}, F_1) = \min_{\mathbf{y}\in F_1}\mathrm{dist}_H(\mathbf{x},\mathbf{y})$. Clearly, $K_0 = F_1$ and $K_d \cap F_2 = \emptyset$ for all $d < D/\ell$. Then, since K_d's are disjoint, we have

$$\sum_{\substack{1\le d<\frac{D}{\ell}-1\\ d:\text{odd}}} p(K_d) + p(K_{d+1}) \le 1 - \big(p(F_1) + p(F_2)\big) \le 1 - 2\mu\,. \tag{157}$$

Therefore, there exists d_0 such that

$$p(K_{d_0}) + p(K_{d_0+1}) \le \frac{1-2\mu}{\frac{D}{2\ell}-2}\,. \tag{158}$$

Now define

$$\widetilde{F}_1 := \bigcup_{j\le d_0} K_d\,. \tag{159}$$

Observe that by definition $F_1 \subseteq \widetilde{F}_1$ and that $\partial_\ell(\widetilde{F}_1) \subseteq K_{d_0} \cup K_{d_0+1}$. Therefore, using Theorem 3 we obtain

$$\begin{aligned}
\frac{1-2\mu}{\frac{D}{2\ell}-2} &\ge p(K_{d_0}) + p(K_{d_0+1}) && (160)\\
&\ge p(\partial_\ell \widetilde{F}_1) && (161)\\
&\ge \frac{1}{8}p(\widetilde{F}_1) - \frac{5}{4}n^{-\left(\kappa_1(1+\kappa_2)\log(1+\kappa_2)-1\right)} && (162)\\
&\ge \frac{\mu}{8} - \frac{5}{4}n^{-\left(\kappa_1(1+\kappa_2)\log(1+\kappa_2)-1\right)}, && (163)
\end{aligned}$$

and

$$\frac{D}{2\ell} - 2 \leq \frac{1-2\mu}{\mu/8} + \frac{5}{4} n^{-\left(\kappa_1(1+\kappa_2)\log(1+\kappa_2)-1\right)} \left(\frac{\frac{D}{2\ell}-2}{\mu/8}\right) \tag{164}$$

$$< \frac{8}{\mu} - 2 + \frac{5}{4} n^{-\left(\kappa_1(1+\kappa_2)\log(1+\kappa_2)-1\right)} \left(\frac{\frac{D}{2\ell}}{\mu/8}\right). \tag{165}$$

Multiplying both sides by 2ℓ yields the desired inequality.

F Proof of Theorem 5

We follow similar steps as in [28] to prove this theorem. Let $\mathbf{x} \in \{0,1\}^n$ be the measurement outcome on $|\psi_T\rangle$ in the computational basis. Let $\mathrm{Cut}(\mathbf{x})$ be the size of the cut associated to $\mathbf{x}$:

$$\mathrm{Cut}(\mathbf{x}) = -\langle \mathbf{x}| \mathbf{C} |\mathbf{x}\rangle . \tag{166}$$

Then we have

$$-\langle \psi_T| \mathbf{C} |\psi_T\rangle = \mathbb{E}[\mathrm{Cut}(\mathbf{x})]. \tag{167}$$

Here a crucial observation is that $\mathrm{Cut}(\mathbf{x}) = \mathrm{Cut}(\bar{\mathbf{x}})$ where $\bar{\mathbf{x}} \in \{0,1\}^n$ is obtained from $\mathbf{x}$ by flipping each coordinate: $\bar{\mathbf{x}}_i = 1 - \mathbf{x}_i$. On the other hand, the distribution of $\bar{\mathbf{x}}$ is the same as the distribution of $\mathbf{x}$. This is because $\mathbf{X}^{\otimes n} |\psi_T\rangle = |\psi_T\rangle$ which itself can be proven using the fact that the starting state of QA satisfies $\mathbf{X}^{\otimes n} |\psi_0\rangle = |\psi_0\rangle$ and that $\mathbf{X}^{\otimes n}$ commutes with $\mathbf{H}(t)$ for any t.

