A FULL CHARACTERIZATION OF QUANTUM ADVICE*

SCOTT AARONSON† AND ANDREW DRUCKER‡

Abstract. We prove the following surprising result: given any quantum state \( \rho \) on \( n \) qubits, there exists a local Hamiltonian \( H \) on \( \text{poly}(n) \) qubits (e.g., a sum of two-qubit interactions), such that any ground state of \( H \) can be used to simulate \( \rho \) on all quantum circuits of fixed polynomial size. In terms of complexity classes, this implies that \( \text{BQP/qpoly} \subseteq \text{QMA/poly} \), which supersedes the previous result of Aaronson that \( \text{BQP/qpoly} \subseteq \text{PP/poly} \). Indeed, we can exactly characterize quantum advice as equivalent in power to untrusted quantum advice combined with trusted classical advice. Proving our main result requires combining a large number of previous tools—including a result of Alon et al. on learning of real-valued concept classes, a result of Aaronson on the learnability of quantum states, and a result of Aharonov and Regev on “QMA+ super-verifiers”—and also creating some new ones. The main new tool is a so-called majority-certificates lemma, which is closely related to boosting in machine learning, and which seems likely to find independent applications. In its simplest version, this lemma says the following. Given any set \( S \) of Boolean functions on \( n \) variables, any function \( f \in S \) can be expressed as the pointwise majority of \( m = O(n) \) functions \( f_1, \ldots, f_m \in S \), such that each \( f_i \) is the unique function in \( S \) compatible with \( O(\log |S|) \) input/output constraints.

Key words. quantum computation, learning, compression, advice, local Hamiltonians, nonuniform computation, boosting, Karp–Lipton theorem

AMS subject classification. 81P68

DOI. 10.1137/110856939

1. Introduction. How much classical information is needed to specify a quantum state of \( n \) qubits?

This question has inspired a rich and varied set of responses, in part because it can be interpreted in many ways. If we want to specify a quantum state \( \rho \) exactly, then of course the answer is “an infinite amount,” since amplitudes in quantum mechanics are continuous. A natural compromise is to try to specify \( \rho \) approximately, i.e., to give a description which yields a state \( \tilde{\rho} \) whose statistical behavior is close to that of \( \rho \) under every measurement. (This statement is captured by the requirement that \( \rho \) and \( \tilde{\rho} \) are close under the so-called trace distance metric.) But it is not hard to see that even for this task, we still need to use an exponential (in \( n \)) number of classical bits.

This fact can be viewed as a disappointment, but also as an opportunity, since it raises the prospect that we might be able to encode massive amounts of information in physically compact quantum states: for example, we might hope to store \( 2^n \) classical bits in \( n \) qubits. But an obvious practical requirement is that we be able to retrieve the information reliably, and this rules out the hope of significant “quantum compression” of classical strings, as shown by a landmark result of Holevo [20] from 1973. Consider

---

*Received by the editors November 28, 2011; accepted for publication (in revised form) February 3, 2014; published electronically May 15, 2014. A preliminary extended abstract of this work appeared in ACM STOC 2010.

http://www.siam.org/journals/sicomp/43-3/85693.html

†Department of Electrical Engineering and Computer Science, MIT EECS, Cambridge, MA 02139 (aaronson@csail.mit.edu). This material is based upon work supported by the National Science Foundation under grant 0844626 and was also supported by a DARPA YFA grant, a Sloan Fellowship, and the Keck Foundation.

‡Institute for Advanced Study, School of Mathematics, Princeton, NJ 08540 (andy.drucker@gmail.com). This author was supported by the NSF under Princeton University prime award CCF-0832797 and sub-contract 00001583. This material is based upon work supported by an Akamai Presidential Graduate Fellowship while the author was a graduate student at MIT.
a sender Alice and a recipient Bob, with a one-way quantum channel between them. Then Holevo's theorem says that if Alice wants to encode an $n$-bit classical string $x$ into an $m$-qubit quantum state $\rho_x$, in such a way that Bob can retrieve $x$ (with probability $2/3$, say) by measuring $\rho_x$, then Alice must take $m \geq n - O(1)$ (or $m \geq n/2 - O(1)$ if Alice and Bob share entanglement). In other words, for this communication task, quantum states offer essentially no advantage over classical strings. In 1999, Nayak [27], improving on Ambainis et al. [11] (see [12]), generalized Holevo's result as follows: even if Bob wants to learn only a single bit $x_i$ of $x = x_1 \ldots x_n$ (for some $i \in [n]$ unknown to Alice) and is willing to destroy the state $\rho_x$ in the process of learning that bit, Alice still needs to send $m = \Omega(n)$ qubits for Bob to succeed with high probability.

These results say that the exponential descriptive complexity of quantum states cannot be effectively harnessed for classical data storage, but they do not bound the number of practically meaningful “degrees of freedom” in a quantum state used for purposes other than storing data. For example, a quantum state could be useful for computation, or it could be a physical system worthy of study in its own right. The question then becomes, What useful information can we give about an $n$-qubit state using a “reasonable” number (say, poly$(n)$) of classical bits?

One approach to this question is to identify special subclasses of quantum states for which a faithful approximation can be specified using only poly$(n)$ bits. This has been done, for example, with matrix product states [31] and “tree states” [1].

A second approach is to try to describe an arbitrary $n$-qubit state $\rho$ concisely, in such a way that the state $\tilde{\rho}$ recovered from the description is close to $\rho$ with respect to some natural subclass of measurements. This has been done for specific classes like the “pretty good measurements” of Hausladen and Wootters [19].

A more ambitious goal in this vein, explored by Aaronson in two previous works [2, 5] and continued in the present paper, is to give a description of an $n$-qubit state $\rho$ which yields a state $\tilde{\rho}$ that behaves approximately like $\rho$ with respect to all (binary) measurements performable by quantum circuits of reasonable size, say, of size at most $n^c$ for some fixed $c > 0$. Then if $c$ is taken large enough, $\tilde{\rho}$ is arguably “just as good” as $\rho$ for practical purposes.

Certainly we can achieve this goal using $2^{n+O(1)}$ bits: simply give approximations to the measurement statistics for every size-$n^c$ circuit. However, the results of Holevo [20] and Ambainis et al. [12] suggest that a much more succinct description might be possible. This hope was realized by Aaronson [2], who gave a description scheme in which an $n$-qubit state can be specified using poly$(n)$ classical bits. There is a significant catch in Aaronson’s result, though: the encoder Alice and decoder Bob both need to invest exponential amounts of computation.

In a subsequent paper [5], Aaronson gave a closely related result which significantly reduces the computational requirements: now Alice can generate her message in polynomial time (for fixed $c$). Also, while Bob cannot necessarily construct the state $\tilde{\rho}$ efficiently on his own, if he is presented with such a state (by an untrusted prover, say), Bob can verify the state in polynomial time. The catch in this result is a weakened approximation guarantee: Bob cannot use $\tilde{\rho}$ to predict the outcomes of all the measurements defined by size-$n^c$ circuits but only most of them (with respect to a samplable distribution used by Alice in the encoding process). Aaronson conjectured [5] that the tradeoff between the results of [5] and of [2] revealed an inherent limit to quantum compression.

1.1. Our quantum information result. The main result of this paper is that Aaronson’s conjecture was false: one really can get the best of both worlds and simulate an arbitrary quantum state $\rho$ on all small circuits, using a different state.
quantum circuit of size \( n \) that is easy to recognize. Indeed, we can even take \( \rho \) to be the ground state of a local Hamiltonian, that is, a pure state \( \rho = |\psi\rangle\langle\psi| \) on \( \text{poly}(n) \) qubits minimizing the disagreement with \( \text{poly}(n) \) local constraints, each involving a constant number of qubits. In a sense, then, this paper completes a “trilogy” of which [2, 5] were the first two installments.

Here is a formal statement of our result.

**Theorem 1.** Let \( c, \delta > 0 \), and let \( \rho^* \) be any \( n \)-qubit quantum state. Then there exists a 2-local Hamiltonian \( H \) on \( \text{poly}(n, 1/\delta) \) qubits, and a transformation \( C \rightarrow C' \) of quantum circuits, computable in time \( \text{poly}(n, 1/\delta) \) given \( H \), such that the following holds: for any ground state \( |\psi\rangle \) of \( H \), and for any measurement \( C \) definable by a quantum circuit of size \( n^c \), we have \( |\mathbb{E}[C'(|\psi\rangle\langle\psi|)] - \mathbb{E}[C(\rho^*)]| \leq \delta \).

In other words, the ground states of local Hamiltonians are “universal quantum states” in a very nonobvious sense. For example, if \( \rho \) is supposed to help a customer’s quantum computer \( Q \) evaluate some Boolean function \( f : \{0, 1\}^n \rightarrow \{0, 1\} \), then \( Q(\rho, x) \) should output \( f(x) \) for every input \( x \in \{0, 1\}^n \). By contrast, any \( k \)-local Hamiltonian \( H \) can be described as a set of at most \( \binom{n}{k} = O(n^k) \) constraints.

One can also interpret Theorem 1 as a statement about communication over quantum channels. Suppose Alice (who is computationally unbounded) has a classical description of an \( n \)-qubit state \( \rho^* \). She would like to describe this state to Bob (who is computationally bounded), at least well enough for Bob to be able to simulate \( \rho^* \) on all quantum circuits of some fixed polynomial size. However, Alice cannot just send \( \rho^* \) to Bob, since her quantum communication channel is noisy and there is a chance that the state might get corrupted along the way. Nor can she send a faithful classical description of \( \rho^* \), since that would require an exponential number of bits. Our result provides an alternative: Alice can send a different quantum state \( \sigma \), of poly(\( n \))-bit classical string \( x \). Then, Bob can use \( x \) to verify that \( \sigma \) can be used to accurately simulate \( \rho^* \) on all small measurements.

We believe Theorem 1 makes a significant contribution to the study of the effective information content of quantum states. It does, however, leave open whether a quantum state of \( n \) qubits can be efficiently encoded and decoded in polynomial time in a way that is “good enough” to preserve the measurement statistics of measurements defined by circuits of fixed polynomial size. This remains an important problem for future work.

**1.2. Impact on quantum complexity theory.** The questions addressed in this paper, and our results, are naturally phrased and proved in terms of complexity classes. In recent years, researchers have defined quantum complexity classes as a way to study the “useful information” embodied in quantum states. One approach is to study the power of nonuniform quantum advice. The class \( \text{BQP}/\text{qpoly} \), defined by Nishimura and Yamakami [28], consists of all languages decidable in polynomial time by a quantum computer, with the help of a poly(\( n \))-qubit advice state that depends only on the input length \( n \). This class is analogous to the classical class \( \text{P}/\text{poly} \). To understand the role of quantum information in determining the power of \( \text{BQP}/\text{qpoly} \), a useful benchmark of comparison is the class \( \text{BQP}/\text{poly} \) of decision problems efficiently solvable by a quantum algorithm with poly(\( n \)) bits of classical
advice (or equivalently, by a nonuniform family of poly(n)-sized quantum circuits). It is open whether BQP/qpoly = BQP/poly.

A second approach studies the power of quantum proof systems, by analogy with the classical class NP. Kitaev (unpublished, 1999 (see [26])) defined the complexity class now called QMA, for “quantum Merlin–Arthur.” This is the class of decision problems for which a “yes” answer can be proved by exhibiting a quantum witness state (or quantum proof) |ψ⟩ on poly(n) qubits, which is then checked by a skeptical polynomial-time quantum verifier. A useful benchmark class is QCMA (for “quantum classical Merlin–Arthur”), defined by Aharonov and Naveh [7]. This is the class of decision problems for which a “yes” answer can be checked by a quantum verifier who receives a classical witness. Here the natural open question is whether QMA = QCMA.

In this paper we prove a new upper bound on BQP/qpoly.

**Theorem 2.** BQP/qpoly ⊆ QMA/poly.

Previously Aaronson showed in [2] that BQP/qpoly ⊆ PP/poly and showed in [5] that BQP/qpoly is contained in the “heuristic” class HeurQMA/poly; Theorem 2 supersedes both of these earlier results.

Theorem 2 says that one can always replace polynomial-size quantum advice by polynomial-size classical advice, together with a polynomial-size untrusted quantum witness. Indeed, we can characterize the class BQP/qpoly as equal to the subclass of QMA/poly in which the quantum witness state |ψ_n⟩ can only depend on the input length n.¹

Using Theorem 2, we also obtain several other results for quantum complexity theory:

1. Without loss of generality, every quantum advice state can be taken to be the ground state of some local Hamiltonian H. In essence, this result follows by combining our BQP/qpoly ⊆ QMA/poly result with the result Kitaev (see [26]) that Local Hamiltonians is QMA-complete. The proof, however, requires a close analysis of the structure of low-energy states of the Hamiltonian H in Kitaev’s 5-local reduction (not proved or needed in [26]). To show that the locality of H can be reduced to 2, we use gadgets and a perturbation-theoretic result of Oliveira and Terhal [29], which built on Kempe, Kitaev, and Regev’s original proof of the QMA-completeness of 2-local Hamiltonians [25].²

2. It is open whether for every local Hamiltonian H on n qubits, there exists a quantum circuit of size poly(n) that prepares a ground state of H. It is easy to show that an affirmative answer would imply QMA = QCMA. As a consequence of Theorem 2, we can show that an affirmative answer would also imply BQP/qpoly = BQP/poly—thereby establishing a previously unknown connection between quantum proofs and quantum advice.

3. We generalize Theorem 2 to show that QCMA/qpoly ⊆ QMA/poly.

4. We use our new characterization of BQP/qpoly to prove a quantum analogue of the Karp–Lipton theorem [24]. Recall that the Karp–Lipton theorem says that if NP ⊆ P/poly, then the polynomial hierarchy collapses to the second level. Our “quantum Karp–Lipton theorem” says that if NP ⊆ BQP/qpoly

---

¹We call this restricted class YQP/poly. Its definition is closely related to the earlier notion of input-oblivious nondeterminism; this concept was used to define several other complexity classes in works of Chakravarthy and Roy [16] and Fortnow, Santhanam, and Williams [17]. We have made a significant alteration to the definition of YQP/poly from prior versions of this work, as discussed in section 1.3.

²Related results appear in [22], although these seem not to give what we need.
(that is, NP-complete problems are efficiently solvable with the help of quantum advice), then \( \Pi^P_2 \subseteq \text{QMA}^{\text{PromiseQMA}} \). As far as we know, this is the first nontrivial result to derive unlikely consequences from a hypothesis about quantum machines being able to solve NP-complete problems in polynomial time.

Finally, using our result, we are able to provide an illuminating perspective on a 2000 paper of Watrous [32]. Watrous gave a simple example of a purely classical problem in QMA that is not obviously in QCMA, that is, for which quantum proofs actually seem to help.\(^3\) This problem is called Group Non-Membership and is defined as follows: Arthur is given a finite black-box group \( G \) and a subgroup \( H \leq G \) (specified by their generators), as well as an element \( x \in G \). His task is to verify that \( x \not\in H \). It is known that, as a black-box problem, this problem is not in MA. But Watrous showed that Group Non-Membership is in QMA, by a protocol in which Merlin is “expected” to send the following quantum proof:

\[
|H\rangle = \frac{1}{\sqrt{|H|}} \sum_{h \in H} |h\rangle.
\]

Arthur’s verification procedure consists of two tests. In the first test, Arthur assumes that Merlin sent \(|H\rangle\) and then uses \(|H\rangle\) to decide whether \( x \in H \). The test is a simple, beautiful illustration of the power of quantum algorithms. The second test in Watrous’s protocol confirms that Merlin really sent \(|H\rangle\), or at least a state which is “equivalent” for purposes of the first test. This second test and its analysis are considerably more involved and seem less “natural.”

