THE COMPLEXITY OF THE LOCAL HAMILTONIAN PROBLEM*

JULIA KEMPE[†], ALEXEI KITAEV[‡], AND ODED REGEV[§]

Abstract. The k-LOCAL HAMILTONIAN problem is a natural complete problem for the complexity class QMA, the quantum analogue of NP. It is similar in spirit to MAX-k-SAT, which is NP-complete for $k \geq 2$. It was known that the problem is QMA-complete for any $k \geq 3$. On the other hand, 1-LOCAL HAMILTONIAN is in P and hence not believed to be QMA-complete. The complexity of the 2-LOCAL HAMILTONIAN problem has long been outstanding. Here we settle the question and show that it is QMA-complete. We provide two independent proofs; our first proof uses only elementary linear algebra. Our second proof uses a powerful technique for analyzing the sum of two Hamiltonians; this technique is based on perturbation theory and we believe that it might prove useful elsewhere. Using our techniques we also show that adiabatic computation with 2-local interactions on qubits is equivalent to standard quantum computation.

Key words. quantum computation, local Hamiltonian problem, complete problems, adiabatic computation

AMS subject classifications. 81P68, 68Q17

DOI. 10.1137/S0097539704445226

1. Introduction. Quantum complexity theory has emerged alongside the first efficient quantum algorithms in an attempt to formalize the notion of an *efficient* algorithm. In analogy to classical complexity theory, several new quantum complexity classes have appeared. A major challenge today consists in understanding their structure and the interrelation between classical and quantum classes.

One of the most important classical complexity classes is NP—nondeterministic polynomial time. This class comprises languages that can be *verified* in polynomial time by a deterministic verifier. The celebrated Cook–Levin theorem (see, e.g., [17]) shows that this class has *complete* problems. More formally, it states that SAT is NP-complete; i.e., it is in NP and any other language in NP can be reduced to it with polynomial overhead. In SAT we are given a set of clauses (disjunctions) over n variables and asked whether there is an assignment that satisfies all clauses. One can consider the restriction of SAT in which each clause consists of at most k literals.

This is known as the k-SAT problem. It is known that 3-SAT is still NP-complete while 2-SAT is in P, i.e., has a polynomial time solution. We can also consider the MAX-k-SAT problem: here, given a k-SAT formula and a number m we are asked whether there exists an assignment that satisfies at least m clauses. It turns out that MAX-2-SAT is already NP-complete; MAX-1-SAT is clearly in P.

The class QMA is the quantum analogue of NP in a probabilistic setting, i.e., the class of all languages that can be probabilistically verified by a quantum verifier in polynomial time (the name is derived from the classical class MA, which is the randomized analogue of NP). This class, which is also called BQNP, was first studied in [13, 12]; the name QMA was given to it by Watrous [20]. Several problems in QMA have been identified [20, 12, 9]. For a good introduction to the class QMA, see the book by Kitaev, Shen, and Vyalvi [12] and the paper by Watrous [20].

Kitaev, inspired by ideas due to Feynman, defined the quantum analogue of the classical SAT problem, the LOCAL HAMILTONIAN problem [12].[1] An instance of k-LOCAL HAMILTONIAN can be viewed as a set of local constraints on n qubits, each involving at most k of them. We are asked whether there is a state of the n qubits such that the expected number of violated constraints is either below a certain threshold or above another, with a promise that one of the two cases holds and both thresholds are at least a constant apart. More formally, we are to determine whether the *groundstate* energy of a given k-local Hamiltonian is below one threshold or above another.

Kitaev proved [12] that the 5-LOCAL HAMILTONIAN problem is QMA-complete. Later, Kempe and Regev showed that even 3-LOCAL HAMILTONIAN is complete for QMA [11]. In addition, it is easy to see that 1-LOCAL HAMILTONIAN is in P. The complexity of the 2-LOCAL HAMILTONIAN problem was left as an open question in [2, 21, 11, 7]. It is not hard to see that the k-LOCAL HAMILTONIAN problem contains the MAX-k-SAT problem as a special case.[2] Using the known NP-completeness of MAX-2-SAT, we obtain that 2-LOCAL HAMILTONIAN is NP-hard; i.e., any problem in NP can be reduced to it with polynomial overhead. But is it also QMA-complete? Or perhaps it lies in some intermediate class between NP and QMA? Some special cases of the problem were considered by Bravyi and Vyalyi [7]; however, the question still remained open.

In this paper we settle the question of the complexity of 2-LOCAL HAMILTONIAN and show that the following theorem holds.

THEOREM 1.1. *The* 2-LOCAL HAMILTONIAN *problem is* QMA-*complete.*

In [12] it was shown that the k-LOCAL HAMILTONIAN problem is in QMA for any constant k (and in fact even for $k = O(\log n)$, where n is the total number of qubits). Hence, our task in this paper is to show that any problem in QMA can be reduced to the 2-LOCAL HAMILTONIAN problem with a polynomial overhead. We give two self-contained proofs for this.

Our first proof is based on a careful selection of gates in a quantum circuit and several applications of a lemma called the *projection lemma*. The proof is quite involved; however, it uses only elementary linear algebra and hence might appeal to some readers.

---

[1]For a survey of the LOCAL HAMILTONIAN problem, see [2].

[2]The idea is to represent the n variables by n qubits and represent each clause by a Hamiltonian. Each Hamiltonian is diagonal and acts on the k variables that appear in its clause. It "penalizes" the assignment that violates the clause by increasing its eigenvalue. Therefore, the lowest eigenvalue of the sum of the Hamiltonians corresponds to the maximum number of clauses that can be satisfied simultaneously.

Our second proof is based on perturbation theory—a collection of techniques that are used to analyze sums of Hamiltonians. This proof is more mathematically involved. Nevertheless, it might give more intuition as to why the 2-LOCAL HAMILTONIAN problem is QMA-complete. Unlike the first proof, which shows how to represent any QMA circuit by a 2-local Hamiltonian, the second proof shows a reduction from the 3-LOCAL HAMILTONIAN problem (which is already known to be QMA-complete [11]) to the 2-LOCAL HAMILTONIAN problem. To the best of our knowledge, this is the first reduction *inside* QMA (i.e., not from the circuit problem). This proof involves what is known as *third order* perturbation theory (interestingly, the projection lemma used in our first proof can be viewed as an instance of *first order* perturbation theory). We are not aware of any similar application of perturbation theory in the literature and we hope that our techniques will be useful elsewhere.

Adiabatic computation. It has been shown in [3] that the model of adiabatic computation with 3-local interactions is equivalent to the standard model of quantum computation (i.e., the quantum circuit model).[3] We strengthen this result by showing that 2-local interactions suffice.[4] Namely, the model of adiabatic computation with 2-local interactions is equivalent to the standard model of quantum computation. We obtain this result by applying the technique of perturbation theory, which we develop in the second proof of the main theorem.

Recent work. After a preliminary version of our paper appeared [10], Oliveira and Terhal [16] generalized our results and have shown that the 2-LOCAL HAMILTONIAN problem remains QMA-complete even if the Hamiltonians are restricted to nearest neighbor interactions between qubits on a two-dimensional grid. Similarly, they show that the model of adiabatic computation with 2-local Hamiltonians between nearest neighbor qubits on a two-dimensional grid is equivalent to standard quantum computation. Their proof applies the perturbation theory techniques that we develop in this paper and introduces several novel "perturbation gadgets" akin to our three-qubit gadget in section 6.2.

Structure. We start by describing our notation and some basics in section 2. Our first proof is developed in sections 3, 4, and 5. The main tool in this proof, which we name the projection lemma, appears in section 3. Using this lemma, we rederive in section 4 some of the previously known results. Then we give the first proof of our main theorem in section 5. In section 6 we give the second proof of our main theorem. This proof does not require the projection lemma and is in fact independent of the first proof. Hence, some readers might choose to skip sections 3, 4, and 5 and go directly to section 6. In section 7 we show how to use our techniques to prove that 2-local adiabatic computation is equivalent to standard quantum computation. Some open questions are mentioned in section 8.

2. Preliminaries. QMA is naturally defined as a class of promise problems: A promise problem L is a pair (L_{yes}, L_{no}) of disjoint sets of strings corresponding to YES and NO instances of the problem. The problem is to determine, given a string $x \in L_{yes} \cup L_{no}$, whether $x \in L_{yes}$ or $x \in L_{no}$. Let $\mathcal{B}$ be the Hilbert space of a qubit.

DEFINITION 2.1 (QMA). *Fix $\varepsilon = \varepsilon(|x|)$ such that $\varepsilon = 2^{-\Omega(|x|)}$. Then, a promise problem L is in QMA if there exists a quantum polynomial time verifier V and a polynomial p such that*

---

[3]Interestingly, their proof uses ideas from the proof of QMA-completeness of the LOCAL HAMILTONIAN problem.

[4]The main result of [3] is that 2-local adiabatic computation on *six-dimensional particles* is equivalent to standard quantum computation. This result is incomparable to ours since the particles in [3] are set on a two-dimensional grid and all 2-local interactions are between closest neighbors.

- $\forall x \in L_{yes}$ $\exists |\xi\rangle \in \mathcal{B}^{\otimes p(|x|)}$ *such that* $\Pr(V(|x\rangle, |\xi\rangle) = 1) \geq 1 - \varepsilon$,
- $\forall x \in L_{no}$ $\forall |\xi\rangle \in \mathcal{B}^{\otimes p(|x|)}$ *such that* $\Pr(V(|x\rangle, |\xi\rangle) = 1) \leq \varepsilon$,

where $\Pr(V(|x\rangle, |\xi\rangle) = 1)$ *denotes the probability that* V *outputs 1 given* $|x\rangle$ *and* $|\xi\rangle$.

We note that in the original definition ε was defined to be $2^{-\Omega(|x|)} \leq \varepsilon \leq 1/3$. By using amplification methods, it was shown in [12] that for any choice of ε in this range the resulting classes are equivalent. Hence our definition is equivalent to the original one. In a related result, Marriott and Watrous [14] showed that exponentially small ε can be achieved without amplification with a polynomial overhead in the verifier's computation.

A natural choice for the quantum analogue of SAT is the LOCAL HAMILTONIAN problem. As we will see later, this problem is indeed a complete problem for QMA.

DEFINITION 2.2. *We say that an operator* $H : \mathcal{B}^{\otimes n} \to \mathcal{B}^{\otimes n}$ *on* n *qubits is a* k-*local Hamiltonian if* H *is expressible as* $H = \sum_{j=1}^{r} H_j$ *where each term is a Hermitian operator acting on at most* k *qubits.*

DEFINITION 2.3. *The (promise) problem* k-LOCAL HAMILTONIAN *is defined as follows. We are given a* k-*local Hamiltonian on* n *qubits* $H = \sum_{j=1}^{r} H_j$ *with* $r =$ poly(n). *Each* H_j *has a bounded operator norm* $\|H_j\| \leq$ poly(n), *and its entries are specified by* poly(n) *bits. In addition, we are given two constants* a *and* b *with* $a < b$. *In* YES *instances, the smallest eigenvalue of* H *is at most* a. *In* NO *instances, it is larger than* b. *We should decide which one is the case.*

We will frequently refer to the lowest eigenvalue of some Hamiltonian H.

DEFINITION 2.4. *Let* $\lambda(H)$ *denote the lowest eigenvalue of the Hamiltonian* H. Another important notion that will be used in this paper is that of a *restriction* of a Hamiltonian.

DEFINITION 2.5. *Let* H *be a Hamiltonian and let* Π *be a projection on some subspace* $\mathcal{S}$. *Then we say that the Hamiltonian* $\Pi H \Pi$ *on* $\mathcal{S}$ *is the restriction of* H *to* $\mathcal{S}$. *We denote this restriction by* $H|_{\mathcal{S}}$.

3. Projection lemma. Our main technical tool is the *projection* lemma. This lemma (in a slightly different form) was already used in [11] and [3] but not as extensively as it is used in this paper (in fact, we apply it four times in the first proof of our main theorem). The lemma allows us to successively *cut out* parts of the Hilbert space by giving them a large *penalty*. More precisely, assume we work in some Hilbert space $\mathcal{H}$ and let H_1 be some Hamiltonian. For some subspace $\mathcal{S} \subseteq \mathcal{H}$, let H_2 be a Hamiltonian with the property that $\mathcal{S}$ is an eigenspace of eigenvalue 0 and $\mathcal{S}^{\perp}$ has eigenvalues at least J for some large $J \gg \|H_1\|$. In other words, H_2 gives a very high penalty to states in $\mathcal{S}^{\perp}$. Now consider the Hamiltonian $H = H_1 + H_2$. The projection lemma says that $\lambda(H)$, the lowest eigenvalue of H, is very close to $\lambda(H_1|_{\mathcal{S}})$, the lowest eigenvalue of the restriction of H_1 to $\mathcal{S}$. The intuitive reason for this is the following. By adding H_2 we give a very high penalty to any vector that has even a small projection in the $\mathcal{S}^{\perp}$ direction. Hence, all eigenvectors with low eigenvalue (and in particular the one corresponding to $\lambda(H)$) have to lie very close to $\mathcal{S}$. From this it follows that these eigenvectors correspond to the eigenvectors of $H_1|_{\mathcal{S}}$.

The strength of this lemma comes from the following fact. Even though H_1 and H_2 are local Hamiltonians, $H_1|_{\mathcal{S}}$ is not necessarily so. In other words, the projection lemma allows us to approximate a nonlocal Hamiltonian by a local Hamiltonian.

LEMMA 3.1. *Let* $H = H_1 + H_2$ *be the sum of two Hamiltonians operating on some Hilbert space* $\mathcal{H} = \mathcal{S} + \mathcal{S}^{\perp}$. *The Hamiltonian* H_2 *is such that* $\mathcal{S}$ *is a zero eigenspace*

and the eigenvectors in $\mathcal{S}^\perp$ have eigenvalue at least $J > 2\|H_1\|$. Then,

$$\lambda(H_1|_{\mathcal{S}}) - \frac{\|H_1\|^2}{J - 2\|H_1\|} \leq \lambda(H) \leq \lambda(H_1|_{\mathcal{S}}).$$

Notice that with, say, $J \geq 8\|H_1\|^2 + 2\|H_1\| = \text{poly}(\|H_1\|)$ we have $\lambda(H_1|_{\mathcal{S}}) - 1/8 \leq \lambda(H) \leq \lambda(H_1|_{\mathcal{S}})$.

Proof. First, we show that $\lambda(H) \leq \lambda(H_1|_{\mathcal{S}})$. Let $|\eta\rangle \in \mathcal{S}$ be the eigenvector of $H_1|_{\mathcal{S}}$ corresponding to $\lambda(H_1|_{\mathcal{S}})$. Using $H_2|\eta\rangle = 0$,

$$\langle\eta|H|\eta\rangle = \langle\eta|H_1|\eta\rangle + \langle\eta|H_2|\eta\rangle = \lambda(H_1|_{\mathcal{S}})$$

and hence H must have an eigenvector of eigenvalue at most $\lambda(H_1|_{\mathcal{S}})$.