Let $\mathbf{x}^*$ be the configuration associated to the bipartition of G. Then we have $\mathrm{Cut}^* = \mathrm{Cut}(\mathbf{x}^*) = \mathrm{Cut}(\bar{\mathbf{x}}^*) = |E| = \Delta n/2$. Let $d = \alpha n$ and define

$$F_1 = \{\mathbf{x} : \mathrm{dist}_H(\mathbf{x}, \mathbf{x}^*) \leq d\}, \tag{168}$$

where as before, $\mathrm{dist}_H(\mathbf{x}, \mathbf{x}^*)$ denotes the hamming distance between $\mathbf{x}$ and $\mathbf{x}^*$. Similarly, define

$$F_2 = \{\mathbf{x} : \mathrm{dist}_H(\mathbf{x}, \bar{\mathbf{x}}^*) \leq d\} = \{\mathbf{x} : \bar{\mathbf{x}} \in F_1\}. \tag{169}$$

Now using Appendix E and the fact that $D := \mathrm{dist}_H(F_1, F_2) \geq n - 2d$ we have

$$p(F_1) = p(F_2) = \min\{p(F_1), p(F_2)\} \leq \frac{16\ell}{D} + 10n^{-\kappa_3}, \tag{170}$$

where the first equality follows from the aforementioned symmetry and ℓ is given by Eq. (15) for some $\kappa_2 > 0$ and

$$\kappa_3 = \kappa_1(1+\kappa_2)\log(1+\kappa_2) - 1. \tag{171}$$

Suppose that $\mathbf{x}$ is not in $F_1 \cup F_2$. This means that either $d < \mathrm{dist}_H(\mathbf{x}, \mathbf{x}^*) \leq n/2$ or $d < \mathrm{dist}_H(\mathbf{x}, \bar{\mathbf{x}}^*) \leq n/2$. Then, e.g., in the first case, using the definition of the Cheeger constant we have

$$\mathrm{Cut}(\mathbf{x}) \leq |E| - h(G)d < \mathrm{Cut}(\mathbf{x}^*). \tag{172}$$

This means that, letting

$$J = \{\mathbf{x} : \mathrm{Cut}(\mathbf{x}) \geq |E| - h(G)d\}, \tag{173}$$

we have $J \subseteq F_1 \cup F_2$ and

$$p(J) \leq p(F_1) + p(F_2) \leq 2\Big(\frac{16\ell}{D} + 10n^{-\kappa_3}\Big). \tag{174}$$

Therefore,

$$p\Big(|E| - \mathrm{Cut}(\mathbf{x}) > h(G)d\Big) \geq 1 - \frac{32\ell}{D} - 20n^{-\kappa_3}. \tag{175}$$

Next, using Markov's inequality we have

$$\mathbb{E}\big[|E| - \mathrm{Cut}(\mathbf{x})\big] \geq h(G)d\left(1 - \frac{32\ell}{D} - 20n^{-\kappa_3}\right). \tag{176}$$

Now, note that $D \geq n - 2d = (1-2\alpha)n$ and that for *sufficiently large* n we have

$$1 - \frac{32\ell}{D} - 20n^{-\kappa_3} \geq 1 - \epsilon. \tag{177}$$

Then, for any such n we have

$$\mathbb{E}[\mathrm{Cut}(\mathbf{x})] \leq |E| - \alpha(1-\epsilon)h(G)n. \tag{178}$$

Thus, using $|E| = \frac{n\Delta}{2}$ and $h(G) \geq \frac{\Delta - 2\sqrt{\Delta-1}}{2}$ we obtain

$$\begin{aligned}\frac{\mathbb{E}[\mathrm{Cut}(\mathbf{x})]}{\mathrm{Cut}^*} &\leq 1 - \alpha(1-\epsilon)\Big(\frac{n\Delta}{2|E|} - \frac{n\sqrt{\Delta-1}}{|E|}\Big)\\ &= 1 - \alpha(1-\epsilon) + 2\alpha(1-\epsilon)\frac{\sqrt{\Delta-1}}{\Delta}.\end{aligned} \tag{179}$$

G Proof of Theorem 6

Let $\mathbb{G}_\Delta(n)$ be a random Δ-regular graph on n vertices and $\mathbb{G}^B_\Delta(n)$ be a bipartite random Δ-regular graph on n vertices. Let L be a parameter that grows at most logarithmically with n. Then, it is known [53] that there exists a constant α_Δ independent of n and L such that

$$\mathbb{E}_{\mathbb{G}_\Delta(n)}\left[\#\text{ of cycles of length at most } 2L+1\right] \leq \alpha_\Delta.\Delta^{2L+1}, \tag{180}$$

$$\mathbb{E}_{\mathbb{G}^B_\Delta(n)}\left[\#\text{ of cycles of length at most } 2L+1\right] \leq \alpha_\Delta.\Delta^{2L+1}. \tag{181}$$