Using our results, we see that a slightly weaker version of Watrous’s result can be derived in an almost automatic way from his first test, as follows. If we assume that the black-box group \( H = H_n \) is fixed for each input length, then Group Non-Membership is in \( \text{BQP}/\text{qpoly} \), by letting \(|H_n\rangle\) as above be the trusted advice for length \( n \) and using Watrous’s first test as the \( \text{BQP}/\text{qpoly} \) algorithm. Then Theorem 2 (which can be readily adapted to the black-box setting) tells us that Group Non-Membership is in \( \text{QMA}/\text{poly} \) as well.

1.3. Changes to the paper. We have corrected some significant issues with previous drafts. First, the definition of so-called YQP machines needed to be amended to correct a deficiency in the previous definition that prevented completeness- and soundness-amplification techniques from working as claimed. This change appears necessary to preserve the claim \( \text{BQP}/\text{qpoly} = \text{YQP}/\text{poly} \). The revised definition of YQP/poly is actually more natural and has the same intuitive interpretation: now as before, a YQP/poly machine receives trusted classical advice plus untrusted quantum advice, each determined solely by the input length, and applies two computations—a first which \( \text{tests} \) the quantum advice \( \rho \) by some measurement process, and a second which \( \text{uses} \) \( \rho \) to compute to decide membership of an input \( x \) in some language \( L \).

The necessary change is that, rather than testing one copy of \( \rho \) and separately using another copy for the computation (an unnatural scenario, due to the no-cloning theorem of quantum mechanics), a YQP/poly algorithm first tests \( \rho \) then uses the \( \text{modified} \) post-measurement state \( \rho' \) for computing \( L(x) \). The revised correctness requirement is that, for any quantum advice \( \rho \) which has a noticeable chance of passing

\[^3\text{Aaronson and Kuperberg [6], however, give evidence that this problem might be in QCMA, under conjectures related to the classification of finite simple groups.}\]
the test, the post-test state $\rho'$ is useful for computation, \textit{conditioned} on passing the test.

The second significant issue we have addressed (pointed out to us by a journal referee) is that the analysis of local-Hamiltonian reductions for QMA in [26, 25] does not immediately supply enough information about the structure of ground states to prove Theorem 1. In particular, ground states of the Hamiltonians produced need not be “history states” encoding QMA verifier computations in the intended format, as we had erroneously claimed.

In the present version, we instead establish some properties of existing local-Hamiltonian reductions that suffice for our original application. First, we show that when Kitaev’s reduction [25] is applied to a 5-local Hamiltonian $V$ and can be used to efficiently obtain a proof state accepted with high probability by $V$. Next, we show that the reductions of Oliveira and Terhal [29], which can be used to transform a 5-local Hamiltonian $H^{(5)}$ into a 2-local $H^{(2)}$, are such that from any nearly minimal energy state for $H^{(2)}$ we can obtain a nearly minimal energy state for $H^{(5)}$. While this property is not immediate from past work, it can be obtained by applying a powerful theorem in [29] (building on [25]) which describes the behavior of $H^{(2)}$ on its low-energy subspaces.

1.4. Proof overview. We now give an overview of the proof of Theorem 2, that BQP/qpoly $\subseteq$ QMA/poly. As we will explain, our proof rests on a new idea we call the “majority-certificates” technique, which is not specifically quantum and which seems likely to find other applications.

We begin with a language $L \in$ BQP/qpoly and, for $n > 0$, a poly($n$)-size quantum circuit $Q(x, \xi)$ that computes $L(x)$ with high probability when given the “correct” advice state $\xi = \rho_n$ on poly($n$) qubits. The challenge, then, is to force Merlin to supply a witness state $\rho'$ that behaves like $\rho_n$ on every input $x \in \{0, 1\}^n$.

Every potential advice state $\xi$ defines a function $f_\xi : \{0, 1\}^n \rightarrow \{0, 1\}$, by $f_\xi(x) := \Pr[Q(x, \xi) = 1]$. For each such $\xi$, let $\hat{f}_\xi(x) := [f_\xi(x) \geq 1/2]$ be the Boolean function obtained by rounding $f_\xi$. As a simplification, suppose that Merlin is restricted to sending an advice state $\xi$ for which $f_\xi(x) \notin (1/3, 2/3)$, that is, an advice state which renders a “clear opinion” about every input $x$. (This simplification helps to explain the main ideas but does not follow the actual proof.) Let $S$ be the set of all Boolean functions $f : \{0, 1\}^n \rightarrow \{0, 1\}$ that are expressible as $\hat{f}_\xi$ for some such advice state $\xi$. Then $S$ includes the “target function” $f^* := L_n$ (the restriction of $L$ to inputs of length $n$), as well as a potentially large number of other functions. However, we claim that $S$ is not too large: $|S| \leq 2^{\text{poly}(n)}$. This bound on the “effective information content” of quantum states was derived previously by Aaronson [2, 5], building on the work of Ambainis et al. [12].

One might initially hope that, just by virtue of the size bound on $S$, we could find some set of poly($n$) values

\[(x_1, f^*(x_1)), \ldots, (x_k, f^*(x_1))\]

which \textit{isolate} $f^*$ in $S$, that is, which differentiate $f^*$ from all other members of $S$. In that case, the trusted classical advice could simply specify those values, as “tests” for
Arthur to perform on the quantum state sent by Merlin. Alas, this hope is unfounded in general. Consider the case where $f^*$ is the identically zero function and $S$ consists of $f^*$ along with the “point function” $f_y$ (which equals 1 on $y$ and 0 elsewhere) for all $y \in \{0,1\}^n$. Then $f^*$ can only be isolated in $S$ by specifying its value at every point!

Luckily, this counterexample leads us to a key observation. Although $f^*$ is not isolatable in $S$ by a small number of values, each point function $f_y$ can be isolated (by its value at $y$), and moreover, $f_y$ is quite “close” to $f^*$. In fact, if we choose any three distinct strings $x, y, z$, then $f^* \equiv \text{MAJ}(f_x, f_y, f_z)$. (Of course if $f^*$ were the identically zero function, it could be easily specified with classical advice! But $f^*$ could have been any function in this example.)

This suggests a new, more indirect approach to our general problem: we try to express $f$ as the pointwise majority vote

$$f^*(x) \equiv \text{MAJ}(f_1(x), \ldots, f_m(x))$$

of a small number ($m = O(n)$) of other functions $f_1, \ldots, f_m$ in $S$, where each $f_i$ is isolatable in $S$ by specifying at most $k = O(\log |S|)$ of its values. Indeed, we will show this can always be done. We call this key result the majority-certificates lemma; we will say more about its proof and its relation to earlier work in section 1.5.

With this lemma in hand, we can solve our (artificially simplified) problem: in the QMA/poly protocol for $L$, we use certificates which isolate $f_1, \ldots, f_m \in S$ as above as the classical advice for Arthur. Arthur requests from Merlin each of the $m$ states $\xi_1, \ldots, \xi_m$ such that $f_i = f_\xi_i$, and verifies that he receives appropriate states by checking them against the certificates. This involves multiple measurements of each $\xi_i$—and an immediate difficulty is that, since measurements are irreversible in quantum mechanics, the process of verifying the witness state might also destroy it.

We get around this difficulty by a somewhat more complicated protocol asking for multiple copies of each state $\xi_i$. Our analysis builds on ideas of Aharonov and Regev [9] used to prove the complexity-class equality $\text{QMA} = \text{QMA}^\perp$; informally, this result says that protocols in which Arthur is granted the (physically unrealistic) ability to perform “nondestructive measurements” on his witness state can be efficiently simulated by ordinary QMA protocols.

To build intuition, we will begin (in section 2) by proving the majority-certificates lemma for Boolean functions, as described above. However, to remove the artificial simplification we made and prove Theorem 2, we will need to generalize the lemma substantially, to a statement about possibly-infinite sets of real-valued functions $f : \{0, 1\}^n \to [0, 1]$. In the general version, the hypothesis that $S$ is finite and not too large will be replaced by a more subtle assumption, namely, an upper bound on the so-called fat-shattering dimension of $S$. To prove our generalization, we use powerful results of Alon et al. [10] and Bartlett and Long [13] on the learnability of real-valued functions. We then use a bound on the fat-shattering dimension of real-valued functions defined by quantum states (from Aaronson [5], building on Ambainis et al. [12]). Figure 1 shows the overall dependency structure of the proof.

1.5. Majority-certificates lemma in context. The majority-certificates lemma is closely related to the seminal notion of boosting [30] from computational learning theory. Boosting is a broad topic with a vast literature, but a common “generic” form of the boosting problem is as follows: we want to learn some target function $f^*$, given sample data of the form $(x, f^*(x))$. We assume we have a weak learning algorithm $A^{f^*,D}$ with the property that, for any probability distribution
Fig. 1. Dependency structure of our proof that quantum advice states can be expressed as ground states of local Hamiltonians.

\(D\) over inputs \(x,\) with high probability \(A\) finds a hypothesis \(f \in \mathcal{F}\) which predicts \(f^*(x)\) “reasonably well” when \(x \sim D\). The task is to boost this weak learner into a strong learner \(B^{f^*}\). The strong learner should output a collection of functions \(f_1, \ldots, f_m \in \mathcal{F}\), such that a (possibly weighted) majority vote over \(f_1(x), \ldots, f_m(x)\) predicts \(f^*(x)\) “extremely well.” It turns out [30, 18] that this goal can be achieved in a very general setting.

Our majority-certificates lemma has strengths and weaknesses compared to boosting. Our assumptions are much milder than those of boosting: rather than needing only that the hypothesis class \(\mathcal{S}\) is “not too large.” Also, we represent our target function \(f^*\) exactly by \(\text{MAJ}(f_1, \ldots, f_m)\), not just approximately. On the other hand, we do not give an efficient algorithm to find our majority-representation. Also, the \(f_i\)’s are not “explicitly given”: we only give a way to recognize each \(f_i\), under the assumption that the function purporting to be \(f_i\) is in fact drawn from the original hypothesis class.

The proof of our lemma also has similarities to boosting. As an analogue of a weak learner, we show that for every distribution \(D\), there exists a function \(f \in \mathcal{S}\) which agrees with the target function \(f^*\) on most \(x \sim D\), and which is isolatable in \(\mathcal{S}\) by specifying \(O(\log |\mathcal{S}|)\) queries. Using the minimax theorem, we then nonconstructively boost this fact into the desired majority-representation of \(f^*\). We note that Nisan used the minimax theorem for boosting in a similar way, in his alternative proof of Impagliazzo’s “hard-core set theorem” (see [21]).

The majority-certificates lemma is also reminiscent of Bshouty et al.’s algorithm [14] for learning small circuits in the complexity class \(\text{ZPP}^{\text{NP}}\). Our lemma lacks the algorithmic component of this earlier work, but unlike Bshouty et al., we do not
1.6. Organization of the paper. In section 2, we prove the Boolean majority-certificates lemma. In section 3, we give our real-valued generalization of this lemma, and in section 4 we use it to prove Theorem 2, and state some consequences for quantum complexity classes. Theorem 1 is proved in sections 5 through 7. Section 8 contains some further applications to quantum complexity theory.

2. The majority-certificates lemma. A Boolean concept class is a family of sets \( \{S_n\}_{n \geq 1} \), where each \( S_n \) consists of Boolean functions \( f : \{0, 1\}^n \rightarrow \{0, 1\} \) on \( n \) variables. Abusing notation, we will often use \( S \) to refer directly to a set of Boolean functions on \( n \) variables, with the quantification over \( n \) being understood.

By a certificate, we mean a partial Boolean function \( C : \{0, 1\}^n \rightarrow \{0, 1, *\} \). The size of \( C \), denoted \( |C| \), is the number of inputs \( x \) such that \( C(x) \in \{0, 1\} \). A Boolean function \( f : \{0, 1\}^n \rightarrow \{0, 1\} \) is consistent with \( C \) if \( f(x) = C(x) \) whenever \( C(x) \in \{0, 1\} \). Given a set \( S \) of Boolean functions and a certificate \( C \), let \( S[C] \) be the set of all functions \( f \in S \) that are consistent with \( C \). Say that a function \( f \in S \) is isolated in \( S \) by the certificate \( C \) if \( S[C] = \{f\} \).

We now prove a lemma that represents one of the main tools of this paper (although it will be generalized, rather than used directly).

**Lemma 3** (majority-certificates lemma). Let \( S \) be a set of Boolean functions \( f : \{0, 1\}^n \rightarrow \{0, 1\} \), and let \( f^* \in S \). Then there exist \( m = O(n) \) certificates \( C_1, \ldots, C_m \), each of size \( k = O(\log |S|) \), and functions \( f_1, \ldots, f_m \in S \), such that

\[
\begin{align*}
(1) \quad S[C_i] &= \{f_i\} \text{ all } i \in [m]; \\
(2) \quad \text{MAJ}(f_1(x), \ldots, f_m(x)) &= f^*(x) \text{ for all } x \in \{0, 1\}^n.
\end{align*}
\]

**Proof.** Our proof of Lemma 3 relies on the following claim.

**Claim 4.** Let \( D \) be any distribution over inputs \( x \in \{0, 1\}^n \). Then there exists a function \( f \in S \) such that

\[
\begin{align*}
(1) \quad f \text{ is isolatable in } S \text{ by a certificate } C \text{ of size } k = O(\log |S|); \\
(2) \quad \Pr_{x \sim D}[f(x) \neq f^*(x)] \leq \frac{1}{16}.
\end{align*}
\]

Lemma 3 follows from Claim 4 by a boosting-type argument, as follows. Consider a two-player game where

- Alice chooses a certificate \( C \) of size \( k \) that isolates some \( f \in S \) and
- Bob simultaneously chooses an input \( x \in \{0, 1\}^n \).

Alice wins the game if \( f(x) = f^*(x) \). Claim 4 tells us that for every mixed strategy of Bob (i.e., distribution \( D \) over inputs), there exists a pure strategy of Alice that succeeds with probability at least 0.9 against \( D \). Then by the minimax theorem, there exists a mixed strategy for Alice—that is, a probability distribution \( C \) over certificates—that allows her to win with probability at least 0.9 against every pure strategy of Bob.