We now show the lower bound on $\lambda(H)$. We can write any unit vector $|v\rangle \in \mathcal{H}$ as $|v\rangle = \alpha_1|v_1\rangle + \alpha_2|v_2\rangle$ where $|v_1\rangle \in \mathcal{S}$ and $|v_2\rangle \in \mathcal{S}^\perp$ are two unit vectors, $\alpha_1, \alpha_2 \in \mathbb{R}$, $\alpha_1, \alpha_2 \geq 0$ and $\alpha_1^2 + \alpha_2^2 = 1$. Let $K = \|H_1\|$. Then we have

$$
\begin{aligned}
\langle v|H|v\rangle &\geq \langle v|H_1|v\rangle + J\alpha_2^2 \\
&= (1 - \alpha_2^2)\langle v_1|H_1|v_1\rangle + 2\alpha_1\alpha_2 \text{Re}\langle v_1|H_1|v_2\rangle + \alpha_2^2\langle v_2|H_1|v_2\rangle + J\alpha_2^2 \\
&\geq \langle v_1|H_1|v_1\rangle - K\alpha_2^2 - 2K\alpha_2 - K\alpha_2^2 + J\alpha_2^2 \\
&= \langle v_1|H_1|v_1\rangle + (J - 2K)\alpha_2^2 - 2K\alpha_2 \\
&\geq \lambda(H_1|_{\mathcal{S}}) + (J - 2K)\alpha_2^2 - 2K\alpha_2,
\end{aligned}
$$

where we used $\alpha_1^2 = 1 - \alpha_2^2$ and $\alpha_1 \leq 1$. Since $(J - 2K)\alpha_2^2 - 2K\alpha_2$ is minimized for $\alpha_2 = K/(J - 2K)$, we have

$$\langle v|H|v\rangle \geq \lambda(H_1|_{\mathcal{S}}) - \frac{K^2}{J - 2K}. \qquad \square$$

4. Kitaev's construction. In this section we reprove Kitaev's result that $O(\log n)$-LOCAL HAMILTONIAN is QMA-complete. The difference between our version of the proof and the original one in [12] is that we do not use the authors' geometrical lemma to obtain the result, but rather apply our Lemma 3.1. This paves the way to the later proof that 2-LOCAL HAMILTONIAN is QMA-complete.

As mentioned before, the proof that $O(\log n)$-LOCAL HAMILTONIAN is in QMA appears in [12]. Hence, our goal is to show that any problem in QMA can be reduced to $O(\log n)$-LOCAL HAMILTONIAN. Let $V_x = V(|x\rangle, \cdot) = U_T \cdots U_1$ be a quantum verifier circuit of size $T = \text{poly}(|x|)$ operating on $N = \text{poly}(|x|)$ qubits.[5] Here and in what follows we assume without loss of generality that each U_i is either a one-qubit gate or a two-qubit gate. We further assume that $T \geq N$ and that initially, the first $m = p(|x|)$ qubits contain the proof and the remaining ancillary $N - m$ qubits are zero (see Definition 2.1). Finally, we assume that the output of the circuit is written into the first qubit (i.e., it is $|1\rangle$ if the circuit accepts). See Figure 4.1.

---

[5] For ease of notation we hardwire the dependence on the input x into the circuit.
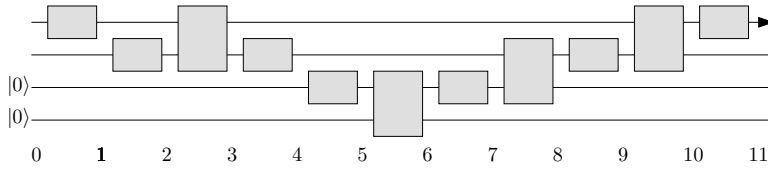
FIG. 4.1. *A circuit with $T = 11$, $N = 4$, and $m = 2$.*

The constructed Hamiltonian H operates on a space of $n = N + \log(T+1)$ qubits. The first N qubits represent the computation and the last $\log(T+1)$ qubits represent the possible values $0, \ldots, T$ for the clock:

$$H = H_{out} + J_{in}H_{in} + J_{prop}H_{prop}.$$

The coefficients J_{in} and J_{prop} will be chosen later to be some large polynomials in N. The terms are given by

$$H_{in} = \sum_{i=m+1}^{N} |1\rangle\langle 1|_i \otimes |0\rangle\langle 0|, \qquad H_{out} = (T+1)|0\rangle\langle 0|_1 \otimes |T\rangle\langle T|,$$

(4.1) $\qquad H_{prop} = \sum_{t=1}^{T} H_{prop,t},$

and

(4.2) $\qquad H_{prop,t} = \dfrac{1}{2}\left(I \otimes |t\rangle\langle t| + I \otimes |t\text{-}1\rangle\langle t\text{-}1| - U_t \otimes |t\rangle\langle t\text{-}1| - U_t^\dagger \otimes |t\text{-}1\rangle\langle t|\right)$

for $1 \le t \le T$, where $|\alpha\rangle\langle\alpha|_i$ denotes the projection on the subspace in which the ith qubit is $|\alpha\rangle$. It is understood that the first part of each tensor product acts on the space of the N computation qubits and the second part acts on the clock qubits. U_t and $U_t^\dagger$ in $H_{prop,t}$ act on the same computational qubits as U_t does when it is employed in the verifier's circuit V_x. Intuitively, each Hamiltonian "checks" a certain property by increasing the eigenvalue if the property does not hold: The Hamiltonian H_{in} checks that the input of the circuit is correct (i.e., none of the last $N - m$ computation qubits is 1), H_{out} checks that the output bit indicates acceptance, and H_{prop} checks that the propagation is according to the circuit. Notice that these Hamiltonians are $O(\log n)$-local since there are $\log(T+1) = O(\log n)$ clock qubits.

To show that a problem in QMA reduces to the $O(\log n)$-LOCAL HAMILTONIAN problem with H chosen as above, we prove the following lemma.

LEMMA 4.1. *If the circuit V_x accepts with probability more than $1 - \varepsilon$ on some input $|\xi, 0\rangle$, then the Hamiltonian H has an eigenvalue smaller than ε. If the circuit V_x accepts with probability less than ε on all inputs $|\xi, 0\rangle$, then all eigenvalues of H are larger than $\frac{3}{4} - \varepsilon$.*

Proof. Assume the circuit V_x accepts with probability more than $1 - \varepsilon$ on some $|\xi, 0\rangle$. Define

$$|\eta\rangle = \frac{1}{\sqrt{T+1}}\sum_{t=0}^{T} U_t \cdots U_1 |\xi, 0\rangle \otimes |t\rangle.$$

It can be seen that $\langle\eta|H_{prop}|\eta\rangle = \langle\eta|H_{in}|\eta\rangle = 0$ and that $\langle\eta|H_{out}|\eta\rangle < \varepsilon$. Hence, the smallest eigenvalue of H is less than ε. It remains to prove the second part of

the lemma. So now assume the circuit V_x accepts with probability less than ε on all inputs $|\xi, 0\rangle$.

Let $\mathcal{S}_{prop}$ be the groundspace of the Hamiltonian H_{prop}. It is easy to see that $\mathcal{S}_{prop}$ is a 2^N-dimensional space whose basis is given by the states

$$(4.3) \qquad |\eta_i\rangle = \frac{1}{\sqrt{T+1}} \sum_{t=0}^{T} U_t \cdots U_1 |i\rangle \otimes |t\rangle,$$

where $i \in \{0, \ldots, 2^N - 1\}$ and $|i\rangle$ represents the ith vector in the computational basis on the N computation qubits. These states have eigenvalue 0. The states in $\mathcal{S}_{prop}$ represent the correct propagation from an initial state on the N computation qubits according to the verifier's circuit V_x.

We would like to apply Lemma 3.1 with the space $\mathcal{S}_{prop}$. For that, we need to establish that $J_{prop} H_{prop}$ gives a sufficiently large $(\text{poly}(N))$ penalty to states in $\mathcal{S}_{prop}^{\perp}$. In other words, the smallest nonzero eigenvalue of H_{prop} has to be lower bounded by some inverse polynomial in N. This has been shown in [12], but we wish to briefly recall it here, as it will apply in several instances throughout this paper.

CLAIM 4.2 ([12]). *The smallest nonzero eigenvalue of H_{prop} is at least c/T^2 for some constant $c > 0$.*

Proof. We first apply the change of basis

$$W = \sum_{t=0}^{T} U_t \cdots U_1 \otimes |t\rangle\langle t|,$$

which transforms H_{prop} to

$$W^{\dagger} H_{prop} W = \sum_{t=1}^{T} I \otimes \frac{1}{2} \left(|t\rangle\langle t| + |t\text{-}1\rangle\langle t\text{-}1| - |t\rangle\langle t\text{-}1| - |t\text{-}1\rangle\langle t| \right).$$

The eigenspectrum of H_{prop} is unchanged by this transformation. The resulting Hamiltonian is block-diagonal with 2^N blocks of size $T + 1$:

$$(4.4) \qquad W^{\dagger} H_{prop} W = I \otimes \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & & \cdots & & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & \ddots & & \vdots \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & \ddots & \vdots \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ \vdots & & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ & & & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} \\ 0 & & \cdots & & 0 & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Using standard techniques, one can show that the smallest nonzero eigenvalue of each $(T + 1) \times (T + 1)$ block matrix is bounded from below by c/T^2 for some constant $c > 0$. ☐

Hence any eigenvector of $J_{prop} H_{prop}$ orthogonal to $\mathcal{S}_{prop}$ has eigenvalue at least $J = c J_{prop}/T^2$. Let us apply Lemma 3.1 with

$$H_1 = H_{out} + J_{in} H_{in}, \qquad\qquad H_2 = J_{prop} H_{prop}.$$

Note that $\|H_1\| \le \|H_{out}\| + J_{in}\|H_{in}\| \le T + 1 + J_{in}N \le \text{poly}(N)$ since H_{in} and H_{out} are sums of orthogonal projectors and $J_{in} = \text{poly}(N)$. Lemma 3.1 implies that we can choose $J_{prop} = JT^2/c = \text{poly}(N)$, such that $\lambda(H)$ is lower bounded by $\lambda(H_1|_{\mathcal{S}_{prop}}) - \frac{1}{8}$. With this in mind, let us now consider the Hamiltonian $H_1|_{\mathcal{S}_{prop}}$ on $\mathcal{S}_{prop}$.

Let $\mathcal{S}_{in} \subset \mathcal{S}_{prop}$ be the groundspace of $H_{in}|_{\mathcal{S}_{prop}}$. Then $\mathcal{S}_{in}$ is a 2^m-dimensional space whose basis is given by states as in (4.3) with $|i\rangle = |j, 0\rangle$, where $|j\rangle$ is a computational basis state on the first m computation qubits. We apply Lemma 3.1 again *inside* $\mathcal{S}_{prop}$ with

$$H_1 = H_{out}|_{\mathcal{S}_{prop}}, \qquad\qquad H_2 = J_{in}H_{in}|_{\mathcal{S}_{prop}}.$$

This time, $\|H_1\| \le \|H_{out}\| = T + 1 = \text{poly}(N)$. Any eigenvector of H_2 orthogonal to $\mathcal{S}_{in}$ inside $\mathcal{S}_{prop}$ has eigenvalue at least $J_{in}/(T + 1)$. Hence, there is a $J_{in} = \text{poly}(N)$ such that $\lambda(H_1 + H_2)$ is lower bounded by $\lambda(H_{out}|_{\mathcal{S}_{in}}) - \frac{1}{8}$.

Since the circuit V_x accepts with probability less than ε on all inputs $|\xi, 0\rangle$, we have that all eigenvalues of $H_{out}|_{\mathcal{S}_{in}}$ are larger than $1 - \varepsilon$. Hence the smallest eigenvalue of H is larger than $1 - \varepsilon - \frac{2}{8} = \frac{3}{4} - \varepsilon$, proving the second part of the lemma. $\square$

5. The 2-local construction.

Previous constructions. Let us give an informal description of ideas used in previous improvements on Kitaev's construction; these ideas will also appear in our proof. The first idea is to represent the clock register in *unary notation.* Then, the clock register consists of T qubits, and time step $t \in \{0, \ldots, T\}$ is represented by $|1^t 0^{T-t}\rangle$. The crucial observation is that clock terms that used to involve $\log(T + 1)$ qubits can now be replaced by 3-local terms that are essentially equivalent. For example, a term like $|t\text{-}1\rangle\langle t|$ can be replaced by the term $|100\rangle\langle 110|_{t-1,t,t+1}$. Since the gates U_t involve at most two qubits, we obtain a 5-local Hamiltonian. This is essentially the way 5-LOCAL HAMILTONIAN was shown to be QMA-complete in [12]. The only minor complication is that we need to get rid of illegal clock states (i.e., ones that are not a unary representation). This is done by the addition of a (2-local) Hamiltonian H_{clock} that penalizes a clock state whenever 1 appears after 0.

This result was further improved to 3-LOCAL HAMILTONIAN in [11]. The main idea there is to replace a 3-local clock term like $|100\rangle\langle 110|_{t-1,t,t+1}$ by the 1-local term $|0\rangle\langle 1|_t$. These one-qubit terms are no longer equivalent to the original clock terms. Indeed, it can be seen that they have unwanted transitions into illegal clock states. The main idea in [11] was that by giving a large penalty to illegal clock states (i.e., by multiplying H_{clock} by some large number) and applying the projection lemma, we can essentially project these one-qubit terms to the subspace of legal clock states. Inside this subspace, these terms become the required clock terms.

The 2-*local construction.* Most of the terms that appear in the construction of [11] are already 2-local. The only 3-local terms are terms as in (4.2) that correspond to two-qubit gates (those corresponding to one-qubit gates are already 2-local). Hence, in order to prove our main theorem, it is enough to find a 2-local Hamiltonian that checks for the correct propagation of two-qubit gates. This seems difficult because the Hamiltonian must somehow couple two computation qubits to a clock qubit. We circumvent this problem in the following manner. First, we isolate from the propagation Hamiltonian those terms that correspond to one-qubit gates and we multiply these terms by some large factor. Using the projection lemma, we can project the remaining Hamiltonians into a space where the one-qubit-gate propagation is correct. In other words, at this stage we can assume that our space is spanned by states that correspond to legal propagation according to the one-qubit gates. This allows us to couple clock qubits *instead* of computation qubits. To see this, consider the circuit in

Figure 5.1 at time t and at time $t + 2$. A Z gate flips the phase of a qubit if its state is $|1\rangle$ and leaves it unchanged otherwise. Hence, the phase difference between time t and time $t + 2$ corresponds to the parity of the two qubits. This phase difference can be detected by a 2-local term such as $|00\rangle\langle 11|_{t+1,t+2}$. The crucial point here is that by using a term involving only two clock qubits, we are able to check the state of two computation qubits (in this case, their parity) at a certain time. This is the main idea in our proof.