Given a graph $G = (V, E)$ let $R_L(G)$ be the set of edges in G that have a cycle in their L-neighborhood:

$$R_L(G) = \{e \in E \,|\, \exists \text{ a cycle in the } L\text{-neighborhood of } e\}.$$

It is easy to verify that if e belong to $R_L(G)$, then there is cycle of length at most $2L+1$ in the L-neighborhood of e. On the other hand, for any such cycle, the number of edges in its L-neighborhood is at most:

$$(2L+1)\Big[1 + \Delta + \Delta(\Delta-1) + \cdots \Delta(\Delta-1)^{L-1}\Big] \leq 2(2L+1)\Delta^L. \tag{182}$$

Therefore, by the above bound on the expectation of the number of cycles of length at most $2L+1$ we have

$$\mathbb{E}_{\mathbb{G}_\Delta(n)}\left[|R_L(\mathbb{G}_\Delta(n))|\right] \leq 2\alpha_\Delta(2L+1)\Delta^{3L+1} \leq \alpha_\Delta n^\kappa, \tag{183}$$

$$\mathbb{E}_{\mathbb{G}^B_\Delta(n)}\left[\Big|R_L(\mathbb{G}^B_\Delta(n))\Big|\right] \leq 2\alpha_\Delta(2L+1)\Delta^{3L+1} \leq \alpha_\Delta n^\kappa. \tag{184}$$

Let

$$L = \frac{\kappa}{4\log\Delta}\log n. \tag{185}$$

For an edge $e = \{i, j\}$ in G define

$$\mathcal{E}(e, L) = \Big\langle \psi_0 \Big| \mathbf{V}_e(T)^\dagger \mathbf{C}_e \mathbf{V}_e(T) \Big| \psi_0 \Big\rangle, \tag{186}$$

where

$$\mathbf{C}_e = \frac{1}{2}(\mathbf{I} - \mathbf{Z}_i\mathbf{Z}_j), \tag{187}$$

and $\mathbf{V}_e(T)$ is the unitary evolution associated to the Hamiltonian that includes the terms of $\mathbf{H}(t)$ that correspond to edges/vertices at distance at most L from e. Note that by Theorem 1 we have

$$\|\boldsymbol{\Upsilon}\| \leq \epsilon(L) = \sqrt{\frac{2}{\pi}}\, \mathrm{e}^{-L\left(\log L - \log T - \log(4\Delta)\right) - \frac{1}{2}\log L}\,, \tag{188}$$

where

$$\boldsymbol{\Upsilon}_e = \mathbf{U}(T)^\dagger \mathbf{C}_e \mathbf{U}(T) - \mathbf{V}_e(T)^\dagger \mathbf{C}_e \mathbf{V}_e(T)\,. \tag{189}$$

On the other hand, observe that the L-neighborhood of any edge e not in $R_L(G)$ is a tree, and that this is a fixed tree. Then, since $\mathbf{V}_e(T)$ depends only on the structure of the graph around e, and $|\psi_0\rangle$ is symmetric, for any $e \notin R_L(G)$, the quantity

$$\mathcal{E}(e, L) = \mathcal{E}_{\mathrm{tree}}(L)\,, \tag{190}$$

is independent of e. Thus, using the fact that $\|\mathbf{C}_e\| \leq 1$, we have