Now suppose we draw \( C_1, \ldots, C_m \) independently from \( C \), isolating the functions \( f_1, \ldots, f_m \) in \( S \). Fix an input \( x \in \{0, 1\}^n \); then by the success of Alice’s strategy against \( x \), and applying a Chernoff bound,

\[
\Pr_{C_1, \ldots, C_m \sim C}[\text{MAJ}(f_1(x), \ldots, f_m(x)) \neq f^*(x)] < \frac{1}{2^n},
\]

provided we choose \( m = O(n) \) suitably. But by the union bound, this means there must be a fixed choice of \( C_1, \ldots, C_m \) such that \( \text{MAJ}(f_1, \ldots, f_m) \equiv f^* \), where each \( f_i \) is isolated in \( S \) by \( C_i \). This proves Lemma 3, modulo the claim. \( \square \)
Proof of Claim 4. By symmetry, we can assume without loss of generality that \( f^* \) is the identically zero function. Given the mixed strategy \( \mathcal{D} \) of Bob, we construct the certificate \( C \) as follows. Initially \( C \) is empty, that is, \( C(x) = * \) for all \( x \in \{0, 1\}^n \). In the first stage, we draw \( t = O(\log |S|) \) inputs \( x_1, \ldots, x_t \) independently from \( \mathcal{D} \). For any \( f : \{0, 1\}^n \to \{0, 1\} \), let

\[
    w_f := \Pr_{x \sim \mathcal{D}}[f(x) = 1].
\]

Now suppose \( f \) is such that \( w_f > 0.1 \). Then

\[
    \Pr_{x_1, \ldots, x_t \sim \mathcal{D}}[f(x_1) = 0 \land \cdots \land f(x_t) = 0] < 0.9^t \leq \frac{1}{|S|},
\]

provided \( t \geq \log_{10/9} |S| \). So by the union bound, there must be a fixed choice of \( x_1, \ldots, x_t \) that kills off every \( f \in S \) such that \( w_f > 0.1 \), that is, such that \( f(x_1) = \cdots = f(x_t) = 0 \) implies \( w_f \leq 0.1 \). Fix that \( x_1, \ldots, x_t \), and set \( C(x_i) := 0 \) for all \( i \in [t] \).

In the second stage, our goal is just to isolate some particular function \( f \in S[C] \). We do this recursively as follows. If \( |S[C]| = 1 \), then we are done. Otherwise, there exists an input \( x \) such that \( f(x) \neq f'(x) \) for some pair \( f, f' \in S[C] \). If setting \( C(x) := 0 \) decreases \( |S[C]| \) by at least a factor of 2, then set \( C(x) := 0 \); otherwise set \( C(x) := 1 \). Since \( S[C] \) can halve in size at most \( \log_2 |S| \) times, this procedure terminates after at most \( \log_2 |S| \) steps with \( |S[C]| = 1 \).

The end result is a certificate \( C \) of size \( O(\log |S|) \), which isolates a function \( f \) in \( S \) for which \( w_f \leq 1/10 \). We have therefore found a pure strategy for Alice that fails with probability at most 1/10 against \( \mathcal{D} \), as desired. \( \square \)

3. Extension to real functions. In this section, we extend the majority-certificates lemma from Boolean functions to real-valued functions \( f : \{0, 1\}^n \to [0, 1] \). We will need this extension for the application to quantum advice in section 4. In proving our extension, we will have to confront several new difficulties. First, the concept classes \( S \) that we want to consider can now contain a continuum of functions—so Lemma 3, which assumed that \( S \) was finite and constructed certificates of size \( O(\log |S|) \), is not going to work. In section 3.1, we review notions from computational learning theory, including fat-shattering dimension and \( \varepsilon \)-covers, which (combined with results of Alon et al. [10] and Bartlett and Long [13]) can be used to get around this difficulty. Second, it is no longer enough to isolate a function \( f_i \in S \) that we are interested in; instead we will need to “safely” isolate \( f_i \), which roughly speaking means that (i) \( f_i \) is consistent with some certificate \( C \) and (ii) any \( f \in S \) that is even approximately consistent with \( C \) is close to \( f_i \). In section 3.2, we prove a “safe winnowing lemma” that can be used for this purpose, and we put our ingredients together to prove a real-valued majority-certificates lemma.

3.1. Background from learning theory. A \( p \)-concept class \( S \) over \( \{0, 1\}^n \) is a family of functions \( f : \{0, 1\}^n \to [0, 1] \) (as usual, quantification over all \( n \) is understood). Given functions \( f, g : \{0, 1\}^n \to [0, 1] \) and a subset of inputs \( X \subseteq \{0, 1\}^n \), we will be interested in two measures of the distance between \( f \) and \( g \) restricted to \( X \):

\[
    \Delta_{\infty}(f, g)[X] := \max_{x \in X} |f(x) - g(x)|,
\]

\[
    \Delta_1(f, g)[X] := \sum_{x \in X} |f(x) - g(x)|.
\]
For convenience, we define $\Delta_\infty(f, g) := \Delta_\infty(f, g)[\{0, 1\}^n]$, and similarly for $\Delta_1(f, g)$. Also, given a distribution $D$ over $\{0, 1\}^n$, define
\[
\Delta_1(f, g)(D) := E_{x \sim D} [|f(x) - g(x)|].
\]
Finally, we will need the notions of $\varepsilon$-covers and fat-shattering dimension.

**Definition 5** ($\varepsilon$-covers). Let $S$ be a $p$-concept class over $\{0, 1\}^n$. The subset $C \subseteq S$ is an $\varepsilon$-cover for $S$ if for all $f \in S$, there exists a $g \in C$ such that $\Delta_\infty(f, g) \leq \varepsilon$.

**Definition 6** (fat-shattering dimension). Let $S$ be a $p$-concept class over $\{0, 1\}^n$ and $\varepsilon > 0$ be a real number. We say the set $A \subseteq \{0, 1\}^n$ is $\varepsilon$-shattered by $S$ if there exists a function $r : A \rightarrow [0, 1]$ such that for all $2^{|A|}$ Boolean functions $g : A \rightarrow \{0, 1\}$, there exists a $p$-concept $f \in S$ such that for all $x \in A$, we have $f(x) \leq r(x) - \varepsilon$ whenever $g(x) = 0$ and $f(x) \geq r(x) + \varepsilon$ whenever $g(x) = 1$. Then the $\varepsilon$-fat-shattering dimension of $S$, denoted $\text{fat}_\varepsilon(S)$, is the size of the largest set $\varepsilon$-shattered by $S$.

The $p$-concept classes we consider in this paper will be convex when considered as subsets of $[0, 1]^n$. We remark that for such classes, $\text{fat}_\varepsilon(S)$ measures the largest dimension of any axis-parallel subcube contained in $S$ of side length $2\varepsilon$.

The following central result was shown by Alon et al. [10] (see also [23]).

**Theorem 7** (see [10]). Every $p$-concept class $S$ has an $\varepsilon$-cover $C$ of size $|C| \leq \exp(O(n + \log 1/\varepsilon)\text{fat}_{1/4}(S))$.

Building on the work of Alon et al. [10], Bartlett and Long [13] then proved the following.

**Theorem 8** (see [13]). Let $S$ be a $p$-concept class and $D$ be a distribution over $\{0, 1\}^n$. Fix an $f : \{0, 1\}^n \rightarrow [0, 1]$ (not necessarily in $S$) and an error parameter $\alpha > 0$. Suppose we form a set $X \subseteq \{0, 1\}^n$ by choosing $m$ inputs independently with replacement from $D$. Then there exists a positive constant $K$ such that, with probability at least $1 - \delta$ over $X$, any hypothesis $h \in S$ that minimizes $\Delta_1(h, f)[X]$ also satisfies
\[
\Delta_1(h, f)(D) \leq \alpha + \inf_{g \in S} \Delta_1(g, f)(D),
\]
provided that
\[
m \geq \frac{K}{\alpha^2} \left( \text{fat}_{\alpha/5}(S) \log^2 \frac{1}{\alpha} + \log \frac{1}{\delta} \right).
\]

Theorem 8 has the following corollary, which is similar to Corollary 2.4 of Aaronson [5], but more directly suited to our purposes here.

**Corollary 9**. Let $S$ be a $p$-concept class over $\{0, 1\}^n$ and $D$ be a distribution over $\{0, 1\}^n$. Fix an $f \in S$ and an error parameter $\varepsilon > 0$. Suppose we form a set $X \subseteq \{0, 1\}^n$ by choosing $m$ inputs independently with replacement from $D$. Then there exists a positive constant $K$ such that, with probability at least $1 - \delta$ over $X$, any hypothesis $h \in S$ that satisfies $\Delta_\infty(h, f)[X] \leq \varepsilon$ also satisfies $\Delta_1(h, f)(D) \leq 11\varepsilon$, provided
\[
m \geq \frac{K}{\varepsilon^2} \left( \text{fat}_{\varepsilon}(S) \log \frac{1}{\varepsilon} + \log \frac{1}{\delta} \right).
\]

---

5It would also be possible to apply the bound from [5] “off-the-shelf,” but at the cost of a worse dependence on $1/\varepsilon$. Theorem 8 has the following corollary, which is similar to Corollary 2.4 of Aaronson [5], but more directly suited to our purposes here. 

**Corollary 9**. Let $S$ be a $p$-concept class over $\{0, 1\}^n$ and $D$ be a distribution over $\{0, 1\}^n$. Fix an $f \in S$ and an error parameter $\varepsilon > 0$. Suppose we form a set $X \subseteq \{0, 1\}^n$ by choosing $m$ inputs independently with replacement from $D$. Then there exists a positive constant $K$ such that, with probability at least $1 - \delta$ over $X$, any hypothesis $h \in S$ that satisfies $\Delta_\infty(h, f)[X] \leq \varepsilon$ also satisfies $\Delta_1(h, f)(D) \leq 11\varepsilon$, provided
\[
m \geq \frac{K}{\varepsilon^2} \left( \text{fat}_{\varepsilon}(S) \log \frac{1}{\varepsilon} + \log \frac{1}{\delta} \right).
\]
Proof. Let \( S^* \) be the \( p \)-concept class consisting of all functions \( g : \{0, 1\}^n \to [0, 1] \) for which there exists an \( f \in S \) such that \( \Delta_\infty(g, f) \leq \varepsilon \). Fix an \( f \in S \) and a distribution \( D \), and let \( X \) be chosen as in the statement of the corollary. Suppose we choose a hypothesis \( h \in S \) such that \( \Delta_\infty(h, f)[X] \leq \varepsilon \). Define a function \( g \) by setting \( g(x) := h(x) \) if \( x \in X \) and \( g(x) := f(x) \) otherwise. Note that \( \Delta_\infty(g, f) \leq \varepsilon \) and that \( g \in S^* \). Also note that \( \Delta_1(h, g)[X] = 0 \), which means that \( h \) minimizes the functional \( \Delta_1(h, g)[X] \) over all hypotheses in \( S \) (and indeed in \( S^* \)). By Theorem 8, this implies that with probability at least \( 1 - \delta \) over \( X \),

\[
\Delta_1(h, g)(D) \leq \alpha + \inf_{u \in S^*} \Delta_1(u, g)(D) = \alpha
\]

for all \( \alpha > 0 \), provided we take

\[
m \geq K \frac{K}{\alpha^2} \left( \text{fat}_{\alpha/5}(S^*) \log^2 \frac{1}{\alpha} + \log \frac{1}{\delta} \right).
\]

Here we have used the fact that \( g \in S^* \), and hence

\[
\inf_{u \in S^*} \Delta_1(u, g)(D) = 0.
\]

So by the triangle inequality,

\[
\Delta_1(h, f)(D) \leq \Delta_1(h, g)(D) + \Delta_1(g, f)(D)
\]

\[
\leq \alpha + \Delta_\infty(g, f)
\]

\[
\leq \alpha + \varepsilon.
\]

Next, we claim that \( \text{fat}_{\alpha/5}(S^*) \leq \text{fat}_{\alpha/5-\varepsilon}(S) \). The reason is simply that, if a given set is \( \beta \)-fat-shattered by \( S^* \), then it must also be \((\beta - \varepsilon)\)-fat-shattered by \( S \), by the triangle inequality. Setting \( \alpha := 10\varepsilon \) now yields the desired statement. \( \square \)

3.2. The “safe winnowing lemma” and the real-valued majority-certificates lemma. A key technical step toward proving the real-valued majority-certificates lemma is our so-called safe winnowing lemma. This lemma says intuitively that, given any set \( S \) of real-valued functions with a small \( \varepsilon \)-cover (or equivalently, with polynomially-bounded fat-shattering dimension), and given any \( f^* \in S \) and subset \( Y \subseteq \{0, 1\}^n \) of inputs to \( f^* \), it is possible to find a set of \( k = \text{poly}(n) \) constraints \(|f(x_1) - a_1| \leq \varepsilon, \ldots, |f(x_k) - a_k| \leq \varepsilon \), and another function \( f \in S \), such that \( f \) is close to \( f^* \) in \( L_\infty \) norm on \( Y \), and \( f \) is essentially the only function in \( S \) compatible with the constraints. Here “essentially” means that (i) any function that satisfies the constraints is close to \( f^* \) in \( L_\infty \)-norm, and (ii) \( f^* \) itself not only satisfies the constraints, but does so with a “margin to spare.”

Lemma 10 (safe winnowing lemma). Let \( S \) be a \( p \)-concept class over \( \{0, 1\}^n \). Fix a function \( f^* \in S \) and subset \( Y \subseteq \{0, 1\}^n \). For some parameter \( \varepsilon > 0 \), let \( C \) be a finite \( \varepsilon \)-cover for \( S \). Then there exists an \( f \in S \), as well as a subset \( Z \subseteq \{0, 1\}^n \) of size at most \( k = \log_2 \frac{1}{\varepsilon} |C| \), such that the following hold:

(i) Every \( g \in S \) that satisfies \( \Delta_\infty(f, g)(Y \cup Z) \leq \frac{\varepsilon}{10} \) also satisfies \( \Delta_\infty(f, g) \leq 3\varepsilon \).

(ii) \( \Delta_\infty(f, f^*)([Y]) \leq \varepsilon / 5 \).

We defer the proof of Lemma 10, showing first how it helps us to prove our generalization of Lemma 3 to the case of real-valued functions:

Lemma 11 (real majority-certificates). Let \( S \) be a \( p \)-concept class over \( \{0, 1\}^n \), let \( f^* \in S \), and let \( \varepsilon > 0 \). Then for some \( m = O(n/\varepsilon^2) \), there exist functions
Let \( f_1, \ldots, f_m \in S \), sets \( X_1, \ldots, X_m \subseteq \{0,1\}^n \) all of some equal size \( |X_i| = k = O((n + \log^{2/3} n) \text{fat}_{48}(S)) \), and an \( \alpha = \Omega((n + \log^{1/3} n) \text{fat}_{48}(S)) \) for which the following holds. All \( g_1, \ldots, g_m \in S \) that satisfy \( \Delta_\infty(f_i, g_i) |X_i| \leq \alpha \) for \( i \in [m] \) also satisfy \( \Delta_\infty(f^*, g) \leq \varepsilon \), where

\[
g(x) := \frac{g_1(x) + \cdots + g_m(x)}{m}.
\]

**Proof.** Let

\[
\beta := \frac{\varepsilon}{48}, \\
t := C \left(n + \log \frac{1}{\beta}\right) \text{fat}_\beta(S), \\
\alpha := \frac{0.4\beta}{t},
\]

where \( C \) is a suitably large constant. Also, let \( S_{\text{fin}} \) be a finite \( \alpha \)-cover for \( S \), that is, a finite subset \( S_{\text{fin}} \subseteq S \) such that for all \( f \in S \), there exists a \( g \in S_{\text{fin}} \) such that \( \Delta_\infty(f, g) \leq \alpha \).

Given \( f \) and \( X \), let \( S[f, X] \) be the set of all \( g \in S \) such that \( \Delta_\infty(f, g) |X| \leq \alpha \).