We now present the proof of the main theorem in detail. We start by making some further assumptions on the circuit V_x, all without loss of generality. First, we assume that in addition to one-qubit gates, the circuit contains only the controlled phase gate, C_ϕ. This two-qubit gate is diagonal in the computational basis and flips the sign of the state $|11\rangle$,

$$C_\phi = C_\phi{}^\dagger = |00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10| - |11\rangle\langle 11|.$$

It is known [5, 15] that quantum circuits consisting of one-qubit gates and C_ϕ gates are universal[6] and can simulate any other quantum circuit with only polynomial overhead. Second, we assume that each C_ϕ gate is both preceded and followed by two Z gates, one on each qubit, as in Figure 5.1. The Z gate is defined by $|0\rangle\langle 0| - |1\rangle\langle 1|$; i.e., it is a diagonal one-qubit gate that flips the sign of $|1\rangle$. Since both the Z gate and the C_ϕ gate are diagonal, they commute and the effect of the Z-gates cancels out. This assumption makes the circuit at most five times bigger. Finally, we assume that the C_ϕ gates are applied at regular intervals. In other words, if T_2 is the number of C_ϕ gates and L is the interval length, then a C_ϕ gate is applied at steps $L, 2L, \ldots, T_2 L$. Before the first C_ϕ gate, after the last C_ϕ gate, and between any two consecutive C_ϕ gates we have $L - 1$ one-qubit gates. This makes the total number of gates in the resulting circuit $T = (T_2 + 1)L - 1$.
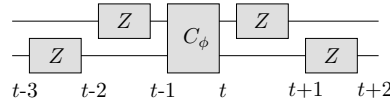
FIG. 5.1. *A modified C_ϕ gate applied at step t.*

We construct a Hamiltonian H that operates on a space of $N + T$ qubits. The first N qubits represent the computation and the last T qubits represent the clock. We think of the clock as represented in unary,

$$(5.1) \qquad |\widehat{t}\rangle \stackrel{def}{=} |\underbrace{1\ldots1}_{t}\underbrace{0\ldots0}_{T-t}\rangle.$$

Let T_1 be the time steps in which a one-qubit gate is applied. Namely, $T_1 = \{1, \ldots, T\} \setminus \{L, 2L, \ldots, T_2 L\}$. Then

$$H = H_{out} + J_{in}H_{in} + J_2 H_{prop2} + J_1 H_{prop1} + J_{clock}H_{clock},$$

---

[6]The original universal gate set in [5] consists of one-qubit gates and CNOT gates. It is, however, easy to see that a CNOT gate can be obtained from a C_ϕ gate by conjugating the second qubit with Hadamard gates (see [15]).

where

$$H_{in} = \sum_{i=m+1}^{N} |1\rangle\langle 1|_i \otimes |0\rangle\langle 0|_1, \qquad H_{out} = (T+1)|0\rangle\langle 0|_1 \otimes |1\rangle\langle 1|_T,$$

$$H_{clock} = \sum_{1 \leq i < j \leq T} I \otimes |01\rangle\langle 01|_{ij}.$$

The terms H_{prop1} and H_{prop2}, which represent the correct propagation according to the one-qubit gates and two-qubit gates, respectively, are defined as

$$H_{prop1} = \sum_{t \in T_1} H_{prop,t}, \qquad H_{prop2} = \sum_{l=1}^{T_2} (H_{qubit,lL} + H_{time,lL})$$

with

$$H_{prop,t} = \frac{1}{2}\left(I \otimes |10\rangle\langle 10|_{t,t+1} + I \otimes |10\rangle\langle 10|_{t-1,t} - U_t \otimes |1\rangle\langle 0|_t - U_t^\dagger \otimes |0\rangle\langle 1|_t\right)$$

for $t \in T_1 \cap \{2, \ldots, T-1\}$ and

$$H_{prop,1} = \frac{1}{2}\left(I \otimes |10\rangle\langle 10|_{1,2} + I \otimes |0\rangle\langle 0|_1 - U_1 \otimes |1\rangle\langle 0|_1 - U_1^\dagger \otimes |0\rangle\langle 1|_1\right),$$

$$H_{prop,T} = \frac{1}{2}\left(I \otimes |1\rangle\langle 1|_T + I \otimes |10\rangle\langle 10|_{T-1,T} - U_T \otimes |1\rangle\langle 0|_T - U_T^\dagger \otimes |0\rangle\langle 1|_T\right)$$

and, with f_t and s_t being the first and second qubits of the C_ϕ gate at time t,

$$H_{qubit,t} = \frac{1}{2}\left(-2|0\rangle\langle 0|_{f_t} - 2|0\rangle\langle 0|_{s_t} + |1\rangle\langle 1|_{f_t} + |1\rangle\langle 1|_{s_t}\right) \otimes (|1\rangle\langle 0|_t + |0\rangle\langle 1|_t),$$

$$\begin{aligned}
H_{time,t} = \frac{1}{8}I \otimes \Big(& |10\rangle\langle 10|_{t,t+1} + 6|10\rangle\langle 10|_{t+1,t+2} + |10\rangle\langle 10|_{t+2,t+3} \\
& + 2|11\rangle\langle 00|_{t+1,t+2} + 2|00\rangle\langle 11|_{t+1,t+2} \\
& + |1\rangle\langle 0|_{t+1} + |0\rangle\langle 1|_{t+1} + |1\rangle\langle 0|_{t+2} + |0\rangle\langle 1|_{t+2} \\
& + |10\rangle\langle 10|_{t-3,t-2} + 6|10\rangle\langle 10|_{t-2,t-1} + |10\rangle\langle 10|_{t-1,t} \\
& + 2|11\rangle\langle 00|_{t-2,t-1} + 2|00\rangle\langle 11|_{t-2,t-1} \\
& + |1\rangle\langle 0|_{t-2} + |0\rangle\langle 1|_{t-2} + |1\rangle\langle 0|_{t-1} + |0\rangle\langle 1|_{t-1}\Big).
\end{aligned}$$

At this point, these last two expressions might look strange. Let us say that later, when we consider their restriction to a smaller space, the reason for this definition should become clear. Note that all the above terms are at most 2-local. We will later choose $J_{in} \ll J_2 \ll J_1 \ll J_{clock} \leq \text{poly}(N)$. As in section 4, we have to prove the following lemma.

LEMMA 5.1. *Assume that the circuit V_x accepts with probability more than $1 - \varepsilon$ on some input $|\xi, 0\rangle$. Then H has an eigenvalue smaller than ε. If the circuit V_x accepts with probability less than ε on all inputs $|\xi, 0\rangle$, then all eigenvalues of H are larger than $\frac{1}{2} - \varepsilon$.*

Proof. If the circuit V_x accepts with probability more than $1 - \varepsilon$ on some input $|\xi, 0\rangle$, then the state

$$|\eta\rangle = \frac{1}{\sqrt{T+1}}\sum_{t=0}^{T} U_t \cdots U_1 |\xi, 0\rangle \otimes |\widehat{t}\rangle$$

satisfies $\langle \eta | H | \eta \rangle \leq \varepsilon$. In order to see this, one can check that

$$\langle \eta | H_{clock} | \eta \rangle = \langle \eta | H_{prop1} | \eta \rangle = \langle \eta | H_{prop2} | \eta \rangle = \langle \eta | H_{in} | \eta \rangle = 0$$

and $\langle \eta | H_{out} | \eta \rangle \leq \varepsilon$. However, verifying that $\langle \eta | H_{prop2} | \eta \rangle = 0$ can be quite tedious. Later in the proof, we will mention an easier way to see this.

In the following, we will show that if the circuit V_x accepts with probability less than ε on all inputs $|\xi, 0\rangle$, then all eigenvalues of H are larger than $\frac{1}{2} - \varepsilon$. The proof of this is based on four applications of Lemma 3.1. Schematically, we proceed as follows:

$$\mathcal{H} \supset \mathcal{S}_{legal} \supset \mathcal{S}_{prop1} \supset \mathcal{S}_{prop} \supset \mathcal{S}_{in},$$

where $\mathcal{S}_{legal}$ corresponds to states with *legal* clock states written in unary, and $\mathcal{S}_{prop1}$ is spanned by states in the legal clock space whose propagation at time steps corresponding to *one-qubit* gates (that is, in T_1) is correct. Finally, $\mathcal{S}_{prop}$ and $\mathcal{S}_{in}$ are defined in almost the same way as in section 4. These spaces will be described in more detail later.

Norms. Note that all relevant norms, as needed in Lemma 3.1, are polynomial in N. Indeed, we have $\|H_{out}\| = T + 1$ and $\|H_{in}\| \leq N$ as in section 4, $\|H_{prop1}\| \leq \sum_{t \in T_1} \|H_{prop,t}\| \leq 2T$ (each term in H_{prop1} has norm at most 2), and $\|H_{prop2}\| \leq \sum_{t=1}^{T_2}(\|H_{qubit,lL}\| + \|H_{time,lL}\|) \leq O(T_2) \leq O(T)$.

1. *Restriction to legal clock states in $\mathcal{S}_{legal}$.* Let $\mathcal{S}_{legal}$ be the $(T+1)2^N$-dimensional space spanned by states with a legal unary representation on the T clock qubits, i.e., by states of the form $|\widetilde{\xi}\rangle \otimes |\widehat{t}\rangle$ with $|\widehat{t}\rangle$ as in (5.1). In this first stage we apply Lemma 3.1 with

$$H_1 = H_{out} + J_{in}H_{in} + J_2 H_{prop2} + J_1 H_{prop1}, \qquad H_2 = J_{clock} H_{clock}.$$

Notice that $\mathcal{S}_{legal}$ is an eigenspace of H_2 of eigenvalue 0 and that states orthogonal to $\mathcal{S}_{legal}$ have eigenvalue at least J_{clock}. Lemma 3.1 implies that we can choose $J_{clock} = \text{poly}(\|H_1\|) = \text{poly}(N)$ such that $\lambda(H)$ can be lower bounded by $\lambda(H_1|_{\mathcal{S}_{legal}}) - \frac{1}{8}$. Hence, in the remainder of the proof, it is enough to study $H_1|_{\mathcal{S}_{legal}}$ inside the space $\mathcal{S}_{legal}$. This can be written as

$$H_{out}|_{\mathcal{S}_{legal}} + J_{in}H_{in}|_{\mathcal{S}_{legal}} + J_2 H_{prop2}|_{\mathcal{S}_{legal}} + J_1 H_{prop1}|_{\mathcal{S}_{legal}}$$

with

$$H_{in}|_{\mathcal{S}_{legal}} = \sum_{i=m+1}^{N} |1\rangle\langle 1|_i \otimes |\widehat{0}\rangle\langle\widehat{0}|, \qquad H_{out}|_{\mathcal{S}_{legal}} = (T+1)|0\rangle\langle 0|_1 \otimes |\widehat{T}\rangle\langle\widehat{T}|,$$

$$H_{prop,t}|_{\mathcal{S}_{legal}} = \frac{1}{2}\left(I \otimes |\widehat{t}\rangle\langle\widehat{t}| + I \otimes |\widehat{t\text{-}1}\rangle\langle\widehat{t\text{-}1}| - U_t \otimes |\widehat{t}\rangle\langle\widehat{t\text{-}1}| - U_t^{\dagger} \otimes |\widehat{t\text{-}1}\rangle\langle\widehat{t}| \right),$$

$$H_{qubit,t}|_{\mathcal{S}_{legal}} = \frac{1}{2}\left(-2|0\rangle\langle 0|_{f_t} - 2|0\rangle\langle 0|_{s_t} + |1\rangle\langle 1|_{f_t} + |1\rangle\langle 1|_{s_t} \right) \otimes \left(|\widehat{t}\rangle\langle\widehat{t\text{-}1}| + |\widehat{t\text{-}1}\rangle\langle\widehat{t}| \right),$$

$$H_{time,t}|_{\mathcal{S}_{legal}} = \frac{1}{8} I \otimes \Big(|\widehat{t}\rangle\langle\widehat{t}| + 6|\widehat{t+1}\rangle\langle\widehat{t+1}| + |\widehat{t+2}\rangle\langle\widehat{t+2}|$$

$$+ 2|\widehat{t+2}\rangle\langle\widehat{t}| + 2|\widehat{t}\rangle\langle\widehat{t+2}| + |\widehat{t+1}\rangle\langle\widehat{t}| + |\widehat{t}\rangle\langle\widehat{t+1}| + |\widehat{t+2}\rangle\langle\widehat{t+1}| + |\widehat{t+1}\rangle\langle\widehat{t+2}|$$

$$+ |\widehat{t\text{-}3}\rangle\langle\widehat{t\text{-}3}| + 6|\widehat{t\text{-}2}\rangle\langle\widehat{t\text{-}2}| + |\widehat{t\text{-}1}\rangle\langle\widehat{t\text{-}1}|$$

$$+ 2|\widehat{t\text{-}1}\rangle\langle\widehat{t\text{-}3}| + 2|\widehat{t\text{-}3}\rangle\langle\widehat{t\text{-}1}| + |\widehat{t\text{-}2}\rangle\langle\widehat{t\text{-}3}| + |\widehat{t\text{-}3}\rangle\langle\widehat{t\text{-}2}| + |\widehat{t\text{-}1}\rangle\langle\widehat{t\text{-}2}| + |\widehat{t\text{-}2}\rangle\langle\widehat{t\text{-}1}| \Big).$$

The above was obtained by noting that the projection of a term like, say, $|10\rangle\langle10|_{t,t+1}$ on $\mathcal{S}_{legal}$ is exactly $|\hat{t}\rangle\langle\hat{t}|$. Similarly, the projection of the term $|1\rangle\langle0|_{t+1}$ is $|\widehat{t+1}\rangle\langle\hat{t}|$.[7] By rearranging terms, $H_{time,t}|_{\mathcal{S}_{legal}}$ can be written as a sum of projectors:

$$\frac{1}{8}I \otimes \left\{ 2\left(|\hat{t}\rangle + |\widehat{t+1}\rangle\right)\left(\langle\hat{t}| + \langle\widehat{t+1}|\right) + 2\left(|\widehat{t+1}\rangle + |\widehat{t+2}\rangle\right)\left(\langle\widehat{t+1}| + \langle\widehat{t+2}|\right) \right.$$

$$+ \left(|\hat{t}\rangle - |\widehat{t+1}\rangle\right)\left(\langle\hat{t}| - \langle\widehat{t+1}|\right) + \left(|\widehat{t+1}\rangle - |\widehat{t+2}\rangle\right)\left(\langle\widehat{t+1}| - \langle\widehat{t+2}|\right)$$

$$- 2\left(|\hat{t}\rangle - |\widehat{t+2}\rangle\right)\left(\langle\hat{t}| - \langle\widehat{t+2}|\right)$$

$$+ 2\left(|\widehat{t\text{-}3}\rangle + |\widehat{t\text{-}2}\rangle\right)\left(\langle\widehat{t\text{-}3}| + \langle\widehat{t\text{-}2}|\right) + 2\left(|\widehat{t\text{-}2}\rangle + |\widehat{t\text{-}1}\rangle\right)\left(\langle\widehat{t\text{-}2}| + \langle\widehat{t\text{-}1}|\right)$$

$$+ \left(|\widehat{t\text{-}3}\rangle - |\widehat{t\text{-}2}\rangle\right)\left(\langle\widehat{t\text{-}3}| - \langle\widehat{t\text{-}2}|\right) + \left(|\widehat{t\text{-}2}\rangle - |\widehat{t\text{-}1}\rangle\right)\left(\langle\widehat{t\text{-}2}| - \langle\widehat{t\text{-}1}|\right)$$

$$\left. - 2\left(|\widehat{t\text{-}3}\rangle - |\widehat{t\text{-}1}\rangle\right)\left(\langle\widehat{t\text{-}3}| - \langle\widehat{t\text{-}1}|\right) \right\}.$$

(5.2)

Notice that the above expression is symmetric around $t - \frac{1}{2}$ (i.e., switching $t-1$ with t, $t-2$ with $t+1$, and $t-3$ with $t+2$ does not change the expression). Let us also mention that the fact that we have terms like $|\hat{t}\rangle - |\widehat{t+2}\rangle$ is crucial in our proof. They allow us to compare the state at time t to the state at time $t+2$.