$$\begin{aligned}
-\langle\psi_T|\mathbf{C}|\psi_T\rangle &= -\left\langle\psi_0\middle|\mathbf{U}(T)^\dagger\mathbf{C}\mathbf{U}(T)\middle|\psi_0\right\rangle && (191)\\
&= \sum_{e\notin R_L(G)} \left\langle\psi_0\middle|\mathbf{U}(T)^\dagger\mathbf{C}_e\mathbf{U}(T)\middle|\psi_0\right\rangle \\
&\quad + \sum_{e\in R_L(G)} \left\langle\psi_0\middle|\mathbf{U}(T)^\dagger\mathbf{C}_e\mathbf{U}(T)\middle|\psi_0\right\rangle && (192)\\
&\leq |E|\cdot\left(\mathcal{E}_{\mathrm{tree}}(L) + \epsilon(L)\right) + |R_L(G)| && (193)\\
&\leq \frac{n\Delta}{2}\left(\mathcal{E}_{\mathrm{tree}}(L) + \epsilon(L)\right) + |R_L(G)|\,. && (194)
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\mathbb{E}_{\mathbb{G}_\Delta^B(n)}\left[-\langle\psi_T|\mathbf{C}|\psi_T\rangle\right] \\
&\qquad \leq \frac{n\Delta}{2}\left(\mathcal{E}_{\mathrm{tree}}(L) + \epsilon(L)\right) + 2\alpha_\Delta(2L+1)\Delta^{3L+1}\,. && (195)
\end{aligned}$$

On the other hand, we also have

$$\begin{aligned}
-\langle\psi_T|\mathbf{C}|\psi_T\rangle &= \sum_{e\notin R_L(G)} \left\langle\psi_0\middle|\mathbf{U}(T)^\dagger\mathbf{C}_e\mathbf{U}(T)\middle|\psi_0\right\rangle \\
&\quad + \sum_{e\in R_L(G)} \left\langle\psi_0\middle|\mathbf{U}(T)^\dagger\mathbf{C}_e\mathbf{U}(T)\middle|\psi_0\right\rangle && (196)\\
&\geq \left(|E| - |R_L(G)|\right)\left(\mathcal{E}_{\mathrm{tree}}(L) - \epsilon(L)\right), && (197)
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}_{\mathbb{G}_\Delta(n)}\left[-\langle\psi_T|\mathbf{C}|\psi_T\rangle\right] \\
&\qquad \geq \left(\frac{n\Delta}{2} - 2\alpha_\Delta(2L+1)\Delta^{3L+1}\right)\left(\mathcal{E}_{\mathrm{tree}}(L) - \epsilon(L)\right). && (198)
\end{aligned}$$

It is known that for any Δ, there exists a constant ρ_Δ such that

$$\mathbb{E}_{\mathbb{G}_\Delta(n)}\left[\text{size of MaxCut}\right] = \rho_\Delta n + o(n)\,. \tag{199}$$

Therefore, since $-\langle\psi_T|\mathbf{C}|\psi_T\rangle$ itself is an expectation of sizes of certain cuts, we have

$$\mathbb{E}_{\mathbb{G}_\Delta(n)}\left[-\langle\psi_T|\mathbf{C}|\psi_T\rangle\right] \leq \rho_\Delta n + o(n)\,. \tag{200}$$

Comparing this with Eq. (198) yields

$$\Delta\left(\frac{n}{2}-2\alpha_\Delta(2L+1)\Delta^{3L}\right)\left(\mathcal{E}_{\text{tree}}(L)-\epsilon(L)\right)\leq\rho_\Delta n+o(n)\,. \tag{201}$$

Next, using the definition of L and the bound on T, we find that for any $\nu>0$ and sufficiently large n we have

$$\frac{\Delta}{2}\mathcal{E}_{\text{tree}}\leq\rho_\Delta+\nu\,, \tag{202}$$

with

$$\mathcal{E}_{\text{tree}}=\limsup_{n\to\infty}\mathcal{E}_{\text{tree}}(L)\,, \tag{203}$$

where we note that L can grow with n. Using this in (195) we find that for sufficiently large n

$$\mathbb{E}_{\mathbb{G}_\Delta^B(n)}\big[-\langle\psi_T|\mathbf{C}|\psi_T\rangle\big]\leq(\rho_\Delta+2\nu)n\,. \tag{204}$$

It is known that $\rho_3\leq 1.4026<\frac{3}{2}$, and that for large Δ [54, 55, 56]

$$\rho_\Delta\leq\frac{\Delta}{4}+O(\sqrt{\Delta})\,. \tag{205}$$

Thus, for sufficiently large n, short time QA on a random bipartite graphs gives a cut of size of at most

$$\left(\frac{\Delta}{4}+O(\sqrt{\Delta})+2\nu\right)n\,, \tag{206}$$

which is far from the optimal value of $\frac{\Delta}{2}n$.