Now consider a two-player game where Alice chooses a function \( f \in S_{\text{fin}} \) and a set \( X \subseteq \{0,1\}^n \) of size \( k \), and Bob simultaneously chooses an input \( x \in \{0,1\}^n \). Alice’s penalty in this game (the number she is trying to minimize) equals

\[
\sup_{g \in S[f, X]} |f^*(x) - g(x)|.
\]

We claim that there exists a mixed strategy for Alice—that is, a probability distribution \( P \) over \((f, X)\) pairs—that gives her an expected penalty of at most \( \varepsilon/2 \) against every pure strategy of Bob.

Let us see why Lemma 11 follows from this claim. Fix an input \( x \in \{0,1\}^n \) and suppose Alice draws \((f_1, X_1), \ldots, (f_m, X_m)\) independently from \( P \). Then for all \( i \in [m] \),

\[
\mathbb{E}_{(f_i, X_i)} \left[ \sup_{g \in S[f, X]} \left| f^*(x) - g(x) \right| \right] \leq \frac{\varepsilon}{2}.
\]

Thus, letting \( z_1, \ldots, z_m \) be independent random variables in \([0,1]\), each with expectation at most \( \varepsilon/2 \), the expression

\[
\Pr_{(f_i, X_i) \in [m]} \left[ \exists g_1 \in S[f_1, X_1], \ldots, g_m \in S[f_m, X_m] : \left| f^*(x) - \frac{g_1(x) + \cdots + g_m(x)}{m} \right| > \varepsilon \right]
\]

is at most \( \Pr[z_1 + \cdots + z_m > \varepsilon m] \) using the triangle inequality. This, in turn, is less than

\[
2 \exp \left( -\frac{2(\varepsilon m/2)^2}{m} \right) < 2^{-n}
\]

\(^6\)We will need \( S_{\text{fin}} \) for the technical reason that the basic minimax theorem only works with finite strategy spaces.
by Hoeffding’s inequality, provided we choose \( m = O(n/\varepsilon^2) \) suitably. By the union bound, this means that there must be a fixed choice of \( f_1, \ldots, f_m \) and \( X_1, \ldots, X_m \) such that

\[
\left| f^*(x) - \frac{g_1(x) + \cdots + g_m(x)}{m} \right| \leq \varepsilon
\]

for all \( g_i \in S[f_1, X_1], \ldots, g_m \in S[f_m, X_m] \) and all inputs \( x \in \{0,1\}^n \) simultaneously, as desired.

We now prove the claim. By the minimax theorem, our task is equivalent to the following: given any mixed strategy \( D \) of Bob, find a pure strategy of Alice that achieves a penalty of at most \( \varepsilon/2 \) against \( D \). In other words, given any distribution \( D \) over inputs \( x \in \{0,1\}^n \), we want a fixed function \( f \in S_{\text{fin}} \), and a set \( X \subseteq \{0,1\}^n \) of size \( k \), such that

\[
\mathbb{E}_{x \sim D} \sup_{g \in S[f,X]} |f^*(x) - g(x)| \leq \frac{\varepsilon}{2}.
\]

We construct this \((f,X)\) pair as follows. In the first stage, we let \( Y \) be a set, of size at most

\[
M := K \left( \text{fat}_\beta(S) \log \frac{1}{\beta} + \log \frac{1}{\delta} \right),
\]

formed by choosing \( M \) inputs independently with replacement from \( D \). Here \( \beta = \varepsilon/48 \) as defined earlier, \( \delta = 1/2 \), and \( K \) is the constant from Corollary 9. Then by Corollary 9, with probability at least \( 1 - \delta = 1/2 \) over the choice of \( Y \), any \( g \in S \) that satisfies \( \Delta_\infty(f^*,g)[Y] \leq \beta \) also satisfies \( \Delta_1(f^*,g)[D] \leq 11\beta \). So there must be a fixed choice of \( Y \) with that property. Fix that \( Y \), and let \( S' \) be the set of all \( g \in S \) such that \( \Delta_\infty(f^*,g)[Y] \leq \beta \).

In the second stage, our goal is just to use Lemma 10 to winnow \( S' \) down to a particular function \( f \). More precisely, we want to find an \( f \in S' \cap S_{\text{fin}} \), and a set \( X \subseteq \{0,1\}^n \) containing \( Y \), such that any \( g \in S \) that satisfies \( \Delta_\infty(f,g)[X] \leq \alpha \) also satisfies \( \Delta_\infty(f,g) \leq 11\beta \). We assert that such a pair \((f,X)\) can be found. It will then follow that

\[
\mathbb{E}_{x \sim D} \sup_{g \in S[f,X]} |f^*(x) - g(x)| \leq \Delta_1(f^*,f)[D] + \sup_{g \in S[f,X]} \Delta_\infty(f,g) \leq 11\beta + 13\beta = \frac{\varepsilon}{2},
\]

which proves that \((f,X)\) give a strategy for Alice having the needed quality against the mixed strategy \( D \) for Bob.

We find the desired \((f,X)\) pair as follows. By Theorem 7, the class \( S' \) has a \( 4\beta \)-cover of size

\[
N = \exp \left[ O \left( \left( n + \log \frac{1}{4\beta} \right) \text{fat}_\beta(S') \right) \right] \leq \exp \left[ O \left( \left( n + \log \frac{1}{\beta} \right) \text{fat}_\beta(S) \right) \right].
\]

Let \( t := \log_2 N \). Then by Lemma 10, there exists a function \( u \in S' \), as well as a subset \( Z \subseteq \{0,1\}^n \) of size at most \( t \), such that the following hold:
(i) \( \Delta_{\infty}(u, f^*)[Y] \leq 0.8\beta \).

(ii) Every \( g \in S' \) that satisfies \( \Delta_{\infty}(u, g)[Y \cup Z] \leq \frac{0.8\beta}{t} \) also satisfies \( \Delta_{\infty}(u, g) \leq 12\beta \).

Let \( X := Y \cup Z \), and observe that

\[
|X| = O \left( \frac{1}{\beta^2} \operatorname{fat}_{\beta} (S) \log^2 \frac{1}{\beta} + \left( n + \log \frac{1}{\beta} \right) \operatorname{fat}_{\beta} (S) \right)
= O \left( \left( n + \frac{\log^2 \left( \frac{1}{\varepsilon} \right)}{\varepsilon^2} \right) \operatorname{fat}_{\varepsilon/48} (S) \right)
\]

as desired. Now let \( f \) be a function in \( S_{\text{in}} \), such that \( \Delta_{\infty}(f, u) \leq \alpha \). Let us check that \( (f, X) \) have the properties we want. First,

\[
\Delta_{\infty}(f^*, f)[Y] \leq \Delta_{\infty}(f^*, u)[Y] + \Delta_{\infty}(u, f)[Y] \\
\leq 0.8\beta + \alpha \\
< 0.9\beta,
\]

and hence \( f \in S' \) as desired. Next, consider any \( g \in S \) that satisfies \( \Delta_{\infty}(f, g)[X] \leq \alpha \). Then we also have

\[
\Delta_{\infty}(f^*, g)[Y] \leq \Delta_{\infty}(f^*, f)[Y] + \Delta_{\infty}(f, g)[Y] \\
\leq 0.9\beta + \alpha \\
< \beta,
\]

and hence \( g \in S' \), so that (by our construction of \( Y \)) we have \( \Delta_1(f^*, g)(D) \leq 11\beta \).

Next, observe that

\[
\Delta_{\infty}(u, g)[X] \leq \Delta_{\infty}(u, f)[X] + \Delta_{\infty}(f, g)[X] \\
\leq 2\alpha \\
= \frac{0.8\beta}{t},
\]

so that, using our guarantee (ii) above, we have \( \Delta_{\infty}(u, g) \leq 12\beta \). Then we find that

\[
\Delta_{\infty}(f, g) \leq \Delta_{\infty}(f, u) + \Delta_{\infty}(u, g) \\
\leq \alpha + 12\beta \\
\leq 13\beta,
\]

as required. This shows that \( (f, X) \) have the required properties, and it completes the proof of Lemma 11.

Proof of Lemma 10. Let \( \delta := \frac{\varepsilon}{\varepsilon^2} \). We construct \( (f, Z) \) by an iterative procedure. Initially let \( S_0 := S \), let \( f_0 := f^* \), and let \( Z_0 := Y \). We will form new sets \( S_1, S_2, \ldots \) by repeatedly adding constraints of the form \( f(x) \leq \alpha \) or \( f(x) \geq \alpha \) for various \( x, \alpha \), maintaining the invariant that \( \beta \in S_i \). At iteration \( t \), suppose there exists a function \( g \in S_{t-1} \) such that \( \Delta_{\infty}(f_1, g)[Y \cup Z_{t-1}] \leq \delta \), but nevertheless \( |f_1(z_i) - g(z_i)| > 3\varepsilon \) for some input \( z_i \). Then first set \( Z_t := Z_{t-1} \cup \{ z_i \} \) (i.e., add \( z_i \) into our set of inputs if it is not already there). Let \( v := \frac{1}{2}|f_1(z_i) + g(z_i)| \), let \( A \) be the set of all functions \( h \in S_{t-1} \) such that \( h(z_i) < v \), and let \( B \) be the set of all \( h \in S_{t-1} \) such that \( h(z_i) \geq v \). Also, for any given set \( M \), let \( M^0 := M \cap C \). Then clearly \( \min(|A^0|, |B^0|) \leq |S_{t-1}^0|/2 \). If \( |A^0| < |B^0| \), then set \( S_t := A \); otherwise set \( S_t := B \).
Then set $f_1 := f_{t-1}$ if $f_{t-1} \in S_t$, and $f_t := g$ otherwise. Since $|S_T^g|$ can halve at most $k = \log_2 |C|$ times, it is clear that after $T \leq k$ iterations we have $|S_T^g| \leq 1$. Set $f := f_T$ and $Z := Z_T$. Then by the triangle inequality,

$$
\Delta_\infty (f, f^*) [Y] \leq T \delta \leq \varepsilon/5,
$$

and also

$$
|f(z_t) - f_t(z_t)| \leq (T - t) \delta < \varepsilon/5
$$

for all $t \in [T]$. So suppose by contradiction that there still exists a function $g \in S_T$ such that $\Delta_\infty (f, g)[Y \cup Z] \leq \delta$ but $|f(x) - g(x)| > 3\varepsilon$ for some $x$, and consider functions $p, q \in C$ in the cover such that $\Delta_\infty (f, p) \leq \varepsilon$ and $\Delta_\infty (g, q) \leq \varepsilon$. Then $p, g \in S_T^g$ but $p \neq g$, which contradicts the fact that $|S_T^g| \leq 1$. Also notice that for all $g \in S$, if $\Delta_\infty (f, g)[Y \cup Z] \leq \delta$, then $g \in S_T$. Thus $\Delta_\infty (f, g)[Y \cup Z] \leq \delta$ implies $\Delta_\infty (f, g) \leq 3\varepsilon$ as desired.

4. Application to quantum advice classes. In this section, we prove Theorem 2, as well as several other results. We will be defining quantum circuits over some fixed universal basis of 2-local unitary and measurement gates. We use size$(C)$ to denote the number of gates of a classical or quantum circuit (including the input and output gates).

4.1. Classical descriptions for quantum states. Fix a quantum circuit $Q$ taking an $n$-bit string $x$ and a $p$-qubit state $\rho$ and producing a 1-bit output. For a given state $\rho$, let $f_\rho(x) := \mathbb{E}[Q(x, \rho)]$. Let $S$ be the $p$-concept class consisting of $f_\rho$ for all $p$-qubit mixed states $\rho$. Then Aaronson [5] proved the following result, which allows us to apply the real-valued majority-certificates lemma to the study of quantum advice.

**Theorem 12** (see [5]). $\text{fat}_\gamma (S) = O(p/\gamma^2)$.

The next claim gives a useful consequence of Theorem 12 and the majority-certificates lemma.

**Lemma 13.** Let $Q_n(x, \rho)$ be a quantum circuit taking as input a string $x \in \{0, 1\}^n$ and a quantum state $\rho$ on $p$ qubits, and outputting a single bit. Fix any $p$-qubit state $\rho_n^*$. Let $c \geq 1$ be a constant. For suitably chosen integers $m, k \leq \text{poly}(n, p)$ and a real parameter $\alpha \geq 1/\text{poly}(n, p)$, there exists

- a second circuit $Q'_n(x, \sigma)$ of size at most $\text{poly} (\text{size}(Q_n))$ taking as input $x \in \{0, 1\}^n$ and an $m \cdot p$-qubit state $\sigma$;
- a collection $C_n = \{C_{(i,j)}(\sigma)\}_{(i,j) \in [m] \times [k]}$ of circuits, each of size $|C_{(i,j)}| \leq \text{poly} (\text{size}(Q_n))$, and each taking as input a quantum state $\sigma$ on $m \cdot p$ qubits; and finally,
- a collection $\{r_{(i,j)}\}_{(i,j) \in [m] \times [k]}$ of rational numbers in $[0, 1]$, each specified by a decimal expansion of length $O(\log (n + p))$.

(Here, $Q'_n$ can be uniformly constructed in time $\text{poly}(s, n)$ given a description of $Q_n$, while $C_n, \{r_{(i,j)}\}$ are nonuniformly chosen.) We have the following properties:

(i) There exists a state $\sigma$ on $m \cdot p$ qubits, of the form $\sigma = \sigma_1 \otimes \cdots \otimes \sigma_m$, that satisfies $|E[C_{(i,j)}(\sigma)] - r_{(i,j)}| \leq \alpha$ for each $(i, j) \in [m] \times [k]$.

(ii) If we are given any state $\sigma$ on $m \cdot p$ qubits, satisfying

$$
|E[C_{(i,j)}(\sigma)] - r_{(i,j)}| \leq 4\alpha \quad \forall (i, j) \in [m] \times [k],
$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
then it also holds that
\[ |\mathbb{E}[Q_n(x,\sigma)] - \mathbb{E}[Q_n(x,\rho_n^i)]| \leq n^{-c} \quad \forall x \in \{0,1\}^n. \]

Proof. For each \( x \in \{0,1\}^n \) and state \( \xi \) on \( p \) qubits, let \( f_\xi(x) := \mathbb{E}[Q_n(x,\xi)] \). Let \( S \) be collection \( \{f_\xi\} \), ranging over all \( p \)-qubit mixed states \( \xi \). Then Theorem 12 implies that \( \text{fat}_{\gamma}(S) = O(p/\gamma^2) \) for all \( \gamma > 0 \). Set \( \varepsilon := n^{-c}, \gamma := \varepsilon/48 \). By Lemma 11, for some \( m, k \leq \text{poly}(n) \), there exist \( p \)-qubit mixed states \( \rho_1, \ldots, \rho_m \), sets \( X_1, \ldots, X_m \subseteq \{0,1\}^n \) each of size \( k \), and an \( \alpha = \Omega\left(\frac{1}{\text{poly}(n,p)}\right) \) for which the following holds:

(*) All collections \( \sigma_1, \ldots, \sigma_m \) of \( p(n) \)-qubit states that satisfy \( \Delta_{\infty}(f_{\rho_1}, f_{\sigma_i})[X_i] \leq \alpha \) for \( i \in [m] \) also satisfy \( \Delta_{\infty}(f_{\rho_1}, f_{\sigma_{\text{avg}}}) \leq n^{-c} \), where \( \sigma_{\text{avg}} := \frac{1}{m}(\sigma_1 + \cdots + \sigma_m) \).