2. *Restriction to $\mathcal{S}_{prop1}$.* We now apply Lemma 3.1 inside $\mathcal{S}_{legal}$ with

$$H_1 = (H_{out} + J_{in}H_{in} + J_2 H_{prop2})|_{\mathcal{S}_{legal}}, \qquad H_2 = J_1 H_{prop1}|_{\mathcal{S}_{legal}}.$$

Let $\mathcal{S}_{prop1}$ be the $2^N(T_2 + 1)$-dimensional space given by all states that represent correct propagation on all one-qubit gates. More precisely, let

(5.3)
$$|\eta_{l,i}\rangle \overset{def}{=} \frac{1}{\sqrt{L}} \sum_{t=lL}^{(l+1)L-1} U_t \cdots U_1 |i\rangle \otimes |\hat{t}\rangle,$$

where $l \in \{0, \ldots, T_2\}$, $i \in \{0, \ldots, 2^N - 1\}$, and $|i\rangle$ represents the ith vector in the computational basis. Then these states form a basis of $\mathcal{S}_{prop1}$. It is easy to see that each $|\eta_{l,i}\rangle$ is an eigenvector of H_{prop1} of eigenvalue 0. Hence, $\mathcal{S}_{prop1}$ is an eigenspace of eigenvalue 0 of $H_{prop1}|_{\mathcal{S}_{legal}}$. Furthermore, $H_{prop1}|_{\mathcal{S}_{legal}}$ decomposes into $T_2 + 1$ invariant blocks, with the lth block spanned by states of the form $U_t \cdots U_1 |i\rangle \otimes |\hat{t}\rangle$ for $t = lL, \ldots, (l+1)L - 1$. Inside such a block $H_{prop1}|_{\mathcal{S}_{legal}}$ corresponds exactly to H_{prop} of section 4, equations (4.1), (4.2). By Claim 4.2, its nonzero eigenvalues are at least $c/L^2 \geq c/T^2$ for some constant $c > 0$ and hence the smallest nonzero eigenvalue of $H_{prop1}|_{\mathcal{S}_{legal}}$ is also at least c/T^2. Therefore, all eigenvectors of H_2 orthogonal to $\mathcal{S}_{prop1}$ have eigenvalue at least $J = J_1 c/T^2$ and Lemma 3.1 implies that for $J_1 \geq \text{poly}(N)$, $\lambda(H_1 + H_2)$ can be lower bounded by $\lambda(H_1|_{\mathcal{S}_{prop1}}) - \frac{1}{8}$.

Hence, in the remainder of the proof, it is enough to study

$$H_{out}|_{\mathcal{S}_{prop1}} + J_{in}H_{in}|_{\mathcal{S}_{prop1}} + J_2 H_{prop2}|_{\mathcal{S}_{prop1}}.$$

Let us find $H_{prop2}|_{\mathcal{S}_{prop1}}$. Let $t = lL$ be the time at which the lth C_ϕ gate is applied and consider the projection of a state $|\eta_{l,i}\rangle$ onto the space spanned by the computation

---

[7]Notice that we do not have terms like $|1\rangle\langle1|_t$; its projection on $\mathcal{S}_{legal}$ is not $|\hat{t}\rangle\langle\hat{t}|$ but rather $|\hat{t}\rangle\langle\hat{t}| + \cdots + |\widehat{T}\rangle\langle\widehat{T}|$.

qubits and $|\hat{t}\rangle, |\widehat{t+1}\rangle, |\widehat{t+2}\rangle$. Since at time $t+1$ (resp., $t+2$) a Z gate is applied to qubit f_t (resp., s_t), this projection is a linear combination of the following four states:

$$|00\rangle_{f_t,s_t}|\xi_{00}\rangle \otimes \left(|\hat{t}\rangle + |\widehat{t+1}\rangle + |\widehat{t+2}\rangle\right),$$

$$|01\rangle_{f_t,s_t}|\xi_{01}\rangle \otimes \left(|\hat{t}\rangle + |\widehat{t+1}\rangle - |\widehat{t+2}\rangle\right),$$

$$|10\rangle_{f_t,s_t}|\xi_{10}\rangle \otimes \left(|\hat{t}\rangle - |\widehat{t+1}\rangle - |\widehat{t+2}\rangle\right),$$

$$|11\rangle_{f_t,s_t}|\xi_{11}\rangle \otimes \left(|\hat{t}\rangle - |\widehat{t+1}\rangle + |\widehat{t+2}\rangle\right),$$

where $|\xi_{b_1 b_2}\rangle$ is an arbitrary state on the remaining $N-2$ computation qubits. This implies that the restriction to $\mathcal{S}_{prop1}$ of the projector on, say, $|\hat{t}\rangle + |\widehat{t+1}\rangle$ from (5.2) is essentially the same as the restriction to $\mathcal{S}_{prop1}$ of the projector on $|0\rangle_{f_t}|\hat{t}\rangle$. More precisely, for all l_1, l_2, i_1, i_2 we have

$$\frac{1}{4}\langle\eta_{l_1,i_1}|\left(I \otimes \left(|\hat{t}\rangle + |\widehat{t+1}\rangle\right)\left(\langle\hat{t}| + \langle\widehat{t+1}|\right)\right)|\eta_{l_2,i_2}\rangle = \langle\eta_{l_1,i_1}|\left(|0\rangle\langle0|_{f_t} \otimes |\hat{t}\rangle\langle\hat{t}|\right)|\eta_{l_2,i_2}\rangle.$$

Similarly, the term involving $|\hat{t}\rangle - |\widehat{t+2}\rangle$ satisfies

$$\frac{1}{4}\langle\eta_{l_1,i_1}|\left(I \otimes \left(|\hat{t}\rangle - |\widehat{t+2}\rangle\right)\left(\langle\hat{t}| - \langle\widehat{t+2}|\right)\right)|\eta_{l_2,i_2}\rangle$$
$$= \langle\eta_{l_1,i_1}|\left(\left(|01\rangle\langle01|_{f_t,s_t} + |10\rangle\langle10|_{f_t,s_t}\right) \otimes |\hat{t}\rangle\langle\hat{t}|\right)|\eta_{l_2,i_2}\rangle.$$

Observe that the right-hand side involves two computation qubits and the clock register. Being able to obtain such a term from 2-local terms is a crucial ingredient in this proof.

Following a similar calculation, we see that from the terms involving $|\widehat{t\text{-}1}\rangle, |\widehat{t\text{-}2}\rangle, |\widehat{t\text{-}3}\rangle$ we obtain projectors involving $|\widehat{t\text{-}1}\rangle$. To summarize, instead of considering $H_{time,t}|_{\mathcal{S}_{prop1}}$ we can equivalently consider the restriction to $\mathcal{S}_{prop1}$ of

$$\frac{1}{2}\left(2|0\rangle\langle0|_{f_t} + 2|0\rangle\langle0|_{s_t} + |1\rangle\langle1|_{f_t} + |1\rangle\langle1|_{s_t} - 2\,|01\rangle\langle01|_{f_t,s_t} - 2|10\rangle\langle10|_{f_t,s_t}\right)$$
$$\otimes \left(|\widehat{t\text{-}1}\rangle\langle\widehat{t\text{-}1}| + |\hat{t}\rangle\langle\hat{t}|\right).$$

We now add the terms in $H_{qubit,t}$. A short calculation shows that the restriction to $\mathcal{S}_{prop1}$ of $(H_{time,t} + H_{qubit,t})$ is the same as the restriction to $\mathcal{S}_{prop1}$ of

$$|00\rangle\langle00|_{f_t,s_t} \otimes 2\left(|\widehat{t\text{-}1}\rangle - |\hat{t}\rangle\right)\left(\langle\widehat{t\text{-}1}| - \langle\hat{t}|\right) +$$

$$|01\rangle\langle01|_{f_t,s_t} \otimes \frac{1}{2}\left(|\widehat{t\text{-}1}\rangle - |\hat{t}\rangle\right)\left(\langle\widehat{t\text{-}1}| - \langle\hat{t}|\right) +$$

$$|10\rangle\langle10|_{f_t,s_t} \otimes \frac{1}{2}\left(|\widehat{t\text{-}1}\rangle - |\hat{t}\rangle\right)\left(\langle\widehat{t\text{-}1}| - \langle\hat{t}|\right) +$$

$$|11\rangle\langle11|_{f_t,s_t} \otimes \left(|\widehat{t\text{-}1}\rangle + |\hat{t}\rangle\right)\left(\langle\widehat{t\text{-}1}| + \langle\hat{t}|\right).$$

At this point, let us mention how one can show that for the state $|\eta\rangle$ described in the beginning of this proof, $\langle\eta|H_{prop2}|\eta\rangle = 0$. First, observe that $|\eta\rangle \in \mathcal{S}_{prop1}$ (its propagation is correct at all time steps). Next, since $|\eta\rangle$ has a C_ϕ propagation at time t, the above Hamiltonian shows that $\langle\eta|H_{prop2}|\eta\rangle = 0$.

Let us return now to the main proof. Recall that we wish to show a lower bound on the lowest eigenvalue of

$$(5.4) \qquad H_{out}|_{\mathcal{S}_{prop1}} + J_{in}H_{in}|_{\mathcal{S}_{prop1}} + J_2 H_{prop2}|_{\mathcal{S}_{prop1}}.$$

In the following, we show a lower bound on the lowest eigenvalue of the Hamiltonian

$$(5.5) \qquad H_{out}|_{\mathcal{S}_{prop1}} + J_{in}H_{in}|_{\mathcal{S}_{prop1}} + J_2 H'$$

on $\mathcal{S}_{prop1}$, where H' satisfies that $H' \leq H_{prop2}|_{\mathcal{S}_{prop1}}$, i.e., $H_{prop2}|_{\mathcal{S}_{prop1}} - H'$ is positive semidefinite. Hence, any lower bound on the lowest eigenvalue of the Hamiltonian in (5.5) implies the same lower bound on the lowest eigenvalue of the Hamiltonian in (5.4). We define H' as the sum over $t \in \{L, 2L, \ldots, T_2 L\}$ of the restriction to $\mathcal{S}_{prop1}$ of

$$|00\rangle\langle 00|_{f_t, s_t} \otimes \frac{1}{2}\left(|\widehat{t\text{-}1}\rangle - |\widehat{t}\rangle\right)\left(\langle\widehat{t\text{-}1}| - \langle\widehat{t}|\right) +$$

$$|01\rangle\langle 01|_{f_t, s_t} \otimes \frac{1}{2}\left(|\widehat{t\text{-}1}\rangle - |\widehat{t}\rangle\right)\left(\langle\widehat{t\text{-}1}| - \langle\widehat{t}|\right) +$$

$$|10\rangle\langle 10|_{f_t, s_t} \otimes \frac{1}{2}\left(|\widehat{t\text{-}1}\rangle - |\widehat{t}\rangle\right)\left(\langle\widehat{t\text{-}1}| - \langle\widehat{t}|\right) +$$

$$|11\rangle\langle 11|_{f_t, s_t} \otimes \frac{1}{2}\left(|\widehat{t\text{-}1}\rangle + |\widehat{t}\rangle\right)\left(\langle\widehat{t\text{-}1}| + \langle\widehat{t}|\right).$$

Equivalently, H' is the sum over $t \in \{L, 2L, \ldots, T_2 L\}$ of

$$\frac{1}{2}\left(I \otimes |\widehat{t}\rangle\langle\widehat{t}| + I \otimes |\widehat{t\text{-}1}\rangle\langle\widehat{t\text{-}1}| - C_\phi \otimes |\widehat{t}\rangle\langle\widehat{t\text{-}1}| - C_\phi^\dagger \otimes |\widehat{t\text{-}1}\rangle\langle\widehat{t}|\right)\Big|_{\mathcal{S}_{prop1}},$$

which resembles (4.2). Note that this term enforces correct propagation at time step $t = lL$. We claim that

$$(5.6) \qquad H' = \frac{1}{2L} \sum_{i=0}^{2^N-1} \sum_{l=1}^{T_2} \left(|\eta_{l-1,i}\rangle - |\eta_{l,i}\rangle\right)\left(\langle\eta_{l-1,i}| - \langle\eta_{l,i}|\right).$$

The intuitive reason for this is the following. For any i, $|\eta_{l-1,i}\rangle + |\eta_{l,i}\rangle$ can be seen as a correct propagation at time $t = lL$. In other words, consider the projection of $|\eta_{l,i}\rangle$ on clock $|\widehat{t}\rangle$ and the projection of $|\eta_{l-1,i}\rangle$ on clock $|\widehat{t\text{-}1}\rangle$. Then the first state is exactly the second state after applying the lth C_ϕ gate. This means that inside $\mathcal{S}_{prop1}$, checking correct propagation from time $t-1$ to time t is equivalent to checking correct propagation from $|\eta_{l-1,i}\rangle$ to $|\eta_{l,i}\rangle$.