For an \( m \cdot p \)-qubit state \( \sigma \) and \( i \in [m] \), let \( \sigma[i] \) denote the reduced state of \( \sigma \) on the \( i \)th register of \( p \) qubits. Let \( x^{(i,j)} \in \{0,1\}^n \) denote the \( j \)th element in \( X_i \) (under some fixed ordering). The circuits \( \{C_{(i,j)}\}_{(i,j)\in[m] \times [k]} \) are then defined as follows: each \( C_{(i,j)} \), on input state \( \sigma \), simulates \( Q_n(x^{(i,j)}, \sigma[i]) \) (by applying \( Q_n(x^{(i,j)}, \cdot) \) to the \( i \)th register of \( \sigma \) and outputs the resulting bit. The value \( r_{i,j} \) is chosen as a rational approximation to the value \( \mathbb{E}[Q_n(x^{(i,j)}, \rho_i)] \), accurate to within \( \pm 10^{-c} \); this can be achieved with \( O(\log(n + p)) \) bits of precision, since \( \alpha \geq 1/\text{poly}(n, p) \). Finally, for the circuit \( Q'_n(x, \sigma) \), we let \( Q'_n \) choose a uniformly random register \( i \in [m] \) and simulate \( Q_n(x, \sigma[i]) \), outputting the result. All of our efficiency claims for \( Q'_n \) and \( \{C_{(i,j)}\}_{(i,j)\in[m] \times [k]} \), and our uniform constructibility claim for \( Q'_n \), follow from the definitions.

To establish item (i) in the Theorem’s conclusion, it is enough to verify that \( \sigma := \rho_1 \otimes \cdots \otimes \rho_m \) is a suitable choice of \( \sigma \), by our settings to \( \{C_{(i,j)}, r_{(i,j)}\} \). For item (ii), let the \( m \cdot p(n) \)-qubit state \( \sigma \) satisfy the hypothesis in that item. By our definitions and the quality of our rational approximations \( \{r_{(i,j)}\} \), this implies that \( \Delta_{\infty}(f_{\rho_1}, f_{\sigma[i]}[X_i] \leq \alpha \) for \( i \in [m] \). Then by (*), we have \( \Delta_{\infty}(f_{\rho_1}, f_{\sigma_{\text{avg}}}) \leq n^{-c} \), where we here define \( \sigma_{\text{avg}} := \frac{1}{m}(\sigma[1] + \cdots + \sigma[m]) \). Also, for our choice of \( Q'_n \) we have
\[
\mathbb{E}[Q'_n(x, \sigma)] = \frac{1}{m} \sum_{i\in[m]} \mathbb{E}[Q_n(x, \sigma[i])] = \mathbb{E}[Q_n(x, \sigma_{\text{avg}})] = f_{\sigma_{\text{avg}}}(x).
\]

This gives item (ii), completing the proof of Lemma 13. \( \square \)

### 4.2. Advice-testing quantum circuits and input-oblivious testers

Next we define a class of quantum circuits that will play an important role in our work.

**Definition 14.** An advice-testing circuit (for the input length \( n > 0 \)) is a quantum circuit \( Y = Y_n \) with a classical \( n \)-bit input register, along with advice and ancilla registers and two designated 1-qubit “advice-testing” and “output” registers. On input a string \( x \in \{0,1\}^n \), and with the advice register initialized to some advice state \( \rho \), the remaining registers are each initialized to the all-zero state. \( Y \) acts as follows:

1. First \( Y \) applies a subcircuit \( A \) to all registers, after which the advice-testing register is measured, producing a value \( b_{\text{adv}} \in \{0,1\} \).
2. Next, \( Y \) applies a second subcircuit \( B \) to all registers then measures the output register, producing a value \( b_{\text{out}} \in \{0,1\} \).

If in step 1 above, the subcircuit \( A \) ignores the input register, then \( Y \) is said to be an input-oblivious advice-testing circuit.
Next, suppose we have a quantum circuit $Q_n(x, \rho)$ taking a classical string $x \in \{0,1\}^n$ and a quantum state $\rho$ that we wish to simulate for a specific desired setting $\rho := \rho^*$. The next result gives a general method to do so by an input-oblivious advice-testing algorithm with polynomial classical advice. Our use of Lemma 13 in proving this result draws ideas from the proof of Aharonov and Regev of the equality of complexity classes QMA^+ = QMA [8].

**Theorem 15.** Let $Q_n(x, \rho)$ be a quantum circuit taking as input a string $x \in \{0,1\}^n$ and a quantum state $\rho$ on $p \leq s$ qubits, and outputting a single bit. Fix any $p$-qubit state $\rho^*$, and let $d \geq 1$ be a fixed constant.

Then there exists an input-oblivious advice-testing circuit $Y_n$, of size bounded by $|Y_n| \leq \text{poly(size}(Q_n))$, taking an input $x \in \{0,1\}^n$ and a $P$-qubit advice state (for some $P \leq \text{poly}(n,p)$), with the following properties:

(i) There exists an advice state $\sigma^*$ on $P$ qubits such that for all $x \in \{0,1\}^n$, in the execution of $Y_n(x, \sigma^*)$ we have $\Pr[b_{\text{adv}} = 1] \geq 1 - e^{-n}$.

(ii) For each $n$ and advice state $\sigma$ on $P$ qubits, it holds that in the execution of $Y_n(x, \sigma)$ (for each $x \in \{0,1\}^n$) we have

$$\Pr[b_{\text{adv}} = 1] \geq n^{-d} \implies |E[b_{\text{out}} | b_{\text{adv}} = 1] - E[Q_n(x, \rho^*)]| \leq n^{-d}.$$

**Proof of Theorem 15.** For $n > 1$, let

$$m, k, \alpha, Q_n', C_n, \{r_{(i,j)}\}_{i \in [m], j \in [k]}$$

be as given by Lemma 13 applied to $Q, \rho^*$, and with $c := 2d$. We set $M := \lceil 10^{n^c}mk/\alpha \rceil$, $N := \lceil 10\ln M/\alpha^2 \rceil$, and $P := MNmp$. We regard a $P$-qubit state as having $MN$ registers (indexed by $[M] \times [N]$) of $m \cdot p$ qubits each. We refer to the register indexed by $(s,t) \in [M] \times [N]$ as the “$(s,t)$th proof register.”

The subroutine $A$ for $Y_n$ is defined as follows.

**Algorithm A($\sigma, y$).**

1. Set $b_{\text{adv}} := 1$, and choose $S \in [M]$ uniformly;
2. For $s = 1, 2, \ldots, (S - 1)$:
   2.a. Choose $(i(s), j(s)) \in [m] \times [k]$ uniformly;
   2.b. Apply $C_{(i(s), j(s))}$ successively to the proof registers $(s, 1), \ldots, (s, N)$, and let $r_s \in [0, 1]$ be the fraction of these computations that accept;
   2.c. If $|r_s - r_{(i(s), j(s))}| > 0.5\alpha$, set $b_{\text{adv}} := 0$.

Note that in step (2.b), the joint state on the proof registers may change after each application of $C_{(i(s), j(s))}$. If $S = 1$, the proof registers go untouched and $b_{\text{adv}} = 1$.

Next, the subroutine $B$ acts as follows. $B$ measures the value $S$ chosen by $A$ (and stored in the ancilla register). It then chooses $t \in [N]$ uniformly and simulates $Q_n'$ applied to input $x$ and with the $(S,t)$th proof register as the quantum advice state for $Q_n'$, taking the resulting bit as $b_{\text{out}}$.

$Y_n$ can clearly be implemented in size $\text{poly(size}(Q_n))$. Now let us analyze $Y_n$ to establish items (i)–(ii) in the theorem’s conclusion. For item (i), consider the execution $Y_n(x, \sigma)$ on the advice state $\sigma$ which is the tensor product of $MN$ independent copies of the state $\sigma$ guaranteed to exist by item (i) in our application of Lemma 13. Then in the operation of the subroutine $A$, for each execution of step (2.b) (indexed by an $s \in [M]$), the expected fraction $E[r_s]$ is within $\pm 1.0\alpha$ of $r_{(i(s), j(s))}$ after conditioning on $i(s), j(s)$. Also, the outcome of the executions of $C_{(i(s), j(s))}$ are mutually independent, since $\sigma$ is a product state over the $MN$ registers. Chernoff bounds and our setting of $N$ then imply that $r_s$ is within $\pm 5\alpha$ of $r_{(i(s), j(s))}$ with probability $> 1 - e^{-n}/M$. A union bound over all $s \in [M]$ completes the proof of item (i) in the theorem.
We now turn to item (ii). Let \( \mathcal{F} \) be any \( P \)-qubit state for which, in the execution of \( Y_n(x, \mathcal{F}) \), we have \( \mathbb{E}[b_{\text{adv}}] \geq n^{-d} \). (If this holds for some \( x \in \{0,1\}^n \), then it holds for all such \( x \); we fix some such \( x \) in what follows.) For \( s \in [M-1] \), let \( q_s \) denote the probability that \( \hat{r}_s - r_{(i(s),j(s))} \leq 0.5\alpha \) holds in the execution of subroutine \( A \) in the operation of \( Y_n(x, \mathcal{F}) \), conditioned on the following two events:

1. \( S = s + 1 \), so that the For loop in step 2 of \( A \) executes for the value \( s \);
2. \( |\hat{r}_s - r_{(i(s),j(s))}| \leq 0.5\alpha \) for all \( s' < s \).

Note that the value \( q_s \) would be unchanged if in the first item above we instead conditioned on \( [S = s''] \) for any \( s'' > s \). Also, for future use we define \( \mathcal{F}^{(s)} \) as the \( Nmp \)-qubit reduced state on the proof registers \( (s,1),(s,2),\ldots,(s,t) \), conditioned on items 1 and 2 above.

Let \( I_{\text{bad}} \subseteq [M-1] \) be the set of indices \( s \) for which \( q_s < 1 - \alpha/(n^{3d}mk) \). We will upper-bound \( \Pr[S \in I_{\text{bad}} \land b_{\text{adv}} = 1] \). Let \( r_{\text{early}} \) be the first \( W := \lfloor n^{3d}mk/\alpha \rfloor \) elements of \( I_{\text{bad}} \) in increasing order (or if \( I_{\text{bad}} \leq W \), then \( r_{\text{early}} := I_{\text{bad}} \)). Let \( I_{\text{late}} := I_{\text{bad}} \setminus r_{\text{early}} \). We have

\[
\Pr[S \in I_{\text{bad}} \land b_{\text{adv}} = 1] \leq W/M + \Pr[S \in I_{\text{late}} \land b_{\text{adv}} = 1],
\]

since \( \Pr[S \in r_{\text{early}}] \leq W/M \). If \( I_{\text{bad}} \neq \emptyset \), then conditioned on any value of \( S \) with \( S > \max(I_{\text{early}}) \), the probability that \( b_{\text{adv}} \) is not set to 0 in the \( S - 1 \) executions of step 2 of \( A \) equals

\[
\prod_{s < S} q_s \leq \prod_{s \in r_{\text{early}}} q_s (1 - \alpha/(n^{3d}mk))^W \leq n^{-4d}.
\]

Thus, \( \Pr[S \in I_{\text{late}} \land b_{\text{adv}} = 1] \leq n^{-4d} \), and \( \Pr[S \in I_{\text{bad}} \land b_{\text{adv}} = 1] \leq n^{-4d} + W/M \); this is at most \( 2n^{-4d} \), by our setting to \( M \). It follows that

\[
\Pr[S \in I_{\text{bad}} | b_{\text{adv}} = 1] \leq \frac{2n^{-4d}}{\Pr[b_{\text{adv}} = 1]} \leq 2n^{-3d},
\]

using our assumption in item (ii) that \( \Pr[b_{\text{adv}} = 1] \geq n^{-d} \).

Next, we claim that for each \( s \in [M] \setminus I_{\text{bad}} \), the conditional expectation \( \mathbb{E}[b_{\text{out}} = 1 | S = s \land b_{\text{adv}} = 1] \) satisfies

\[
|\mathbb{E}(b_{\text{out}} = 1 | S = s \land b_{\text{adv}} = 1) - \mathbb{E}(Q_n(x, \rho_n^s))| \leq n^{-3d}.
\]

To see this, fix any such \( s \). First note that, if we condition on \( S = s \land b_{\text{adv}} = 1 \), the joint post-conditioned state of the proof registers \( (s,1),(s,2),\ldots,(s,t) \) is precisely \( \mathcal{F}^{(s)} \) as defined previously. Now consider the experiment in which we choose a pair \( (i,j) \) uniformly from \( [m] \times [k] \) and apply \( C_{i,j} \) to each of these proof registers, prepared in the joint state \( \mathcal{F}^{(s)} \), and let \( \hat{r}_{(i,j)} \in [0,1] \) be the fraction of 1s measured. The probability in this experiment that \( |\hat{r}_{(i,j)} - r_{(i,j)}| \leq 0.5\alpha \) is, by the linearity of quantum mechanics, equal to \( q_s \); this is greater than \( 1 - \alpha/(n^{3d}mk) \) since \( s \notin I_{\text{bad}} \). Then by an application of Markov’s inequality, for every \( (i^*,j^*) \in [m] \times [k] \), if we perform this experiment on \( \mathcal{F}^{(s)} \) with the fixed choice \( (i,j) = (i^*,j^*) \), then we see \( |\hat{r}_{(i^*,j^*)} - r_{(i^*,j^*)}| \leq 0.5\alpha \) with probability greater than \( 1 - \alpha n^{-3d} > 1 - 0.2\alpha \). Thus \( |\mathbb{E}(\hat{r}_{(i^*,j^*)}) - r_{(i^*,j^*)}| \leq 0.7\alpha \).

For \( t \in [N] \), let \( \sigma^{(s,t)} \) denote the reduced state of \( \mathcal{F}^{(s)} \) on the \((s,t)\) proof register. Let \( \sigma^{(s,t)\text{avg}} := \frac{1}{N} \sum_{t \in [N]} \sigma^{(s,t)} \), and note that in the experiment above with
fixed pair \((i^*, j^*)\), we have \(\mathbb{E}[\hat{r}_{(i^*, j^*)}] = \mathbb{E}[C_{(i^*, j^*)}(\sigma^{(s, \text{avg})})]\). By our work above, 
\[|\mathbb{E}[C_{(i^*, j^*)}(\sigma^{(s, \text{avg})})] - r_{(i^*, j^*)}| \leq 0.7\alpha.\] As \((i^*, j^*)\) was arbitrary, it follows from item (ii) in our application of Lemma 13 that 
\[|\mathbb{E}[Q_n(x, \sigma^{(s, \text{avg})})] - \mathbb{E}[Q_n(x, \rho_n^*)]| \leq n^{-2d}.
\]

Now let us return to the definition of the algorithm \(Y_n\) and note that, in the execution \(Y_n(x, \overline{s})\), if we condition on \([b_{\text{adv}} = 1 \land S = s]\), then \(Y_n\) simulates \(Q_n\) applied to \(x\) and to an advice state whose density operator is (under our conditioning) precisely that of \(\sigma^{(s, \text{avg})}\), and \(Y_n\) outputs the resulting bit. Thus, 
\[|\mathbb{E}[b_{\text{out}}|b_{\text{adv}} = 1 \land S = s] - \mathbb{E}[Q_n(x, \rho_n^*)]| \leq n^{-2d},
\]
and since \(s\) was an arbitrary element of \([M]\) \(\setminus I_{\text{bad}}\), we also have 
\[|\mathbb{E}[b_{\text{out}}|b_{\text{adv}} = 1 \land S \notin I_{\text{bad}}] - \mathbb{E}[Q_n(x, \rho_n^*)]| \leq n^{-2d}.
\]
Combining our findings, we see that 
\[|\mathbb{E}[b_{\text{out}}|b_{\text{adv}} = 1] - \mathbb{E}[Q_n(x, \rho_n^*)]| \leq \Pr[S \in I_{\text{bad}}|b_{\text{adv}} = 1] + n^{-2d}
\leq 2n^{-3d} + n^{-2d}
\leq n^{-d}
\]
for \(n > 1\). The statement of item (ii) is trivial for \(n = 1\), so this proves item (ii), completing the proof of the theorem. 

4.3. Bestiary of quantum complexity classes. In this section we define some old and new complexity classes which our techniques shed light on. Given a language \(L \subseteq \{0, 1\}^*\), let \(L : \{0, 1\}^* \rightarrow \{0, 1\}\) be the characteristic function of \(L\). We now give a formal definition of the class \(\text{BQP/qpoly}\).

**Definition 16.** A language \(L\) is in \(\text{BQP/qpoly}\) if there exists a polynomial-time quantum algorithm \(A\) and polynomial-time computable function \(p(n) \leq \text{poly}(n)\) such that for all \(n\), there exists an advice state \(\rho_n\) on \(p(n)\) qubits such that \(A(x, \rho_n)\) outputs \(L(x)\) with probability \(\geq 2/3\) for all \(x \in \{0, 1\}^n\).

Closely related to quantum advice are quantum proofs. We now recall the definition of QMA, a quantum version of NP.

**Definition 17.** A language \(L\) is in QMA if there exists a polynomial-time quantum algorithm \(A\) and polynomial-time computable function \(p(n) \leq \text{poly}(n)\) such that for all \(x \in \{0, 1\}^n\) the following hold:

(i) If \(x \in L\), then there exists a witness \(\rho_x\) on \(p(n)\) qubits such that \(A(x, \rho_x)\) accepts with probability \(\geq 2/3\).

(ii) If \(x \notin L\), then \(A(x, \rho)\) accepts with probability \(\leq 1/3\) for all \(\rho\).

We will define some complexity classes involving untrusted (classical or quantum) advice that depends only on the input length. This notion has been studied before: Chakaravarthy and Roy [16] and Fortnow, Santhanam, and Williams [17] defined the complexity class \(\text{ONP}\) ("Oblivious NP"), which is like \(\text{NP}\) except that the witness can depend only on the input length. Independently, Aaronson [5] defined the complexity class \(\text{YP}\), which is easily seen to equal \(\text{ONP} \cap \text{coONP}\). We will adopt the "Y" notation in this paper.

\[\text{YP}\] stands for "Yoda polynomial-time," a nomenclature that seems to make neither more nor less sense than "Arthur–Merlin."
We now give a formal definition of \( Y_P \), as well as a slight variant called \( Y_{P^*} \).

**Definition 18.** A language \( L \) is in \( Y_P \) if there exist polynomial-time algorithms \( A, B \) and a polynomial-time computable function \( p(n) \leq \text{poly}(n) \) such that the following hold.

(i) For all \( n \), there exists an advice string \( y_n \in \{0,1\}^{p(n)} \) such that \( A(x,y_n) = 1 \) for all \( x \in \{0,1\}^n \).

(ii) If \( A(x,y) = 1 \), then \( B(x,y) = L(x) \).

\( L \) is in \( Y_{P^*} \) if moreover \( A \) ignores \( x \), depending only on \( y \).

Clearly \( P \subseteq Y_{P^*} \subseteq Y_P \subseteq P/\text{poly} \cap \text{NP} \cap \text{coNP} \). Also, Aaronson [5] showed that \( \text{ZPP} \subseteq Y_P \). We will be primarily interested in a quantum analogue of \( Y_{P^*} \). This analogue builds on Definition 14. However, it also models a distinctively quantum ingredient: we consider two-phase protocols in which an untrusted quantum advice state is first tested in an input-oblivious fashion and, if accepted, is passed along in altered form to be used in computation with the given input. This model is natural, since quantum measurements unavoidably alter the measured states; the alterations performed by the initial testing are also crucial to the power of these protocols.

(Roughly speaking, this works as follows: if the given quantum advice state is a mixture \( \rho = t\rho_1 + (1-t)\rho_2 \) of a “good state” \( \rho_1 \) which passes our test with high probability and is useful for computation, and a “bad state” \( \rho_2 \) which passes with low probability, then conditioning on passing the test “filters out” the contribution of \( \rho_2 \), making the resulting state more useful. We emphasize, however, that the test involves various measurements that significantly alter even a state that passes with high probability. The technical core of this procedure has already been given in Theorem 15.)

**Definition 19 (\( Y_{QP} \) and \( Y_{QP^*} \)).** A language \( L \) is in \( Y_{QP} \) if there exists a uniform (i.e., polynomial-time constructible) family of advice-testing quantum circuits \( \{Y_n(x,\rho)\}_{n>0} \) (as per Definition 14). Each \( Y_n \) is of size \( \text{poly}(n) \) and takes as input an \( x \in \{0,1\}^n \) and a \( p(n) \)-qubit state \( \rho \) (for some \( p(n) \leq \text{poly}(n) \)). We have the following properties:

(i) For all \( n \), there exists a setting \( \rho_n \) to the quantum advice register such that for any \( x \in \{0,1\}^n \), in the execution of \( Y \) on \( (x,\rho_n) \) we have \( \mathbb{E}[b_{\text{adv}}] \geq 9/10 \).

(ii) If for any settings \( (x,\rho) \) to the input and advice registers we have \( \mathbb{E}[b_{\text{adv}}] \geq 1/10 \), then \( \Pr[b_{\text{out}} = L(x)|b_{\text{adv}} = 1] \geq 9/10 \).

\( L \) is in \( Y_{QP^*} \) if the circuit family \( \{Y_n\}_{n>0} \) can be additionally be chosen to obey the input-oblivious property.

We define the corresponding nonuniform classes \( Y_{QP}/\text{poly}, Y_{QP^*}/\text{poly} \) by removing the requirement that the family \( \{Y_n\}_{n>0} \) be uniform.

Clearly \( BQP \subseteq Y_{QP^*} \subseteq Y_{QP} \subseteq BQP/\text{qpoly} \cap \text{QMA} \cap \text{coQMA} \).

### 4.4. Characterizing quantum advice

We now prove the following characterization of \( BQP/\text{qpoly} \), which immediately implies (and strengthens) Theorem 2.

**Theorem 20.** \( BQP/\text{qpoly} = Y_{QP^*}/\text{poly} \).

**Proof.** One direction \( (Y_{QP^*}/\text{poly} \subseteq BQP/\text{qpoly}) \) is obvious, since untrusted quantum advice and trusted classical advice can both be simulated by trusted quantum advice. We prove that \( BQP/\text{qpoly} \subseteq Y_{QP^*}/\text{poly} \). Let \( L \in BQP/\text{qpoly} \), and let \( Q(x,\rho), \{\rho_n\}_{n>0} \) be a polynomial-time quantum algorithm (given by a uniform circuit family \( \{Q_n\}_{n>0} \) for input length \( n \)) and polynomial-size quantum advice family defining \( L \). We insist that \( Q \) enjoy completeness and soundness parameters \((99/100, 1/100)\) in place of \( 2/3, 1/3 \) in Definition 16; this can be achieved by standard soundness amplification by providing multiple copies of the trusted advice state. We apply Theorem 15 to \( Q_n(x,\rho) \) and \( \{\rho_n\}_{n>0} \) with \( d := 1 \) for each \( n \). We obtain a (nonuniform) family of
input-oblivious advice-testing quantum circuits \( \{Y_n\}_{n>0} \), such that

(i) for each \( n \), there is a state \( \sigma \) such that in the execution of \( Y_n(x, \sigma) \) we have

\[
\Pr[b_{\text{adv}} = 1] \geq 1 - \varepsilon^{-n};
\]

(ii) for any \( n > 1 \) and advice state \( \sigma \), it holds that for each \( x \in \{0,1\}^n \), in the execution of \( Y_n(x, \sigma) \),

\[
\Pr[b_{\text{adv}} = 1] \geq n^{-1} \implies |\mathbb{E}[b_{\text{out}}|b_{\text{adv}} = 1] - \mathbb{E}[Q_n(x, \rho_n^*)]| \leq n^{-1}.
\]

Now by the definitions of \( Q_n \) and \( \rho_n^* \), we have \(|\mathbb{E}[Q_n(x, \rho_n^*)] - L(x)| \leq 1/100 \) for all \( x \in \{0,1\}^n \). Thus, if \( n \) is sufficiently large, we have the following:

(iii) For any advice state \( \sigma \) for length \( n \), it holds that for each \( x \in \{0,1\}^n \), in the execution of \( Y_n(x, \sigma) \), if \( \Pr[b_{\text{adv}} = 1] \geq 1/10 \), then we have

\[
|\mathbb{E}[b_{\text{out}}|b_{\text{adv}} = 1] - L(x)| \leq n^{-1} + 1/100 \leq 1/10.
\]

Thus the family \( \{Y_n\}_{n>0} \) witnesses that \( L \in \text{YQP}^* / \text{poly} \). This proves Theorem 20.

One interesting consequence of Theorem 20 is that \( \text{YQP} / \text{poly} = \text{YQP}^* / \text{poly} \). We do not know of an easier proof of this equality, and we leave as an open question whether, in the uniform setting, the corresponding equality \( \text{YQP} = \text{YQP}^* \) holds.

Since we never critically used the assumption that the \( \text{BQP} / \text{qpoly} \) machine computes a language (i.e., a total Boolean function), a strengthening of Theorem 20 that we can easily observe is the promise-class equality \( \text{PromiseBQP} / \text{qpoly} = \text{PromiseYQP}^* / \text{poly} = \text{PromiseYQP} / \text{poly} \).

4.5. Application to quantum communication. We can also use our Theorem 15 to obtain a new positive result about the possibility of robust communication over fault-prone quantum communication channels (augmented with a trustworthy classical channel). Our result does not assume any particular error model for quantum channels. Rather, it asserts that a successful outcome is achieved by the protocol under a perfect transmission, and that the protocol guards against a certain type of bad outcome under any corruption of the transmitted quantum state.

Theorem 21. Suppose that Alice, who is computationally unbounded, has a classical description of an \( N \)-qubit quantum state \( \rho^* \). She wants to send \( \rho^* \) to Bob, who is computationally bounded. Assume that Alice has at her disposal a noiseless one-way classical channel to Bob, as well as a noisy one-way quantum channel. Bob holds a binary measurement \( E \) for which he wishes to learn \( \mathbb{E}[E(\rho^*)] \) to within an accuracy \( \varepsilon > 0 \). We assume \( E \) is implemented by a circuit with at most \( m \) gates (under some fixed finite basis); here \( m \) is known to Alice, but \( E \) is known only to Bob.

Then for all \( \varepsilon > 0 \), there exists a protocol whereby

- Alice sends Bob a classical string \( z \) of \( \text{poly}(N, m, 1/\varepsilon) \) bits, as well as a state \( \sigma \) of \( \text{poly}(N, m, 1/\varepsilon) \) qubits;
- Bob receives \( z \) together with a possibly-corrupted version \( \tilde{\sigma} \) of \( \sigma \) and performs a (nonbinary) measurement \( f_z(E) \) on \( \tilde{\sigma} \), outputting a real value \( \beta \in [0, 1] \) along with a “success bit” \( b_{\text{succ}} \in \{0, 1\} \). This \( f_z(E) \) can be computed and performed in \( \text{poly}(N, m, 1/\varepsilon) \) steps, given \( z \) together with a description of \( E \).

The following properties hold:

(i) If \( \tilde{\sigma} = \sigma \), then with probability greater than \( 1 - 2^{-N} \) we have \( |\beta - \mathbb{E}[E(\rho^*)]| \leq \varepsilon \) and \( b_{\text{succ}} = 1 \).

(ii) For every \( \tilde{\sigma} \) and every measurement \( E \) as described above, with probability at least \( 1 - 2^{-N} \), Bob either sets \( b_{\text{succ}} = 0 \) or outputs a \( \beta \in [0, 1] \) such that \( |\beta - \mathbb{E}[E(\rho)]| \leq \varepsilon \).
Proof. We will apply Theorem 15 to the communication setting. The string $z$ plays the role of the trusted classical advice; the state $\sigma$ plays the role of the untrusted quantum advice; the measurement $E$ plays the role of the input $x$; Bob plays the role of the advice-testing algorithm $Y$. We will perform multiple trials to increase our confidence.

We prove the result under the assumption that $\varepsilon$ is at least inverse-polynomial in $N$, which allows us to apply our prior work more directly. We will assume that $\varepsilon \geq N^{-1}$; the general result will follow, since in our construction we may begin by padding the quantum register with $1/\varepsilon$ dummy qubits. The protocol will succeed for sufficiently large $N$—smaller values of $N$ can be handled by brute force.

Let $n > 0$ be a fixed description length adequate to describe any $m$-gate measurement $E$ that may be held by Bob in our communication scenario, for our specific values of interest $m, N$; here we can take $N \leq n \leq \text{poly}(N, m)$. Let $Q_n(E, \xi)$ be a quantum circuit which receives a description of a binary measurement $E$ of description length $n$, described by a circuit in our fixed finite basis. $Q_n$ also receives a quantum state $\xi$ on $N$ qubits and outputs the result of $E(\xi)$. This $Q_n$ can be implemented in size $\text{poly}(n, N) \leq \text{poly}(N, m)$. Let $Y_n = Y_n(E, \sigma)$ be the input-oblivious advice-testing circuit of size $\text{poly}(N, m)$ given by Theorem 15 for $(Q_n, \rho^*, d := 2)$.

In our protocol, Alice sends a description of $Y_n$ as the reliable classical message $z$ to Bob, and for the fault-prone quantum state $\sigma$, Alice sends $T := n^4$ independent copies of the $P$-qubit advice state $\sigma^*$ guaranteed to exist by item (i) of Theorem 15; we have $|z| \leq \text{poly}(N, m)$ and $\sigma$ is on $\text{poly}(N, m)$ qubits, as needed.

Bob receives the (correct) string $z$, and a quantum state $\sigma$ on $T \cdot P$ qubits, where we consider this state to be defined over $T$ registers called the “transmission registers.” Bob acts as follows (these steps define the measurement $f_i(E)$): For $i = 1, 2, \ldots, T$, Bob executes $Y_n$ applied to input bitstring $E$, classical advice $z$, and with the $i$th transmission register used as the quantum advice state. For each such application of $Y_n$ in turn, Bob measures the bits $b_{\text{adv},i}, b_{\text{out},i}$. (Here, we use $b_{\text{adv},i}$ to denote the value of $b_{\text{adv}}$ on the $i$th trial, and similarly for $b_{\text{out},i}$.) If $b_{\text{adv},i} = 0$ for any $i$, Bob sets $b_{\text{suc}} := 0$ (and sets $\beta := 0$, say). Otherwise, Bob sets $b_{\text{suc}} := 1$ and outputs the value $\beta := \frac{1}{T} \sum_{i \in [T]} b_{\text{out},i}$.