More precisely, fix some l and $t = lL$. Then, using (5.3), we get that for all l_1, l_2, i_1, i_2 such that either $l_1 \neq l$, $l_2 \neq l$, or $i_1 \neq i_2$,

$$\langle\eta_{l_1,i_1}|\left(I \otimes |\widehat{t}\rangle\langle\widehat{t}|\right)|\eta_{l_2,i_2}\rangle = 0.$$

Otherwise, $l_1 = l_2 = l$ and $i_1 = i_2 = i$ for some i and we have

$$\langle\eta_{l,i}|\left(I \otimes |\widehat{t}\rangle\langle\widehat{t}|\right)|\eta_{l,i}\rangle = \frac{1}{L}.$$

Hence we obtain

$$I \otimes |\widehat{t}\rangle\langle\widehat{t}||_{\mathcal{S}_{prop1}} = \frac{1}{L} \sum_{i=0}^{2^N-1} |\eta_{l,i}\rangle\langle\eta_{l,i}|$$

and similarly,

$$I \otimes |\widehat{t\text{-}1}\rangle\langle\widehat{t\text{-}1}|\,\|_{\mathcal{S}_{prop1}} = \frac{1}{L} \sum_{i=0}^{2^N-1} |\eta_{l-1,i}\rangle\langle\eta_{l-1,i}|.$$

For the off-diagonal terms we see that

$$\langle\eta_{l_1,i_1}| \left(C_\phi \otimes |\widehat{t}\rangle\langle\widehat{t\text{-}1}| \right) |\eta_{l_2,i_2}\rangle = 0$$

if $l_1 \neq l$ or $l_2 \neq l-1$. If $l_1 = l$ and $l_2 = l-1$, then using $C_\phi = U_{lL}$, we get

$$\langle\eta_{l,i_1}| \left(C_\phi \otimes |\widehat{t}\rangle\langle\widehat{t\text{-}1}| \right) |\eta_{l-1,i_2}\rangle = \frac{1}{L}\langle i_1| \left(U_{lL}\cdots U_1\right)^\dagger C_\phi U_{lL-1}\cdots U_1|i_2\rangle = \frac{1}{L}\langle i_1|i_2\rangle,$$

which is 0 if $i_1 \neq i_2$ and $\frac{1}{L}$ otherwise. Hence $C_\phi \otimes |\widehat{t}\rangle\langle\widehat{t\text{-}1}|\,\|_{\mathcal{S}_{prop1}} = \frac{1}{L}\sum_{i=0}^{2^N-1} |\eta_{l,i}\rangle\langle\eta_{l-1,i}|$ and similarly for its Hermitian adjoint. This establishes (5.6).

3. *Restriction to* $\mathcal{S}_{prop}$. Let $\mathcal{S}_{prop}$ be the 2^N-dimensional space whose basis is given by the states

$$|\eta_i\rangle = \frac{1}{\sqrt{T+1}} \sum_{t=0}^{T} U_t\cdots U_1|i\rangle \otimes |\widehat{t}\rangle = \frac{1}{\sqrt{T_2+1}} \sum_{l=0}^{T_2} |\eta_{l,i}\rangle$$

for $i \in \{0,\ldots,2^N-1\}$. Equation (5.6) shows that $\mathcal{S}_{prop}$ is an eigenspace of H' of eigenvalue 0. Moreover, H' is block-diagonal with 2^N blocks of size T_2+1. Each block is a matrix as in (4.4), multiplied by $1/L$. As in Claim 4.2 we see that the smallest nonzero eigenvalue of this Hamiltonian is $c/LT_2^2 \geq c/T^2$ for some constant c. Now we can apply Lemma 3.1. This time, we apply it inside $\mathcal{S}_{prop1}$ with

$$H_1 = \left(H_{out} + J_{in}H_{in}\right)|_{\mathcal{S}_{prop1}}, \qquad\qquad H_2 = J_2 H'.$$

Eigenvectors of H_2 orthogonal to $\mathcal{S}_{prop}$ have eigenvalue at least $J = J_2 c/T^2$. As before, we can choose $J_2 = \text{poly}(N)$ such that $\lambda(H_1 + H_2)$ is lower bounded by $\lambda(H_1|_{\mathcal{S}_{prop}}) - \frac{1}{8}$. Hence, in the remainder we consider

$$H_{out}|_{\mathcal{S}_{prop}} + J_{in}H_{in}|_{\mathcal{S}_{prop}}.$$

4. *Restriction to* $\mathcal{S}_{in}$. The rest of the proof proceeds in the same way as in section 4. Indeed, the subspace $\mathcal{S}_{prop}$ is isomorphic to the one in section 4 and both $H_{out}|_{\mathcal{S}_{prop}}$ and $H_{in}|_{\mathcal{S}_{prop}}$ are the same Hamiltonians. So by another application of Lemma 3.1 we get that the lowest eigenvalue of $H_{out}|_{\mathcal{S}_{prop}} + J_{in}H_{in}|_{\mathcal{S}_{prop}}$ is lower bounded by $\lambda(H_{out}|_{\mathcal{S}_{in}}) - \frac{1}{8}$. As in section 4, we have that $\lambda(H_{out}|_{\mathcal{S}_{in}}) > 1 - \varepsilon$ if the circuit accepts with probability less than ε. Hence $\lambda(H)$, the lowest eigenvalue of the original Hamiltonian H, is larger than $1 - \varepsilon - \frac{4}{8} = \frac{1}{2} - \varepsilon$. $\quad\square$

6. Perturbation theory proof. In this section we give an alternative proof of our main theorem. In section 6.1, we develop our perturbation theory technique. Since this technique might constitute a useful tool in other Hamiltonian constructions, we keep the presentation as general as possible. Then, in section 6.2, we present a specific application of our technique, the three-qubit gadget. Finally, in section 6.3, we use this gadget to complete the proof of the main theorem.

6.1. Perturbation theory. The goal in perturbation theory is to analyze the spectrum of the sum of two Hamiltonians $\widetilde{H} = H + V$ in the case that V has a small norm compared to the spectral gap of H. One setting was described in the projection lemma. Specifically, assume H has a zero eigenvalue with the associated eigenspace $\mathcal{S}$, whereas all other eigenvalues are greater than $\Delta \gg \|V\|$. The projection lemma shows that in this case, the lowest eigenvalue of $\widetilde{H}$ is close to that of $V|_{\mathcal{S}}$. In this section we find a better approximation to $\operatorname{Spec}\widetilde{H}$ by considering certain correction terms that involve higher powers of V. It turns out that these higher order correction terms include interesting interactions, which will allow us to create an effective 3-local Hamiltonian from 2-local terms. We remark that the projection lemma (for the entire lower part of the spectrum) can be obtained by following the development done in this section up to the first order.

Before giving a more detailed description of the technique, we need to introduce a certain amount of notation. For two Hermitian operators H and V, let $\widetilde{H} = H + V$. We refer to H as the *unperturbed Hamiltonian* and to V as the *perturbation Hamiltonian.* Let λ_j, $|\psi_j\rangle$ be the eigenvalues and eigenvectors of H, whereas the eigenvalues and eigenvectors of $\widetilde{H}$ are denoted by $\widetilde{\lambda}_j$, $|\widetilde{\psi}_j\rangle$. In case of multiplicities, some eigenvalues might appear more than once. We order the eigenvalues in a nondecreasing order

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_{\dim \mathcal{H}}, \qquad \widetilde{\lambda}_1 \le \widetilde{\lambda}_2 \le \cdots \le \widetilde{\lambda}_{\dim \mathcal{H}}.$$

In general, everything related to the perturbed Hamiltonian is marked with a tilde.

An important component in our proof is the *resolvent* of $\widetilde{H}$, defined as

$$(6.1) \qquad \widetilde{G}(z) = \left(zI - \widetilde{H}\right)^{-1} = \sum_j \left(z - \widetilde{\lambda}_j\right)^{-1} |\widetilde{\psi}_j\rangle\langle\widetilde{\psi}_j|.$$

It is a meromorphic[8] operator-valued function of the complex variable z with poles at $z = \widetilde{\lambda}_j$. In fact, for our purposes, it is sufficient to consider real z.[9] Its usefulness comes from the fact that poles can be preserved under projections (while eigenvalues are usually lost). Similarly, we define the resolvent of H as $G(z) = (zI - H)^{-1}$.[10]

Let $\lambda_* \in \mathbb{R}$ be some cutoff on the spectrum of H.

DEFINITION 6.1. *Let $\mathcal{H} = \mathcal{L}_+ \oplus \mathcal{L}_-$, where $\mathcal{L}_+$ is the space spanned by eigenvectors of H with eigenvalues $\lambda \ge \lambda_*$ and $\mathcal{L}_-$ is spanned by eigenvectors of H of eigenvalue $\lambda < \lambda_*$. Let $\Pi_\pm$ be the corresponding projection onto $\mathcal{L}_\pm$. For an operator X on $\mathcal{H}$ define the operator $X_{++} = X|_{\mathcal{L}_+} = \Pi_+ X \Pi_+$ on $\mathcal{L}_+$ and similarly $X_{--} = X|_{\mathcal{L}_-}$. We also define $X_{+-} = \Pi_+ X \Pi_-$ as an operator from $\mathcal{L}_-$ to $\mathcal{L}_+$, and similarly X_{-+}.*

With these definitions, in a representation of $\mathcal{H} = \mathcal{L}_+ \oplus \mathcal{L}_-$ both H and G are block-diagonal and we will omit one index for their blocks, i.e., $H_+ \overset{def}{=} H_{++}$,

---

[8] A meromorphic function is analytic in all but a discrete subset of $\mathbb{C}$, and these singularities must be poles and not essential singularities.

[9] The resolvent is the main tool in abstract spectral theory [19]. In physics, it is known as the *Green's function.* Physicists actually use slightly different Green's functions that are suited for specific problems.

[10] We can express $\widetilde{G}$ in terms of G (where we omit the variable z): $\widetilde{G} = (G^{-1} - V)^{-1} = G(I - VG)^{-1} = G + GVG + GVGVG + GVGVGVG + \cdots$. This expansion of $\widetilde{G}$ in powers of V may be represented by Feynman diagrams [1].

$G_+ \overset{def}{=} G_{++}$ and so on. Note that $G_\pm^{-1} = zI_\pm - H_\pm$. To summarize, we have

$$\widetilde{H} = \begin{pmatrix} \widetilde{H}_{++} & \widetilde{H}_{+-} \\ \widetilde{H}_{-+} & \widetilde{H}_{--} \end{pmatrix}, \quad V = \begin{pmatrix} V_{++} & V_{+-} \\ V_{-+} & V_{--} \end{pmatrix}, \quad H = \begin{pmatrix} H_+ & 0 \\ 0 & H_- \end{pmatrix},$$

$$\widetilde{G} = \begin{pmatrix} \widetilde{G}_{++} & \widetilde{G}_{+-} \\ \widetilde{G}_{-+} & \widetilde{G}_{--} \end{pmatrix}, \quad G = \begin{pmatrix} G_+ & 0 \\ 0 & G_- \end{pmatrix}.$$

We similarly write $\mathcal{H} = \widetilde{\mathcal{L}}_+ \oplus \widetilde{\mathcal{L}}_-$ according to the spectrum of $\widetilde{H}$ and the cutoff λ_*. Finally, we define

$$\Sigma_-(z) = zI_- - \widetilde{G}_{--}^{-1}(z).$$

This operator-valued function is called *self-energy*.[11]

　　With these notations in place, we can now give an overview of what follows. Our goal is to approximate the spectrum of $\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$. We will do this by showing that in some sense, the spectrum of $\Sigma_-(z)$ gives such an approximation. To see why this arises, notice that by definition of $\Sigma_-(z)$, we have $\widetilde{G}_{--}(z) = \left(zI_- - \Sigma_-(z)\right)^{-1}$. In some sense, this equation is the analogue of (6.1), where $\Sigma_-(z)$ plays the role of a Hamiltonian for the projected resolvent $\widetilde{G}_{--}(z)$. However, $\Sigma_-(z)$ is in general z-dependent and not a fixed Hamiltonian. Nonetheless, for certain choices of H and V, $\Sigma_-(z)$ is nearly constant in a certain range of z so we can choose an *effective Hamiltonian* H_{eff} that approximates $\Sigma_-(z)$ in this range. Our main theorem relates the spectrum of H_{eff} to that of $\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$.

　　THEOREM 6.2. *Assume H has a spectral gap Δ around the cutoff λ_*; i.e., all its eigenvalues are in $(-\infty, \lambda_-] \cup [\lambda_+, +\infty)$, where $\lambda_+ = \lambda_* + \Delta/2$ and $\lambda_- = \lambda_* - \Delta/2$. Assume, moreover, that $\|V\| < \Delta/2$. Let $\varepsilon > 0$ be arbitrary. Assume there exists an operator H_{eff} such that $\mathrm{Spec}\, H_{\mathrm{eff}} \subseteq [c,d]$ for some $c < d < \lambda_* - \varepsilon$ and, moreover, the inequality*

$$\|\Sigma_-(z) - H_{\mathrm{eff}}\| \le \varepsilon$$

holds for all $z \in [c - \varepsilon, d + \varepsilon]$. Then each eigenvalue $\widetilde{\lambda}_j$ of $\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$ is ε-close to the jth eigenvalue of H_{eff}.

　　The usefulness of the theorem comes from the fact that $\Sigma_-(z)$ has a natural series expansion, which can be truncated to obtain H_{eff}. This series may give rise to interesting terms; for example, in our application, 2-local terms in H and V lead to 3-local terms in H_{eff}. To obtain this expansion, we start by expressing $\widetilde{G}$ in terms of G as

$$\widetilde{G} = \left(G^{-1} - V\right)^{-1} = \begin{pmatrix} G_+^{-1} - V_{++} & -V_{+-} \\ -V_{-+} & G_-^{-1} - V_{--} \end{pmatrix}^{-1}.$$

Then, using the block matrix identity

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} \left(A - BD^{-1}C\right)^{-1} & -A^{-1}B\left(D - CA^{-1}B\right)^{-1} \\ -D^{-1}C\left(A - BD^{-1}C\right)^{-1} & \left(D - CA^{-1}B\right)^{-1} \end{pmatrix}$$

---

[11] As we will see later, this definition includes an H_- term. This term is usually not considered part of self-energy, but we have included it for notational convenience.

we conclude that

$$\widetilde{G}_{--} = \left(G_-^{-1} - V_{--} - V_{-+}\left(G_+^{-1} - V_{++}\right)^{-1}V_{+-}\right)^{-1}.$$

Finally, we can represent $\Sigma_-(z)$ using the series expansion $(I-X)^{-1} = I+X+X^2+\cdots$,

$$
\begin{aligned}
\Sigma_-(z) &= H_- + V_{--} + V_{-+}\left(G_+^{-1} - V_{++}\right)^{-1}V_{+-} \\
(6.2) \qquad &= H_- + V_{--} + V_{-+}G_+\left(I_+ - V_{++}G_+\right)^{-1}V_{+-} \\
&= H_- + V_{--} + V_{-+}G_+V_{+-} + V_{-+}G_+V_{++}G_+V_{+-} \\
&\qquad\qquad + V_{-+}G_+V_{++}G_+V_{++}G_+V_{+-} + \cdots .
\end{aligned}
$$

Proof of Theorem 6.2. We start with an overview of the proof. We first notice that, by definition, the eigenvalues of $\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$ appear as poles in $\widetilde{G}$. In Lemma 6.4, we show that these poles also appear as poles of $\widetilde{G}_{--}$. As mentioned before, this is the reason we work with resolvents. In Lemmas 6.5 and 6.6 we relate these poles to the eigenvalues of Σ_- by showing that z is a pole of $\widetilde{G}_{--}$ if and only if it is an eigenvalue of $\Sigma_-(z)$. In other words, these are values of z for which $\Sigma_-(z)$ has z as an eigenvalue. Finally, we complete the proof of the theorem by using the assumption that $\Sigma_-(z)$ is close to H_{eff}, and thus any eigenvalue of $\Sigma_-(z)$ must be close to an eigenvalue of H_{eff}. This situation is illustrated in Figure 6.1.
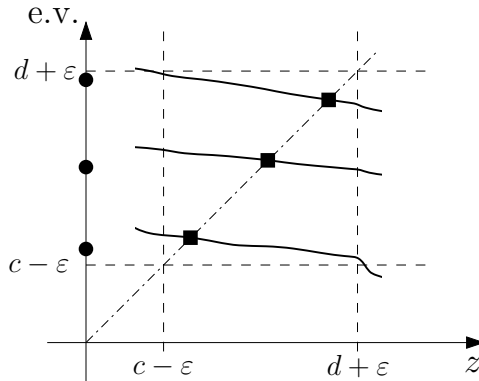

FIG. 6.1. *The spectrum of $\Sigma_-(z)$ as a function of z is indicated with solid curves. The boxes correspond to the spectrum of $\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$; they are those eigenvalues of $\Sigma_-(z)$ that lie on the dashed line $z = e.v.$ The dots indicate the spectrum of H_{eff}, which approximates the spectrum of $\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$.*

We start with a simple lemma that says that if two Hamiltonians H_1, H_2 are close, their spectra must also be close. It is a special case of Weyl's inequalities (see, e.g., section III.2 in [6]).

LEMMA 6.3. *Let H_1, H_2 be two Hamiltonians with eigenvalues $\mu_1 \le \mu_2 \le \cdots$ and $\sigma_1 \le \sigma_2 \le \cdots$. Then, for all j, $|\mu_j - \sigma_j| \le \|H_1 - H_2\|$.*

Proof. We will use a fact from the theory of Hermitian forms: If $X \le Y$ (i.e., if $Y - X$ is positive semidefinite), then the operator Y has at least as many positive and nonnegative eigenvalues as X. Let $\varepsilon = \|H_1 - H_2\|$; then

$$(\mu_j - \varepsilon)I - H_2 \le \mu_j I - H_1 \le (\mu_j + \varepsilon)I - H_2.$$

The operator $\mu_j I - H_1$ has at most $j-1$ positive and at least j nonnegative eigenvalues. Hence $(\mu_j - \varepsilon)I - H_2$ has at most $j - 1$ positive eigenvalues, and $(\mu_j + \varepsilon)I - H_2$ has at least j nonnegative eigenvalues. It follows that $\sigma_j \in [\mu_j - \varepsilon, \ \mu_j + \varepsilon]$.   □

The next lemma asserts that the poles of $\widetilde{G}_{--}$ in the range $(-\infty, \lambda_*)$ are in one-to-one correspondence with the eigenvalues of $\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$. Hence we can recover the eigenvalues of $\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$ from the poles of $\widetilde{G}_{--}$.

LEMMA 6.4. *Let* $\tilde{\lambda}$ *be in* $(-\infty, \lambda_*)$ *and let* $m \geq 0$ *be its multiplicity as an eigenvalue of* $\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$. *Then around* $\tilde{\lambda}$, $\widetilde{G}_{--}$ *is of the form* $(z - \tilde{\lambda})^{-1}A + O(1)$, *where* A *is a rank* m *operator.*

Proof. We first show that $\widetilde{\mathcal{L}}_- \cap \mathcal{L}_+ = \{0\}$. Suppose the contrary; i.e., there is a nonzero vector $|\xi\rangle \in \widetilde{\mathcal{L}}_- \cap \mathcal{L}_+$. Without loss of generality, $\langle\xi|\xi\rangle = 1$. Then we have $\langle\xi|(H + V)|\xi\rangle \leq \lambda_*$ (since $|\xi\rangle \in \widetilde{\mathcal{L}}_-$) and $\langle\xi|H|\xi\rangle \geq \lambda_+$ (since $|\xi\rangle \in \mathcal{L}_+$). Hence $\langle\xi|V|\xi\rangle \leq \lambda_* - \lambda_+ = -\Delta/2$. But this is impossible because $\|V\| < \Delta/2$.

Now, since $\widetilde{\mathcal{L}}_- \cap \mathcal{L}_+ = \{0\}$, we have that $\Pi_-|\xi\rangle \neq 0$ for all nonzero vectors $|\xi\rangle \in \widetilde{\mathcal{L}}_-$. From (6.1) we obtain

$$\widetilde{G}_{--} = \Pi_-\widetilde{G}\Pi_- = \sum_j (z - \tilde{\lambda}_j)^{-1}\Pi_-|\widetilde{\psi}_j\rangle\langle\widetilde{\psi}_j|\Pi_-.$$

If the multiplicity of $\tilde{\lambda}$ is m, then the matrix $\sum|\widetilde{\psi}_j\rangle\langle\widetilde{\psi}_j|$ of the corresponding eigenvectors has rank m. This implies that the matrix $\sum\Pi_-|\widetilde{\psi}_j\rangle\langle\widetilde{\psi}_j|\Pi_-$ also has rank m. Indeed, if there is some linear combination of $\Pi_-|\widetilde{\psi}_j\rangle$ that sums to zero, then taking the same linear combination of $|\widetilde{\psi}_j\rangle$ must also sum to zero.   □

The next two lemmas relate the spectrum of $\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$ to the operator $\Sigma_-(z)$.

LEMMA 6.5. *For any* $z < \lambda_*$, *the multiplicity of* z *as an eigenvalue of* $\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$ *is equal to the multiplicity of* z *as an eigenvalue of* $\Sigma_-(z)$.

Proof. Fix some $z < \lambda_*$ and let m be its multiplicity as an eigenvalue of $\widetilde{H}$ (in particular, $m = 0$ if z is not an eigenvalue of $\widetilde{H}$). In the neighborhood of z the function $\widetilde{G}_{--}(w)$ has the form

$$\widetilde{G}_{--}(w) = (w - z)^{-1}A + B + O\big(|w - z|\big),$$

where by Lemma 6.4, A is an operator of rank m. We now consider $\widetilde{G}_{--}^{-1}(w)$. For any $w < \lambda_+ - \|V\|$ the norm of $G_+(w)$ is strictly less than $1/\|V\|$. Hence, by (6.2) we see that all the poles of $\Sigma_-(w)$ lie on the interval $[\lambda_+ - \|V\|, +\infty)$; in particular $\widetilde{G}_{--}^{-1}(w) = wI_- - \Sigma_-(w)$ is analytic for $w \in (-\infty, \lambda_*]$. Hence we can write

$$\widetilde{G}_{--}^{-1}(w) = wI_- - \Sigma_-(w) = C + D(w - z) + O\big(|w - z|^2\big).$$

We claim that the dimension of the null-space of C is exactly m. Notice that this implies that z is an m-fold eigenvalue of $\Sigma_-(z) = zI_- - C$. By multiplying the two equations above, we obtain

$$I_- = \widetilde{G}_{--}^{-1}(w)\widetilde{G}_{--}(w) = (w - z)^{-1}CA + (DA + CB) + O(|w - z|).$$

By equating coefficients, we obtain $CA = 0$ and $DA + CB = I_-$. On one hand, $CA = 0$ implies that the null-space of C has dimension at least m. On the other hand, the rank of DA is at most rank$(A) = m$. Since I_- has full rank, the dimension

of the null-space of CB must be at most m. This implies that the dimension of the null-space of C must also be at most m.   □

We observe that the function $\Sigma_-(z)$ is monotone decreasing in the operator sense (i.e., if $z_1 \leq z_2$, then $\Sigma_-(z_1) - \Sigma_-(z_2)$ is positive semidefinite):

$$\frac{d\Sigma_-(z)}{dz} = \frac{d}{dz}\Big(H_- + V_{--} + V_{-+}(zI_+ - H_+ - V_{++})^{-1}V_{+-}\Big)$$
$$= -V_{-+}(zI_+ - H_+ - V_{++})^{-2}V_{+-} \leq 0.$$

LEMMA 6.6. *Let $\widetilde{\lambda}_j$ be the jth eigenvalue of $\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$. Then it is also the jth eigenvalue of $\Sigma_-(\widetilde{\lambda}_j)$.*

Proof. For any $z \in \mathbb{R}$, let $f_1(z)$ (resp., $f_2(z)$) be the number of eigenvalues not greater than z of $\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$ (resp., $\Sigma_-(z)$). When $z \to -\infty$, $f_1(z)$ is clearly 0. By the monotonicity of Σ_- we see that $f_2(z)$ is also 0. Using Lemma 6.5 we see that as z increases, both numbers increase together by the same amount m whenever z hits an eigenvalue of $\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$ of multiplicity m (here we used again the monotonicity of Σ_-). Hence, for all z, $f_1(z) = f_2(z)$ and the lemma is proven.   □

We can now complete the proof of the theorem. By Lemma 6.3 and our assumption on H_{eff}, we have that for any $z \in [c - \varepsilon, d + \varepsilon]$, $\operatorname{Spec}\Sigma_-(z)$ is contained in $[c - \varepsilon, d + \varepsilon]$. From this and the monotonicity of Σ_-, we obtain that there is no $z \in (d + \varepsilon, \lambda_*]$ that is an eigenvalue of $\Sigma_-(z)$. Similarly, there is no $z < c - \varepsilon$ that is an eigenvalue of $\Sigma_-(z)$. Hence, using Lemma 6.5 we see that $\operatorname{Spec}\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$ is contained in $[c - \varepsilon, d + \varepsilon]$. Now let $\widetilde{\lambda}_j \in [c - \varepsilon, d + \varepsilon]$ be the jth eigenvalue of $\widetilde{H}|_{\widetilde{\mathcal{L}}_-}$. By Lemma 6.6 it is also the jth eigenvalue of $\Sigma_-(\widetilde{\lambda}_j)$. By Lemma 6.3 it is ε-close to the jth eigenvalue of H_{eff}.   □

6.2. The three-qubit gadget. In this section we demonstrate how Theorem 6.2 can be used to transform a 3-local Hamiltonian into a 2-local one. The complete reduction will be shown in the next section. From now we try to keep the discussion more specialized to our QMA problem rather than presenting it in full generality as was done in section 6.1.

Let Y be some arbitrary 2-local Hamiltonian acting on a space $\mathcal{M}$ of N qubits. Also, let B_1, B_2, B_3 be positive semidefinite Hamiltonians each acting on a different qubit (so they commute). We think of these four operators as having constant norm. Assume we have the 3-local Hamiltonian

$$(6.3) \qquad\qquad\qquad Y - 6B_1B_2B_3.$$

The factor 6 is added for convenience. Recall that in the LOCAL HAMILTONIAN problem we are interested in the lowest eigenvalue of a Hamiltonian. Hence, our goal is to find a 2-local Hamiltonian whose lowest eigenvalue is very close to the lowest eigenvalue of (6.3).

We start by adding three qubits to our system. For $j = 1, 2, 3$, we denote the Pauli operators acting on the jth qubit by σ_j^α. Let $\delta > 0$ be a sufficiently small constant. Our 2-local Hamiltonian is $\widetilde{H} = H + V$, where

$$H = -\frac{\delta^{-3}}{4}I \otimes \big(\sigma_1^z\sigma_2^z + \sigma_1^z\sigma_3^z + \sigma_2^z\sigma_3^z - 3I\big),$$
$$V = X \otimes I - \delta^{-2}\big(B_1 \otimes \sigma_1^x + B_2 \otimes \sigma_2^x + B_3 \otimes \sigma_3^x\big),$$
$$X = Y + \delta^{-1}(B_1^2 + B_2^2 + B_3^2).$$

The unperturbed Hamiltonian H has eigenvalues 0 and $\Delta \stackrel{def}{=} \delta^{-3}$. Associated with the zero eigenvalue is the subspace

$$\mathcal{L}_- = \mathcal{M} \otimes \mathcal{C}, \qquad \text{where} \quad \mathcal{C} = \big(|000\rangle, |111\rangle\big).$$

In the orthogonal subspace $\mathcal{C}^\perp$ we have the states $|001\rangle, |010\rangle$, etc. We may think of the subspace $\mathcal{C}$ as an effective qubit (as opposed to the three physical qubits); the corresponding Pauli operators are denoted by $\sigma_{\text{eff}}^\alpha$.

To obtain H_{eff}, we now compute the self-energy $\Sigma_-(z)$ using the power expansion in (6.2) up to the third order. There is no zeroth order term, i.e., $H_- = 0$. For the remaining terms, notice that $G_+ = (z - \Delta)^{-1} I_{\mathcal{L}_+}$. Hence, we can write $\Sigma_-(z)$ as

$$V_{--} + (z-\Delta)^{-1} V_{-+} V_{+-} + (z-\Delta)^{-2} V_{-+} V_{++} V_{+-} + (z-\Delta)^{-3} V_{-+} V_{++} V_{++} V_{+-} + \cdots.$$

The first term is $V_{--} = X \otimes I_{\mathcal{C}}$ because a σ^x term takes any state in $\mathcal{C}$ to $\mathcal{C}^\perp$. The expressions in the following terms are of the form

$$V_{-+} = -\delta^{-2}\big(B_1 \otimes |000\rangle\langle 100| + B_2 \otimes |000\rangle\langle 010| + B_3 \otimes |000\rangle\langle 001| +$$
$$B_1 \otimes |111\rangle\langle 011| + B_2 \otimes |111\rangle\langle 101| + B_3 \otimes |111\rangle\langle 110|\big),$$

$$V_{++} = X \otimes I_{\mathcal{C}^\perp} - \delta^{-2}\big(B_1 \otimes (|001\rangle\langle 101| + |010\rangle\langle 110| + |101\rangle\langle 001| + |110\rangle\langle 010|) +$$
$$B_2 \otimes (\dots) + B_3 \otimes (\dots)\big),$$

where the dots denote similar terms for B_2 and B_3. Now, in the second term of $\Sigma_-(z)$, V_{+-} flips one of the physical qubits, and V_{-+} must return it to its original state in order to return to the space $\mathcal{C}$. Hence we have $V_{-+} V_{+-} = \delta^{-4}(B_1^2 + B_2^2 + B_3^2) \otimes I_{\mathcal{C}}$. The third term is slightly more involved. Here we have two possible processes. In the first process, V_{+-} flips a qubit, V_{++} acts with $X \otimes I_{\mathcal{C}^\perp}$, and finally V_{-+} flips the qubit back. In the second process, V_{+-}, V_{++}, and V_{-+} flip all three qubits in succession. Thus,

(6.4)
$$\Sigma_-(z) = X \otimes I_{\mathcal{C}} + (z-\Delta)^{-1} \delta^{-4}(B_1^2 + B_2^2 + B_3^2) \otimes I_{\mathcal{C}}$$

$$+ (z-\Delta)^{-2} \delta^{-4}(B_1 X B_1 + B_2 X B_2 + B_3 X B_3) \otimes I_{\mathcal{C}}$$

$$- (z-\Delta)^{-2} \delta^{-6}$$
$$\big(B_3 B_2 B_1 + B_2 B_3 B_1 + B_3 B_1 B_2 + B_1 B_3 B_2 + B_2 B_1 B_3 + B_1 B_2 B_3\big) \otimes \sigma_{\text{eff}}^x$$

$$+ O\big(\|V\|^4 (z-\Delta)^{-3}\big).$$

We now focus on the range $z = O(1) \ll \Delta$. In this range we have

$$(z-\Delta)^{-1} = -\frac{1}{\Delta}\Big(1 - \frac{z}{\Delta}\Big)^{-1} = -\frac{1}{\Delta} + O\Big(\frac{z}{\Delta^2}\Big) = -\delta^3 + O(\delta^6).$$

Simplifying, we obtain

$$\Sigma_-(z) = \underbrace{Y \otimes I_{\mathcal{C}} - 6 B_1 B_2 B_3 \otimes \sigma_{\text{eff}}^x}_{H_{\text{eff}}} + O(\delta).$$

Notice that $\|H_{\text{eff}}\| = O(1)$ and hence we obtain that for all z in, say, $[-2\|H_{\text{eff}}\|, 2\|H_{\text{eff}}\|]$ we have

$$\|\Sigma_-(z) - H_{\text{eff}}\| = O(\delta).$$

We may now apply Theorem 6.2 with $c = -\|H_{\text{eff}}\|$, $d = \|H_{\text{eff}}\|$, and $\lambda_* = \Delta/2$ to obtain the following result: Each eigenvalue $\widetilde{\lambda}_j$ from the lower part of Spec $\widetilde{H}$ is $O(\delta)$-close to the j-th eigenvalue of H_{eff}. In fact, for our purposes, it is enough that the lowest eigenvalue of $\widetilde{H}$ is $O(\delta)$-close to the lowest eigenvalue of H_{eff}. It remains to notice that the spectrum of H_{eff} consists of two parts that correspond to the effective spin states $|+\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big)$ and $|-\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle - |1\rangle\big)$. Since $B_1 B_2 B_3$ is positive semidefinite, the smallest eigenvalue is associated with $|+\rangle$. Hence, the lowest eigenvalue of $\widetilde{H}$ is equal to the lowest eigenvalue of (6.3), as required.