Let us analyze this procedure. First note that when Bob receives the same state $\sigma^*$ sent by Alice, item (i) of Theorem 15 tells us that each $b_{\text{adv},i}$ equals 1 with probability at least $1 - e^{-n}$. Then by a union bound over all $i$, for sufficiently large $N$, each of these bits equals 1 with probability at least $1 - 2^{-(n+1)}$. So $\Pr[b_{\text{suc}} = 1] \geq 1 - 2^{-(n+1)}$.

Also, item (ii) of Theorem 15 tells us that each $b_{\text{out},i}$ satisfies $|E[b_{\text{out},i}] - E_E(\rho^*)| = |E[b_{\text{out},i}] - Q_n(E, \rho^*)| \leq n^{-2}$, and these bits are independent. By Chernoff’s bound, $\Pr[|\beta - E_E(\rho^*)| \leq n^{-1}] \geq 1 - 2^{-(n+1)}$ for large $n$. A union bound completes the proof of item (i) in the Theorem’s statement.

For item (ii), consider any quantum state $\sigma$ on $T \cdot P$ qubits received by Bob. Each execution of Bob’s algorithm determines, for each $i \in [T]$, a mixed state $\xi_i$ on $P$ qubits that describes the reduced state on the $i$th transmission register, immediately after Bob has applied $Y_n$ to the first $(i - 1)$ transmission registers and measured $b_{\text{adv},1}, b_{\text{out},1}, \ldots, b_{\text{adv},i-1}, b_{\text{out},i-1}$. We consider $\xi_i$ as a random variable determined by Bob’s execution (acting on the pair $z, \sigma$).

Say that state $\xi$ on $P$ qubits is good if in the execution of $Y_n(x, \xi)$, we have $\Pr[b_{\text{adv}} = 1] \geq n^{-2}$. Let $G \subseteq [T]$ be the (random) set $\{i : \xi_i \text{ is good}\}$. Conditioned on any outcomes $b_{\text{adv},1}, b_{\text{out},1}, \ldots, b_{\text{adv},i-1}, b_{\text{out},i-1}$ which determine a state $\xi_i$ which is good, item (ii) of Theorem 15 tells us that the expected value of $b_{\text{out},i}$, conditioned on $[b_{\text{adv},i} = 1]$, is within $\pm n^{-2}$ of $E(Q(E, \rho^*)) = E(E(\rho^*))$.
For $i \in [T]$, let the random variable $Z_i \in \{0,1\}$ be defined by

$$Z_i := \begin{cases} b_{out,i} & \text{if } i \in G \text{ and } b_{adv,i} = 1, \\ \text{an independent coin flip with bias } \mathbb{E}[E(\rho^*)] & \text{otherwise.} \end{cases}$$

Note that we have the relation $|\mathbb{E}[Z_i | Z_1, \ldots, Z_{i-1}] - \mathbb{E}[E(\rho^*)]| \leq n^{-2}$. By an application of Azuma’s inequality,

$$\Pr \left[ \frac{1}{T} \sum_{i \in T} Z_i - \mathbb{E}[E(\rho^*)] \geq n^{-2} + 0.5n^{-1} \right] \leq \exp \left(-\Omega((.5n^{-1})^2 \cdot T)\right) \leq e^{-n}$$

for sufficiently large $N$.

Now, it is clear that $\Pr[[T] \setminus G > n \wedge b_{suc} = 1] \leq (n^{-2})^n < e^{-n}$. If $|[T] \setminus G| \leq n$ and $b_{suc} = 1$, then we have $b_{adv,i} = 1$ for all $i$ so that $|\frac{1}{T} \sum_{i \in T} Z_i - \frac{1}{T} \sum_{i \in T} b_{out,i}| \leq n/T$. Combining this with our previous work, it follows that

$$\Pr \left[ b_{suc} = 1 \wedge \frac{1}{T} \sum_{i \in T} b_{out,i} - \mathbb{E}[E(\rho^*)] \geq (n^{-2} + 0.5n^{-1}) + n/T \right] \leq 2 \cdot e^{-n} \leq 2^{-n}$$

for large $N$; for such $N$ we have $n^2 + 0.5n^{-1} + n/T \leq n^{-1}$. As $n^{-1} \leq N^{-1} \leq \epsilon$, this gives item (ii). \(\square\)

5. Local Hamiltonians and the complexity of preparing quantum advice states. In this section we begin the proof of Theorem 1 from the introduction, which we will obtain from a slightly more general result.

Let $B^\otimes N$ denote the $2^N$-dimensional complex Hilbert space whose unit ball consists of the $N$-qubit pure quantum states. Recall that a Hamiltonian on $N$-qubit states is a Hermitian operator $H : B^\otimes N \rightarrow B^\otimes N$. (We will only discuss the action of Hamiltonians on pure states.) $H$ is called a $k$-local Hamiltonian if it can be written as $H = \sum_{i=1}^{N} H_i$, where each $H_i$ is a Hermitian operator acting on at most $k$ qubits.

If we combine Theorem 15 with known QMA-completeness reductions (and some further analysis of these reductions), we can obtain a striking consequence for quantum complexity theory. Namely, the preparation of quantum advice states can always be reduced to the preparation of ground states of 2-local Hamiltonians—despite the fact that quantum advice states involve an exponential number of constraints, while ground states of local Hamiltonians involve only a polynomial number. (In particular, if ground states of local Hamiltonians can be prepared by polynomial-size circuits, then we have not only QMA = QCMA, but also BQP/qpoly = BQP/poly.) Our objective in sections 6 and 7 is to prove the following result.

**Theorem 22.** Let $C^*(z, \rho)$ be a quantum circuit of $T$ gates (each 2-local) taking an input string $z \in \{0,1\}^N$ and a quantum state $\rho$ on $\ell$ qubits (we may assume $\ell \leq 2T$). Let $\rho^*$ be a distinguished state on $\ell$ qubits. For all $\delta > 0$, there exists a second quantum circuit $C'$ and a 2-local Hamiltonian $H$ acting on $\ell' \leq \text{poly}(T, N, 1/\delta)$ qubits, such that for any ground state $|\psi\rangle$ of $H$ and any input $z \in \{0,1\}^N$,

$$|\mathbb{E}[C'(z, |\psi\rangle|\psi\rangle)] - \mathbb{E}[C^*(z, \rho^*)]| \leq \delta.$$  

While a description of $H$ may not be efficiently computable, $C'$ can be constructed in (classical, deterministic) time $\text{poly}(T, N, 1/\delta)$, given $\delta$ and descriptions of $C^*$ and $H$.  

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
Our proof of Theorem 22 combines Theorem 15 with the following result on the expressive power of ground states of 2-local Hamiltonians.

**Theorem 23.** Let $V(\xi)$ be a quantum “verifier” circuit of $T$ gates (each 2-local), which acts on an $m$-qubit quantum state $\xi$ and an ancilla register of $N - m$ qubits (we may assume $N \leq 2T$), with the ancilla register initially in the all-zero state. Suppose that $V$ defines a binary measurement on $\xi$. Fix any $\varepsilon > 0$, and assume that \( \max_p \mathbb{E}[V(p)] \geq 1 - \varepsilon \). Then there exists

- a 2-local Hamiltonian $H_{V,\varepsilon}$ acting on $N'$-qubit states for some value $N' \leq \text{poly}(T, 1/\varepsilon)$, expressed as a sum of 2-local terms $H_i$ with operator norm $\frac{1}{\text{poly}(T, 1/\varepsilon)} \leq ||H_i|| \leq \text{poly}(T, 1/\varepsilon)$ and
- a quantum operation $R_{V,\varepsilon}$ mapping $N'$-qubit states to $m$-qubit states, implemented by a quantum circuit with $\text{poly}(T, 1/\varepsilon)$ gates, for which the following property holds: if $|\psi\rangle\langle\psi|$ is any ground state of $H_{V,\varepsilon}$, then for $\xi := R_{V,\varepsilon}(|\psi\rangle\langle\psi|)$ we have

$$
\mathbb{E}[V(\xi)] \geq 1 - \kappa \cdot T^\kappa \varepsilon^{1/\kappa},
$$

where $\kappa > 1$ is an absolute constant. Furthermore, $H_{V,\varepsilon}$ and $R_{V,\varepsilon}$ can be constructed in (classical, deterministic) time $\text{poly}(T, 1/\varepsilon)$, given a description of $V$.

We will obtain Theorem 23 by a detailed analysis of known QMA-completeness reductions. We defer the proof.

Theorem 1 is now easily obtained.

**Proof of Theorem 1.** Define a circuit $C^*(E, \rho)$ which takes as input a circuit $E$ of size $n^c$ defining a binary measurement, and a quantum state $\rho$ on $n$ qubits, and executes $C(\rho)$. The circuit $C^*$ can be implemented in size $\text{poly}(n)$ using 2-local gates, and we have $\mathbb{E}[C^*(E, \rho)] = \mathbb{E}[E(\rho)]$ for all inputs $(E, \rho)$ to $C^*$. The result follows by an application of Theorem 22 to $C^*$ and $\rho^*$. \(\square\)

**Proof of Theorem 22.** We may (by a padding argument as in the proof of Theorem 21) assume that $\delta \geq 2/N$. We may also assume that $N \geq 2$ and $\delta < .5$. Let $n$ be a value such that for any $z \in \{0, 1\}^N$, a description of length exactly $n$ can be given for the specialized circuit $C^*(z, \cdot)$; here, we can take $N \leq n \leq \text{poly}(T, N)$.

Let $P(C, \xi)$ be a polynomial-time quantum algorithm which receives a description of a circuit $C$, of description length $n$, defining a binary measurement, and which applies $C$ to an $\ell$-qubit input state $\xi$ (where $\ell$ is as in the statement of Theorem 22), outputting the result.

Let $Y_n = Y_n(C, \overline{\sigma})$ be the input-oblivious advice-testing circuit provided by Theorem 15 for $(P, \rho^*, d := 2)$. The number of gates in $Y_n$ is at most $\text{poly}(n) \leq \text{poly}(T, N)$. Let $p$ be the number of qubits in the quantum advice register for $Y_n$. Let $C' = C'(z, \overline{\sigma})$ be the circuit which executes $Y_n(C^*(z, \cdot), \overline{\sigma})$ and outputs the measured bit $b_{\text{out}}$.

Next we will define $H$ as in the theorem statement, using Theorem 23. The circuit $Y_n$ has two subcircuits $A, B$, following Definition 14. Let $V(\overline{\sigma})$ be the circuit which executes $A(\overline{\sigma})$ and outputs the measured bit $b_{\text{adv}}$. By item (i) of Theorem 15, there exists a state $\overline{\sigma}^+$ on $p$ qubits for which $\mathbb{E}[V(\overline{\sigma}^+)] \geq 1 - 2^{-n}$. For large enough $N$ this is greater than $1 - \varepsilon$, where $\varepsilon := (\delta/ (2^c N^c))^k$ for the constant $\kappa > 1$ from Theorem 23.

Theorem 23 now gives us a Hamiltonian $H = H_{V,\varepsilon}$ and quantum operation $R = R_{V,\varepsilon}$. These have the property that for any ground state $|\psi\rangle\langle\psi|$ of $H$, for $\xi := ||\psi\rangle\langle\psi|| = \mathbb{E}[V(\xi)]$. 

\(8\)The state output by $R_{V,\varepsilon}$ may be mixed, even if its input state is pure.
\[ R(|\psi\rangle\langle\psi|) \text{ we have} \]
\[ E[V(\xi)] \geq 1 - \frac{\delta}{2} > n^{-2} \]
(the first inequality holding by of our choice of \(\varepsilon\). By definition of \(V\), this means that in the execution of \(Y_n(\rho, \lambda, \xi)\), we have \(\Pr[b_{adv} = 1] > n^{-2}\). (This holds for any \(z \in \{0, 1\}^N\); the expectation above is independent of \(z\) since \(Y_n\) has the input-oblivious testing property.) By our guarantee for \(Y_n\) given in Theorem 15, item (ii), it follows that in the execution of \(Y_n(\rho, \lambda, \xi)\) on any circuit \(C(\cdot, \cdot)\) of description length \(n\),
\[ |E[b_{out}|b_{adv} = 1] - E[P(\rho, \lambda, \xi)]| \leq n^{-2}. \]
Recall from our definition that the output bit of \(C'(z, \xi)\) is distributed as \(b_{out}\) in the execution of \(Y_n(C^*(z, \xi), \xi)\). Thus,
\[ |E[C'(z, \xi)] - E[P(C^*(z, \cdot), \rho^*)]| \leq n^{-2} + \Pr[b_{adv} = 0], \]
where \(b_{adv}\) is as in the execution of \(Y_n(\rho, \lambda, \xi)\). We have seen that in this execution \(\Pr[b_{adv} = 1] > 1 - \delta/2\), so the right-hand side above is at most \(n^{-2} + \delta/2 \leq \delta\). Also, by our definitions, \(E[P(C(z, \cdot), \rho^*)] = E[C(z, \rho^*)]\). This proves the theorem. \(\square\)

6. Reduction to 5-local Hamiltonians. In sections 6 and 7, we prove Theorem 23. The proof is achieved by a sequence of reductions. Each reduction was defined previously, but we need to establish facts about these reductions not found in previous references [26, 7, 25, 29]. This requires careful work.

For a Hamiltonian \(H\), we use \(\lambda_1(H) \leq \cdots \leq \lambda_M(H)\) to denote the real eigenvalues of \(H\), counted according to their geometric multiplicity\(^9\) and sorted in nondecreasing order. We will use \(||H||\) to denote the operator norm of \(H\).

The energy of a pure state \(|\psi\rangle\rangle \text{ with respect to } H\text{ is defined as } \langle\psi|H|\psi\rangle\). It is a basic fact that for all vectors \(|\psi\rangle\rangle\) we have \(\langle\psi|H|\psi\rangle \geq \lambda_1(H) \cdot |||\psi|||\rangle\), and the ground states of \(H\) are precisely those unit vectors for which equality holds. In proving Theorem 23, a key role will be played by nearly minimal energy states—those unit vectors \(|\psi\rangle\rangle\) for which \(\langle\psi|H|\psi\rangle \approx \lambda_1(H)\).

In this section, we will use the original QMA-completeness reduction, due to Kitaev, Shen, and Vyalyi [26], to prove Theorem 24 below, a variant of Theorem 23. This variant is weaker, in that the Hamiltonian \(H\) produced is only required to have locality 5, rather than 2; but it is stronger in that the reduction \(R\) is required to produce a useful state given any nearly minimal energy state for \(H\) (not just any ground state). This “robust” guarantee will be important in our subsequent construction of 2-local Hamiltonians. Theorem 24 is also stronger in that \(H, R\) are chosen independent of \(\varepsilon\), although this property is not essential for our work.