6.3. Reduction from 3-Local Hamiltonian to 2-Local Hamiltonian. In this section we reduce the 3-LOCAL HAMILTONIAN problem to the 2-LOCAL HAMILTONIAN problem. By the QMA-completeness of the 3-LOCAL HAMILTONIAN problem [11], this establishes Theorem 1.1.

THEOREM 6.7. *There is a polynomial time reduction from the* 3-LOCAL HAMILTONIAN *problem to the* 2-LOCAL HAMILTONIAN *problem.*

Proof. Recall that in the 3-LOCAL HAMILTONIAN problem (see Definition 2.3) we are given two constants a and b and a local Hamiltonian $H^{(3)} = \sum_j H_j$ such that each H_j is a three-qubit term whose norm is at most $\text{poly}(n)$. Our goal in this proof is to transform $H^{(3)}$ into a 2-local Hamiltonian $H^{(2)}$ whose lowest eigenvalue is close to that of $H^{(3)}$. We do this in two steps. The first is a somewhat technical step where we bring $H^{(3)}$ into a convenient form. In the second step, we replace each 3-local term with 2-local terms by using the gadget construction of the previous section. Before we continue with the proof, let us mention that it is crucial that we apply the gadget construction to all 3-local terms *simultaneously*. If instead we tried to apply the gadget construction sequentially, we would end up with an exponential blowup in the norms (since each application of the three-qubit gadget increases the norm by a multiplicative factor).

LEMMA 6.8. *The 3-local Hamiltonian $H^{(3)}$ can be represented as*

$$H^{(3)} = c_r \left(Y - 6 \sum_{m=1}^{M} B_{m1} B_{m2} B_{m3} \right),$$

where Y is a 2-local Hamiltonian with $\|Y\| = O(1/n^6)$, $M = O(n^3)$, each B_{mi} is a one-qubit term of norm $O(1/n^3)$ that satisfies $B_{mi} \geq \frac{1}{n^3} I$, and c_r is a rescaling factor satisfying $1 \leq c_r \leq \text{poly}(n)$.[12]

Proof. First, we can assume without loss of generality that each H_j acts on a different triple of qubits, and hence there are at most n^3 such terms. Recall that any three-qubit Hermitian operator can be written as a linear combination with real coefficients of the basis elements $\sigma^\alpha \otimes \sigma^\beta \otimes \sigma^\gamma$, where each of $\sigma^\alpha, \sigma^\beta, \sigma^\gamma$ ranges over the four possible Pauli matrices $\{I, \sigma^x, \sigma^y, \sigma^z\}$. Hence, for $M = O(n^3)$, we can write

$$H^{(3)} = c_r \left(-6 \sum_{m=1}^{M} c_m \cdot \sigma^{m,\alpha} \otimes \sigma^{m,\beta} \otimes \sigma^{m,\gamma} \right),$$

where each $\sigma^{m,\alpha}$ is a Pauli matrix acting on one of the qubits, and $c_r \leq \text{poly}(n)$ is chosen to be large enough so that $|c_m| \leq \frac{1}{n^9}$ for all $m = 1, \ldots, M$.

---

[12]For the proof of Theorem 6.7 we need only the property $B_{mi} \geq 0$. The stronger property $B_{mi} \geq \frac{1}{n^3} I$ will be used in section 7.

We finish the proof by writing each $c_m \sigma^{m,\alpha} \otimes \sigma^{m,\beta} \otimes \sigma^{m,\gamma}$ as

$$
\underbrace{\left(\frac{2}{n^3}I + n^6 c_m \sigma^{m,\alpha}\right)}_{B_{m1}} \otimes \underbrace{\left(\frac{2}{n^3}I + \frac{1}{n^3}\sigma^{m,\beta}\right)}_{B_{m2}} \otimes \underbrace{\left(\frac{2}{n^3}I + \frac{1}{n^3}\sigma^{m,\gamma}\right)}_{B_{m3}} + D_m,
$$

where D_m is 2-local. Since $|c_m| \leq 1/n^9$ we have that $B_{mi} \geq \frac{1}{n^3}I$ and $\|D_m\| = O(1/n^9)$. $\square$

We now replace each term $-6B_{m1}B_{m2}B_{m3}$ by a three-qubit gadget. More specifically, let δ be a sufficiently small inverse polynomial in n to be chosen later. We consider the Hamiltonian $H^{(2)} = c_r \widetilde{H}$, $\widetilde{H} = H + V$, acting on a system of $n + 3M$ qubits, where

$$
H = -\frac{\delta^{-3}}{4} \sum_{m=1}^{M} I \otimes \left(\sigma^z_{m1}\sigma^z_{m2} + \sigma^z_{m1}\sigma^z_{m3} + \sigma^z_{m2}\sigma^z_{m3} - 3I\right),
$$

$$
V = Y \otimes I + \delta^{-1} \sum_{m=1}^{M} (B^2_{m1} + B^2_{m2} + B^2_{m3}) \otimes I
$$

(6.5)
$$
- \delta^{-2} \sum_{m=1}^{M} \left(B_{m1} \otimes \sigma^x_{m1} + B_{m2} \otimes \sigma^x_{m2} + B_{m3} \otimes \sigma^x_{m3}\right).
$$

As before, let $\Delta = \delta^{-3}$ be the spectral gap of H. Notice that the spectrum of H includes not only 0 and Δ but also $2\Delta, 3\Delta, \ldots, M\Delta$. Associated with the zero eigenvalue is the subspace spanned by all the zero-subspaces of the gadgets. Using $\|B_{mi}\| \leq O(1/n^3)$ and $M = O(n^3)$ we get $\|V\| = O(\delta^{-2}) < \Delta/2$.

The calculation of Σ_- is quite similar to the one-gadget case (cf. (6.4)). Each gadget contributes an independent term. Terms up to the third order can include only processes that involve one gadget. Indeed, in order to involve two gadgets, one has to flip a qubit from one gadget and from another gadget, and then flip both qubits back. Moreover, since only one gadget is involved, G_+ can be replaced by $(z - \Delta)^{-1} I_{\mathcal{L}_+}$ as before. From the fourth order onwards, processes start to include cross-terms between different gadgets. However, we claim that their contribution is only $O(\delta)$, as long as $|z| = O(1)$. Indeed, in this range, the eigenvalues of G_+, which are $(z - \Delta)^{-1}$, $(z - 2\Delta)^{-1}$, $\ldots$, are all at most $O(\delta^3)$ in absolute value, while the norm of each of the V terms is at most $O(\delta^{-2})$. To summarize, for $|z| = O(1)$,

(6.6)
$$
\Sigma_-(z) = \underbrace{Y \otimes I_{\mathcal{C}} - 6 \sum_{m=1}^{M} B_{m1}B_{m2}B_{m3} \otimes \left(\sigma^x_m\right)_{\text{eff}}}_{H_{\text{eff}}} + O(\delta).
$$

Since $\|H_{\text{eff}}\| \leq O(1)$, we can apply Theorem 6.2 with $c = -\|H_{\text{eff}}\|$, $d = \|H_{\text{eff}}\|$, and $\lambda_* = \Delta/2$. We obtain that the smallest eigenvalue of $\widetilde{H}$ is $O(\delta)$-close to that of H_{eff}. The spectrum of H_{eff} consists of 2^M parts, corresponding to subspaces spanned by setting each effective spin state to either $|+\rangle$ or $|-\rangle$. Since $B_{m1}B_{m2}B_{m3} \geq 0$, the smallest eigenvalue of H_{eff} is achieved in the subspace where all effective spin states are in the $|+\rangle$ state. In this subspace, H_{eff} is identical to $H^{(3)}/c_r$. Hence, the smallest eigenvalue of $H^{(2)} = c_r \widetilde{H}$ is $O(c_r \delta)$-close to that of $H^{(3)}$. We complete the proof by choosing $\delta = c'/c_r$ for some small enough constant c'. $\square$

7. 2-local universal adiabatic computation. In this section we show that adiabatic computation with 2-local Hamiltonians is equivalent to "standard" quantum computation in the circuit model. In order to prove such an equivalence, one has to show that each model can simulate the other. One direction is already known: it is not too hard to show that any polynomial time adiabatic computation can be efficiently simulated by a quantum circuit [8]. Hence, it remains to show that adiabatic computation with 2-local Hamiltonians can efficiently simulate any quantum circuit. In [3] it is shown that adiabatic computation with 3-local Hamiltonians can efficiently simulate any quantum circuit. We obtain our result by combining their result with the techniques in our second proof.

Let us briefly mention the main ideas behind adiabatic computation. For more details see [3] and the references therein. In adiabatic computation, we consider a time-dependent Hamiltonian $H(s)$ for $s \in [0, 1]$ acting on a quantum system. We initialize the system in the groundstate of the initial Hamiltonian $H(0)$. This ground-state is required to be some simple quantum state that is easy to create. We then slowly modify the Hamiltonian from $s = 0$ to $s = 1$. We say that the adiabatic computation is *successful* if the final state of the system is close to the groundstate of $H(1)$. The adiabatic theorem (see, e.g., [18, 4]) says that if the Hamiltonian is modified slowly enough, the adiabatic computation is successful. In other words, it gives an upper bound on the running time of an adiabatic computation. For our purposes, it is enough to know that this bound is polynomial if for any $s \in [0, 1]$, the norm of $H(s)$, as well as those of its first and second derivatives, is bounded by a polynomial, and the spectral gap of $H(s)$ is larger than some inverse polynomial.

In [3] it is shown how to transform an arbitrary quantum circuit into an efficient 3-local adiabatic computation. To establish this, they define a 3-local time-dependent Hamiltonian $H^{(3)}(s)$ with the following properties. First, the Hamiltonian acts on a system of n qubits, where n is some constant times the number of gates in the circuit. Second, the groundstate of $H^{(3)}(0)$ is very easy to create (namely, it is the all zero state), and the groundstate of $H^{(3)}(1)$ is some state that encodes the result of the quantum circuit. Third, for all $s \in [0, 1]$, the spectral gap of $H^{(3)}(s)$ is bounded from below by an inverse polynomial in n, and the norm of $H^{(3)}(s)$, as well as those of its first and second derivatives, is bounded by some polynomial in n. Together with the adiabatic theorem, these properties imply that adiabatic computation according to $H^{(3)}(s)$ is efficient. Finally, let us mention that $H^{(3)}(s)$, as defined in [3], is linear in s; that is, $H^{(3)}(s) = (1-s)H^{(3)}(0) + sH^{(3)}(1)$. This property will be useful in our proof.

The following is the main theorem of this section.

THEOREM 7.1. *Any quantum computation can be efficiently simulated by an adiabatic computation with 2-local Hamiltonians.*

Proof. Given a quantum circuit, let $H^{(3)}(s)$ be the time-dependent Hamiltonian of [3] as described above. The idea of the proof is to apply the gadget construction of section 6.3 to $H^{(3)}(s)$ for any $s \in [0, 1]$, thereby creating a 2-local time-dependent Hamiltonian $H^{(2)}(s)$. Some care needs to be taken to ensure that the resulting time-dependent Hamiltonian is smooth enough as a function of s. We therefore describe how this is done in more detail.

We start by writing $H^{(3)}(s)$ in a form similar to that given by Lemma 6.8. Since $H^{(3)}(s)$ is linear in s, we can write

$$H^{(3)}(s) = c_r \left(-6 \sum_{m=1}^{M} c_m(s) \cdot \sigma^{m,\alpha} \otimes \sigma^{m,\beta} \otimes \sigma^{m,\gamma} \right),$$

where $M = O(n^3)$, each $c_m(s)$ is a linear function of s, and $c_r \leq \text{poly}(n)$ is chosen to be large enough so that $|c_m(s)| \leq \frac{1}{n^9}$ for all m and all $s \in [0,1]$. Notice that c_r is a fixed scaling factor used for all $s \in [0,1]$. Following the proof of Lemma 6.8, we write

$$H^{(3)}(s) = c_r \left(Y(s) - 6 \sum_{m=1}^{M} B_{m1}(s) B_{m2} B_{m3} \right),$$

where by our construction, $Y(s)$ and $B_{m1}(s)$ are linear in s, whereas B_{m2} and B_{m3} are independent of s. Finally, we define $H^{(2)}(s) = c_r \tilde{H}(s)$, where $\tilde{H}(s) = H + V(s)$ and the Hamiltonians H and $V(s)$ are defined as in (6.5). The parameter δ will be chosen later to be some small enough inverse polynomial in n.

In the rest of the proof, we show that adiabatic computation according to $H^{(2)}(s)$ can be used to simulate the given quantum circuit. We start by proving two lemmas that, together with the adiabatic theorem, imply that the running time of the adiabatic computation is polynomial in n.

LEMMA 7.2. *For any* $s \in [0,1]$, $\|H^{(2)}(s)\|$, $\|\frac{d}{ds} H^{(2)}(s)\|$, *and* $\|\frac{d^2}{ds^2} H^{(2)}(s)\|$ *are upper bounded by a polynomial in* n.

Proof. Recall that $Y(s)$ and $B_{m1}(s)$ are linear in s. Together with the definition of $H^{(2)}$, this implies that $H^{(2)}(s)$ is a degree-two polynomial in s, i.e., we can write $H^{(2)}(s) = A + sB + s^2C$ for some Hermitian matrices A, B, C. It is not hard to see that the norm of each of these matrices is bounded by some polynomial in n. This implies that the norms of $H^{(2)}(s)$, its first derivative $B + 2sC$, and its second derivative $2C$ are bounded by some polynomial in n. $\square$

LEMMA 7.3. *For any* $s \in [0,1]$, *the spectral gap of* $H^{(2)}(s)$ *is lower bounded by an inverse polynomial in* n.

Proof. As shown in section 6.3, the lower part of the spectrum of $H^{(2)}(s)$ is $O(c_r \delta)$-close to the spectrum of $c_r H_{\text{eff}}(s)$. Hence, by choosing δ to be a small enough inverse polynomial in n, we see that it is enough to show that the spectral gap of $c_r H_{\text{eff}}(s)$ is at least some inverse polynomial in n.

The spectrum of $c_r H_{\text{eff}}(s)$ consists of 2^M parts, corresponding to all possible settings for the effective qubits. The part corresponding to the subspace in which all effective qubits are in the $|+\rangle$ state is identical to the spectrum of $H^{(3)}(s)$. Hence, we know that in this subspace the spectral gap is at least some inverse polynomial in n. We now claim that the lowest eigenvalue in all other $2^M - 1$ subspaces is greater than that in the all $|+\rangle$ subspace by at least some inverse polynomial in n. Indeed, the restriction of $c_r H_{\text{eff}}(s)$ to any such subspace is given by $H^{(3)}(s)$ plus a nonzero number of terms of the form $12 c_r B_{m1}(s) B_{m2} B_{m3}$. The claim follows from the fact that $B_{m1}(s) B_{m2} B_{m3} \geq \frac{1}{n^9} I$. $\square$

To complete the proof, we need to argue about the groundstate of $H^{(2)}(0)$ and that of $H^{(2)}(1)$. To this end, we use the following lemma, which essentially says that if H_{eff} has a spectral gap, then Theorem 6.2 not only implies closeness in spectra but also in the groundstates.

LEMMA 7.4. *Assume that* H, V, H_{eff} *satisfy the conditions of Theorem* 6.2 *with some* $\varepsilon > 0$. *Let* $\lambda_{\text{eff},i}$ *denote the* ith *eigenvalue of* H_{eff} *and* $|\tilde{v}\rangle$ (*resp.*, $|v_{\text{eff}}\rangle$) *denote the groundstate of* $\tilde{H}$ (*resp.*, H_{eff}). *Then, under the assumption* $\lambda_{\text{eff},2} > \lambda_{\text{eff},1}$,

$$|\langle \tilde{v} | v_{\text{eff}} \rangle| \geq 1 - \frac{2\|V\|^2}{(\lambda_+ - \lambda_{\text{eff},1} - \varepsilon)^2} - \frac{4\varepsilon}{\lambda_{\text{eff},2} - \lambda_{\text{eff},1}}.$$

Before we prove the lemma, let us complete the proof of the theorem. Recall that in our case $\varepsilon = O(\delta)$, $\|V\| = O(\delta^{-2})$, $\lambda_+ = \delta^{-3}$, $|\lambda_{\text{eff},1}| \leq O(1)$, and $\lambda_{\text{eff},2} - \lambda_{\text{eff},1} = 1/\text{poly}(n)$. Hence, the first error term in the above bound is $O(\delta^2)$ while the second is $O(\delta \cdot \text{poly}(n))$. Therefore, by choosing δ to be a small enough inverse polynomial in n, we can guarantee that the groundstate of $H^{(2)}(s)$ is close to the groundstate of $H_{\text{eff}}(s)$. In particular, the groundstate of $H^{(2)}(1)$, which is the output of the adiabatic computation, is close to the groundstate of $H_{\text{eff}}(1)$. The latter is $|v_1\rangle \otimes |+\rangle^{\otimes M}$, where $|v_1\rangle$ is the groundstate of $H^{(3)}(1)$. By simply tracing out the $3M$ gadget qubits, we can recover $|v_1\rangle$ from this groundstate and therefore obtain the output of the quantum circuit. Similarly, the groundstate of $H^{(2)}(0)$, which is the state to which the system should be initialized, is close to the groundstate of $H_{\text{eff}}(0)$. The latter is $|v_0\rangle \otimes |+\rangle^{\otimes M}$, where $|v_0\rangle$ is the groundstate of $H^{(3)}(0)$. We therefore initialize the system by setting the original n qubits to $|v_0\rangle$ and the M gadgets to the effective $|+\rangle$ state. This state is close to the groundstate of $H^{(2)}(0)$, and since the adiabatic computation is unitary, this approximation does not affect the output by much.

It remains to prove the lemma.

Proof of Lemma 7.4. Let $|\widetilde{v}_-\rangle = \Pi_- |\widetilde{v}\rangle / \|\Pi_- |\widetilde{v}\rangle\|$ be the normalized projection of $|\widetilde{v}\rangle$ on the space $\mathcal{L}_-$. We first show that $|\widetilde{v}_-\rangle$ is close to $|\widetilde{v}\rangle$. By Theorem 6.2, we know that $\widetilde{\lambda}_1 \leq \lambda_{\text{eff},1} + \varepsilon$. Hence,

$$\|\Pi_+ \widetilde{H} |\widetilde{v}\rangle\| = \widetilde{\lambda}_1 \|\Pi_+ |\widetilde{v}\rangle\| \leq (\lambda_{\text{eff},1} + \varepsilon)\|\Pi_+ |\widetilde{v}\rangle\|$$

and

$$\|\Pi_+ \widetilde{H} |\widetilde{v}\rangle\| = \|\Pi_+ H |\widetilde{v}\rangle + \Pi_+ V |\widetilde{v}\rangle\| \geq \|\Pi_+ H |\widetilde{v}\rangle\| - \|V\| \geq \lambda_+ \|\Pi_+ |\widetilde{v}\rangle\| - \|V\|.$$

By combining the two inequalities we obtain

$$\|\Pi_+ |\widetilde{v}\rangle\| \leq \frac{\|V\|}{\lambda_+ - \lambda_{\text{eff},1} - \varepsilon},$$

from which we see that

$$\alpha \stackrel{def}{=} |\langle \widetilde{v} | \widetilde{v}_- \rangle| = \|\Pi_- |\widetilde{v}\rangle\| \geq \|\Pi_- |\widetilde{v}\rangle\|^2 \geq 1 - \frac{\|V\|^2}{(\lambda_+ - \lambda_{\text{eff},1} - \varepsilon)^2}.$$

Our next step is to show that $|\widetilde{v}_-\rangle$ is close to $|v_{\text{eff}}\rangle$. For this we need to consider the proof of Theorem 6.2. We start by taking Lemma 6.4 with $\widetilde{\lambda} = \widetilde{\lambda}_1$. The lemma says that A is a matrix of rank 1. By looking at the proof, it is easy to see that A is in fact $\Pi_- |\widetilde{v}\rangle\langle\widetilde{v}|\Pi_-$. Next, Lemma 6.5 implies that $\widetilde{\lambda}_1$ is an eigenvalue of multiplicity 1 of $\Sigma_-(\widetilde{\lambda}_1)$. In fact, from the proof it follows that the corresponding eigenvector is exactly $\Pi_- |\widetilde{v}\rangle$ (since the null space of C is equal to the span of A). By normalizing, this is exactly $|\widetilde{v}_-\rangle$. But by our assumption, $\|\Sigma_-(z) - H_{\text{eff}}\| \leq \varepsilon$ for all $z \in [c - \varepsilon, d + \varepsilon]$ and in particular

$$\|\Sigma_-(\widetilde{\lambda}_1) - H_{\text{eff}}\| \leq \varepsilon.$$

From this we obtain that

$$\left| \langle \widetilde{v}_- | (\Sigma_-(\widetilde{\lambda}_1) - H_{\text{eff}}) | \widetilde{v}_- \rangle \right| \leq \varepsilon$$

and hence

$$\langle \widetilde{v}_- | H_{\text{eff}} | \widetilde{v}_- \rangle \leq \widetilde{\lambda}_1 + \varepsilon \leq \lambda_{\text{eff},1} + 2\varepsilon,$$

where we again used that $\widetilde{\lambda}_1 \leq \lambda_{\text{eff},1} + \varepsilon$. Since H_{eff} has a spectral gap, this indicates that $|\widetilde{v}_-\rangle$ must be close to $|v_{\text{eff}}\rangle$. Indeed, let $\beta = |\langle \widetilde{v}_- | v_{\text{eff}}\rangle|$. Then,

$$\langle \widetilde{v}_- | H_{\text{eff}} | \widetilde{v}_- \rangle \geq \beta^2 \lambda_{\text{eff},1} + (1 - \beta^2)\lambda_{\text{eff},2} = \lambda_{\text{eff},1} + (1 - \beta^2)(\lambda_{\text{eff},2} - \lambda_{\text{eff},1}).$$

By combining the two inequalities we obtain

$$1 - \beta^2 \leq \frac{2\varepsilon}{\lambda_{\text{eff},2} - \lambda_{\text{eff},1}}.$$

Summarizing,

$$|\langle \widetilde{v} | v_{\text{eff}}\rangle| = |\langle \widetilde{v} | \widetilde{v}_- \rangle \langle \widetilde{v}_- | v_{\text{eff}}\rangle + \langle \widetilde{v} | (I - |\widetilde{v}_-\rangle\langle \widetilde{v}_-|) | v_{\text{eff}}\rangle|$$

$$\geq \alpha \cdot \beta - \sqrt{(1 - \alpha^2)(1 - \beta^2)} \geq \alpha \cdot \beta - \frac{1}{2}\big((1 - \alpha^2) + (1 - \beta^2)\big)$$

$$\geq \big(1 - (1 - \alpha) - (1 - \beta)\big) - \big((1 - \alpha) + (1 - \beta)\big) = 1 - 2(1 - \alpha) - 2(1 - \beta)$$

$$\geq 1 - \frac{2\|V\|^2}{(\lambda_+ - \lambda_{\text{eff},1} - \varepsilon)^2} - \frac{4\varepsilon}{\lambda_{\text{eff},2} - \lambda_{\text{eff},1}}. \qquad \square$$

8. Conclusion. Some interesting open questions remain. First, perturbation theory has allowed us to perform the first reduction *inside* QMA. What other problems can be solved using this technique? Second, there exists an intriguing class between NP (in fact, MA) and QMA known as QCMA. It is the class of problems that can be verified by a quantum verifier with a *classical* proof. Can one show a separation between QCMA and QMA, or perhaps show that they are equal? Third, Kitaev's original 5-local proof has the following desirable property: For any YES instance produced by the reduction there exists a state such that each individual 5-local term is very close to its groundstate. Note that this is a stronger property than the one required in the LOCAL HAMILTONIAN problem. Using a slight modification of Kitaev's original construction, one can show a reduction to the 4-LOCAL HAMILTONIAN problem that has the same property. However, we do not know if this property can be achieved for the 3-local or the 2-local problem.

REFERENCES

[1] A. A. ABRIKOSOV, L. P. GORKOV, AND I. E. DZYALOSHINSKI, *Methods of Quantum Field Theory in Statistical Physics*, Dover, New York, 1975.

[2] D. AHARONOV AND T. NAVEH, *Quantum NP—A survey*, 2002; available online from http://www.arxiv.org/abs/quant-ph/0210077.

[3] D. AHARONOV, W. VAN DAM, J. KEMPE, Z. LANDAU, S. LLOYD, AND O. REGEV, *Adiabatic quantum computation is equivalent to standard quantum computation*, in Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science, 2004, pp. 42–51.

[4] A. AMBAINIS AND O. REGEV, *An elementary proof of the quantum adiabatic theorem*, 2004; available online from http://www.arxiv.org/abs/quant-ph/0411152.

[5] D. BARENCO, C. H. BENNETT, R. CLEVE, D. P. DIVINCENZO, N. MARGOLUS, P. SHOR, T. SLEATOR, J. SMOLIN, AND H. WEINFURTER, *Elementary gates for quantum computation*, Phys. Rev. A, 52 (1995), pp. 3457–3467.

[6] R. BHATIA, *Matrix Analysis*, Grad. Texts in Math. 169, Springer-Verlag, New York, 1997.

[7] S. BRAVYI AND M. VYALYI, *Commutative version of the k-local Hamiltonian problem and non-triviality check for quantum codes*, Quantum Inf. Comput., 5 (2005), pp. 187–215.

[8] E. FARHI, J. GOLDSTONE, S. GUTMANN, AND M. SIPSER, *Quantum computation by adiabatic evolution*, 2000; available online from http://www.arxiv.org/abs/quant-ph/0001106.

[9] D. Janzing, P. Wocjan, and T. Beth, *"Non-identity-check" is QMA-complete*, Internat. J. Quantum Inform., 3 (2005), pp. 463–473.

[10] J. Kempe, A. Kitaev, and O. Regev, *The complexity of the local Hamiltonian problem*, in Proceedings of the 24th Annual FSTTCS, Lecture Notes in Comput. Sci. 3328, Springer-Verlag, Berlin, 2004, pp. 372–383.

[11] J. Kempe and O. Regev, *3-local Hamiltonian is QMA-complete*, Quantum Inf. Comput., 3 (2003), pp. 258–264.

[12] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and quantum computation*, Grad. Stud. Math. 47, AMS, Providence, RI, 2002.

[13] E. Knill, *Quantum randomness and nondeterminism*, 1996; available online from http://www.arxiv.org/abs/quant-ph/9610012.

[14] C. Marriott and J. Watrous, *Quantum Arthur–Merlin games*, in Proceedings of the 19th IEEE Annual Conference on Computational Complexity (CCC), 2004, pp. 344–354.

[15] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, UK, 2000.

[16] R. Oliveira and B. Terhal, *The complexity of quantum spin systems on a two-dimensional square lattice*, 2005; available online from http://www.arxiv.org/abs/quant-ph/0504050.

[17] C. Papadimitriou, *Computational Complexity*, Addison–Wesley, Reading, MA, 1994.

[18] B. Reichardt, *The quantum adiabatic optimization algorithm and local minima*, in Proceedings of the 36th Annual STOC, ACM, 2004, pp. 502–510.

[19] W. Rudin, *Functional Analysis*, 2nd ed., International Series in Pure and Applied Mathematics, McGraw–Hill Inc., New York, 1991.

[20] J. Watrous, *Succinct quantum proofs for properties of finite groups*, in Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science, 2000, pp. 537–546.

[21] P. Wocjan and T. Beth, *The 2-local Hamiltonian problem encompasses NP*, Internat. J. Quantum Inform., 1 (2003), pp. 349–357.