**THEOREM 24.** Let \(V(\xi)\) be a quantum “verifier” circuit of \(T\) gates (each 2-local), which acts on an \(m\)-qubit quantum state \(\xi\) and an ancilla register of \(N - m\) qubits (we may assume \(N \leq 2T\)), with the ancilla register initially in the all-zero state. Suppose that \(V\) defines a binary measurement on \(\xi\). Then there exists

- a 5-local Hamiltonian \(HV\) acting on \(N'\)-qubit states for some \(N' \leq O(T)\), expressed as a sum of 5-local terms \(H_i\), of operator norm \(\frac{1}{\text{poly}(T, 1/\varepsilon)} \leq ||H_i|| \leq \text{poly}(T, 1/\varepsilon)\), and

\(^9\)That is, an eigenvalue \(\lambda\) appears \(p\) times in the list, where \(p\) is the dimension of the eigenspace for \(\lambda\). By the spectral theorem we have \(M = \dim(B^\otimes N) = 2^N\).
• a quantum operation $R_V$ mapping $N'$-qubit states to $m$-qubit states, implemented by a quantum circuit with $\poly(T)$ gates,

for which the following property holds for any $\varepsilon > 0$: if $\max_\rho \mathbb{E}[V(\rho)] \geq 1 - \varepsilon$, and if $|\psi\rangle$ is any $N'$-qubit state such that

$$\langle \psi | H_V | \psi \rangle < \lambda_1(H_V) + \varepsilon,$$

then for $\xi := R_V(|\psi\rangle\langle\psi|)$ we have

$$\mathbb{E}[V(\xi)] \geq 1 - c \cdot T^{c \cdot \varepsilon^{1/\varepsilon}},$$

where $c > 1$ is an absolute constant. Furthermore, $H_V$ and $R_V$ can be constructed in time $\poly(T)$, given a description of $V$.

Theorems 22 and 23 can be similarly strengthened, so that their guarantees hold for nearly minimal energy states of the local Hamiltonian as well as for ground states. The dependence of the output Hamiltonian upon the choice of error parameters appears necessary in these results, however.

Similarly to Kitaev’s work, it turns out to be convenient to first prove a weakened form of Theorem 24 in which the Hamiltonian is only required to be $O(\log T)$-local. This forms the bulk of our work in this section. It will then be a simple step to reduce the locality to 5.

6.1. The $O(\log T)$-local reduction.

6.1.1. The Hamiltonian. Say that $V$, which expects a proof state $\xi$ on $m$ qubits, acts upon the proof register containing $\xi$ and an $(N-m)$-qubit ancilla register, initialized to the all-zero state, by the sequence $U_1, \ldots, U_T$ of unitary transformations, each of which is 2-local. Here we may assume (by padding, if necessary) that $T + 1$ is a power of $2$. The transformation performed by $V$, applied to a pure input state $|\psi\rangle\langle\psi|$, produces the state $U_T \ldots U_1 \cdot (|\psi\rangle \otimes |0^{N-m}\rangle)$.

Afterward, we assume that the first qubit is measured in the standard basis; $V$ outputs the measured value. We use $V$ to define a Hamiltonian $H = H_V$ acting on $N' := N + D$ qubits, where $D := \log_2(T + 1)$, as follows. We speak of the first $N$ qubits (consisting of the proof and ancilla registers) jointly as the “circuit register,” and the last $N$ qubits as a “clock register.” The local unitaries $U_1, \ldots, U_T$ will be regarded as operators on the Hilbert space of the circuit register. We identify the computational basis states of the clock register with the integers $\{0, 1, \ldots, T\}$, and we write these basis states as $|t\rangle$ for $0 \leq t \leq T$.

To specify projective operators acting on the circuit register, we use the notation $|b\rangle\langle b|$ for $b \in \{0, 1\}$, $i \in [N]$ to denote the projection onto the subspace spanned by all computational basis vectors whose $i$th coordinate is $b$. Formally,

$$|b\rangle\langle b| := I_{i-1} \otimes |b\rangle \otimes I_{N-i}.$$

We define a Hamiltonian operator $H = H_V$ having three terms, $H_{\text{in}}, H_{\text{out}},$ and $H_{\text{prop}}$. For our analysis we will depart slightly from [25] in our definitions; however, each of the three terms will be a positive scalar multiple of the corresponding term in [25]. We define

$$H := H_{\text{in}} + H_{\text{out}} + H_{\text{prop}},$$

(1)
where

\begin{equation}
H_{\text{in}} := \frac{1}{2} \sum_{i=m+1}^{N} |1\rangle_{i} \langle 1| \otimes |0\rangle \langle 0|
\end{equation}

(here the rightmost projector $|0\rangle \langle 0|$ is onto the basis vector $|t = 0\rangle$ for the clock register),

\begin{equation}
H_{\text{out}} := \frac{1}{2} |0\rangle \langle 0|_1 \otimes |T\rangle \langle T|
\end{equation}

and

\begin{equation}
H_{\text{prop}} := \sum_{t=1}^{T} H_{\text{prop},t},
\end{equation}

where the operators $H_{\text{prop},t}$ are defined for $t \in [T]$ by

\begin{equation}
H_{\text{prop},t} := \frac{1}{2} \left( I_N \otimes |t\rangle \langle t| + I_N \otimes |t-1\rangle \langle t-1| - U_t \otimes |t\rangle \langle t-1| - U_t^\dagger \otimes |t-1\rangle \langle t| \right).
\end{equation}

Note immediately that the operator norms of the individual $O(\log T)$-local terms of $H$ are each $\Theta(1)$.

One can verify that $H_{\text{prop},t}$ is Hermitian. More strongly, $H_{\text{in}}, H_{\text{out}},$ and the terms $H_{\text{prop},t}$ are all positive semidefinite (PSD). For the first two this is obvious: $H_{\text{in}}, H_{\text{out}}$ are orthogonal projectors. To see that $H_{\text{prop},t}$ is PSD, it is clearly enough to show that $\langle w|H_{\text{prop},t}|w\rangle \geq 0$ for any $|w\rangle$ of form $|w\rangle = |w_{t-1}\rangle \otimes |t-1\rangle + |w_t\rangle \otimes |t\rangle$. We compute

\begin{align*}
2 \cdot \langle w| H_{\text{prop},t} |w\rangle &= \langle w_t|w_t\rangle + \langle w_{t-1}|w_{t-1}\rangle - \langle w_t|U_t|w_{t-1}\rangle - \langle w_{t-1}|U_t^\dagger|w_t\rangle \\
&= ||w||^2 - \langle w_t|U_t|w_{t-1}\rangle - \langle w_{t-1}|U_t^\dagger|w_t\rangle \\
&= ||w||^2 - 2 ||w_t|| \cdot ||U_t|w_{t-1}|| \\
&\geq ||w||^2 - 2 ||w_t|| \cdot ||U_t|w_{t-1}|| \\
&\geq ||w||^2 - 2 ||w_t||^2 - ||w_{t-1}||^2 \\
&= 0,
\end{align*}

as needed. Thus $H$, a sum of PSD operators, is itself PSD (and $\lambda_1(H) \geq 0$). This will be important for our analysis.

6.1.2. The transformation of quantum states. For our transformation $R = R_T$ of quantum states as in Theorem 23, we use the operation which first measures the clock register, observing some value $t \in [0, T]$, and then applies $U_1^\dagger \ldots U_T^\dagger$ to the circuit register, outputting the resulting $m$-qubit reduced state on the proof register alone (eliminating the ancilla and clock registers). This transformation is implementable in size $\text{poly}(T)$, since an inverse unitary operation $U^\dagger$ is $k$-local whenever $U$ is $k$-local.

6.1.3. Objective of the analysis. It is shown in [26, 7] that, if $\max_{\rho} E[V(\rho)] \geq 1 - \varepsilon$, then the minimal eigenvalue $\lambda_1(H)$ is at most $O(\varepsilon)$. (This fact is unaffected by our scalar-multiple adjustments to the definitions of $H_{\text{in}}, H_{\text{out}}, H_{\text{prop}}$.) In our analysis,
we will assume that \( \lambda_1(H) < .01 \delta / T \), where \( \delta > 0 \) will be defined as a sufficiently small inverse-polynomial in \( T \). This smallness assumption is without loss of generality, since our sought-after bound in Theorem 24 allows a poly(\( T \)) slack factor. We will then show that if \(| \psi \rangle\) is any state satisfying \( \langle \psi | H | \psi \rangle < .02 \delta / T \), the \( m \)-qubit (mixed) state \( \xi := R(| \psi \rangle \langle \psi |) \) satisfies \( E[V(\xi)] \geq 1 - \delta^{O(1)} \). This suffices to prove the weakened version of Theorem 24 in which \( H \) is only required to be \( O(\log T) \)-local.

6.2. Describing the action of \( H \) on a state. Here we introduce notation and derive some useful expressions which describe the action of \( H \) on an arbitrary pure state.

Consider an \( (N + \log_2(T + 1)) \)-qubit state \(| \psi \rangle\), given by

\[
| \psi \rangle = \sum_{y \in \{0,1\}^N, t \in \{0,1,...,T\}} \alpha_{y,t} | y \rangle \otimes | t \rangle,
\]

with \( \sum_{y,t} | \alpha_{y,t} |^2 = 1 \). We may write

\[
| \psi \rangle = \sum_{t \in \{0,1,...,T\}} \psi_t \otimes | t \rangle,
\]

where

\[
| \psi_t \rangle := \sum_{y \in \{0,1\}^N} \alpha_{y,t} | y \rangle
\]

is a state on the circuit register. Note that \(| \psi_t \rangle\) is not in general a unit vector; we have \( \sum_t | | \psi_t \rangle |^2 = 1 \). We define vectors \(| \xi_0 \rangle, \ldots, | \xi_T \rangle\) by the relation

\[
H | \psi \rangle = \sum_{t=0}^{T} | \xi_t \rangle \otimes | t \rangle,
\]

noting that the \(| \xi_t \rangle\) will also not in general be unit vectors (nor will \( H | \psi \rangle \)).

Now for \( t \in [0, T] \) define

\[
| \phi_t \rangle := U_1^† U_2^† \ldots U_t^† | \psi_t \rangle,
\]

so that

\[
| \psi_t \rangle = U_t U_{t-1} \ldots U_1 | \phi_t \rangle.
\]

(Here, \(| \phi_0 \rangle = | \psi_0 \rangle\) and \(| \phi_1 \rangle = U_1^† | \psi_1 \rangle\). Note that \( | | \phi_t \rangle | = | | \psi_t \rangle |\) and

\[
\sum_{t=0}^{T} | | \phi_t \rangle |^2 = \sum_{t} | | \psi_t \rangle |^2 = 1,
\]

as the \( U_t \) are unitary.

With these definitions, we first examine the action of \( H_{\text{prop}} \) on \(| \psi \rangle\). For \( t \in [T] \), the operator \( H_{\text{prop},t} \) acts as

\[
H_{\text{prop},t} | \psi \rangle = \frac{1}{2} \left( | \psi_t \rangle \otimes | t \rangle + | \psi_{t-1} \rangle \otimes | t - 1 \rangle - U_t \psi_{t-1} \otimes | t \rangle - U_t^† | \psi_t \rangle \otimes | t - 1 \rangle \right),
\]
which we can express as

\[ H_{\text{prop},t}(\psi) = \frac{1}{2} \left( U_1 \ldots U_1 (|\phi_t \rangle - |\phi_{t-1} \rangle) \otimes |t \rangle + U_{t-1} \ldots U_1 (|\phi_{t-1} \rangle - |\phi_t \rangle) \otimes |t-1 \rangle \right). \]

Next, observe that \( H_{\text{in}} \) only outputs vectors in the span of the basis vectors with clock-register equal to 0, i.e., in the span of \( \{|y\rangle \otimes |0\rangle \}_y \), and that \( H_{\text{out}} \) outputs vectors in the span of \( \{|y\rangle \otimes |T\rangle \}_y \). Thus for \( t \in [T-1] \), the only contribution of terms of form \( |y\rangle \otimes |t \rangle \) to the output of \( H(\psi) \) comes from \( H_{\text{prop},t} \) and \( H_{\text{prop},t+1} \), and we compute that for such \( t \),

\[ |\xi_t \rangle = U_1 \ldots U_1 (|\phi_t \rangle - .5|\phi_{t-1} \rangle - .5|\phi_{t+1} \rangle). \]

In particular, as \( (U_1 \ldots U_1) \) is unitary we have

\[ |||\xi_t \rangle \otimes |t \rangle|| = |||\xi_t \rangle|| = |||\phi_t \rangle - .5|\phi_{t-1} \rangle - .5|\phi_{t+1} \rangle||. \]

Next we examine the terms in \( H(\psi) \) on clock-value \( t = 0 \), which come solely from the actions of \( H_{\text{in}} \) and \( H_{\text{prop},1} \). Define the orthogonal projector \( \Pi_{\text{in}} \) acting on the \( N \)-qubit circuit register by

\[ \Pi_{\text{in}} := \sum_{i=m+1}^{N} |1\rangle \langle 1|; \]

the operator

\[ \Pi_{\text{in}}' := (I_N - \Pi_{\text{in}}) \]

is also an orthogonal projection. We have

\[ |\xi_0 \rangle = .5|\psi_0 \rangle - .5U_1^\dagger |\psi_1 \rangle + .5\Pi_{\text{in}} |\psi_0 \rangle \]
\[ = .5|\psi_0 \rangle - .5|\phi_1 \rangle + .5\Pi_{\text{in}} |\phi_0 \rangle \]
\[ = |\phi_0 \rangle - .5(|\phi_0 \rangle - \Pi_{\text{in}} |\phi_0 \rangle) - .5|\phi_1 \rangle \]
\[ = |\phi_0 \rangle - .5\Pi_{\text{in}}' |\phi_0 \rangle - .5|\phi_1 \rangle. \]

Thus,

\[ |||\xi_0 \rangle \otimes |0\rangle|| = |||\xi_0 \rangle|| = |||\phi_0 \rangle - .5\Pi_{\text{in}}' |\phi_0 \rangle - .5|\phi_1 \rangle||. \]

Finally, we examine the terms in \( H(\psi) \) on clock-value \( T \), which come solely from the actions of \( H_{\text{out}} \) and \( H_{\text{prop},T} \). Define the projector \( \Pi_{\text{out}} \) acting on the circuit register by \( \Pi_{\text{out}} := |0\rangle \langle 0|_1 \); define the operators

\[ \Phi_{\text{out}} := U_1^\dagger \ldots U_T^\dagger \Pi_{\text{out}} U_T \ldots U_1 \]

and

\[ \Phi_{\text{out}}' := I_N - \Phi_{\text{out}}. \]
acting on $N$ qubits. Then we have

\begin{align}
(12) \quad |\xi_T\rangle &= 0.5|\psi_T\rangle - 0.5U_T|\psi_{T-1}\rangle + 0.5\Pi_{\text{out}}|\psi_T\rangle \\
(13) \quad &= |\psi_T\rangle - 0.5U_T|\psi_{T-1}\rangle - 0.5|\phi_T\rangle + \Pi_{\text{out}}|\psi_T\rangle \\
(14) \quad &= U_T \ldots U_1 \left( |\phi_T\rangle - 0.5|\phi_{T-1}\rangle - 0.5|\phi_T\rangle \right) + 0.5\Pi_{\text{out}}|\psi_T\rangle \\
(15) \quad &= U_T \ldots U_1 \left( |\phi_T\rangle - 0.5|\phi_{T-1}\rangle - 0.5|\phi_T\rangle \right) + 0.5U_T \ldots U_1 \Pi_{\text{out}}U_T \ldots U_1 |\phi_T\rangle \\
(16) \quad &= U_T \ldots U_1 \left( |\phi_T\rangle - 0.5|\phi_{T-1}\rangle - 0.5|\phi_T\rangle + 0.5U_T \ldots U_1 \Pi_{\text{out}}U_T \ldots U_1 |\phi_T\rangle \right) \\
(17) \quad &= U_T \ldots U_1 \left( |\phi_T\rangle - 0.5|\phi_{T-1}\rangle - 0.5\Phi'_{\text{out}}|\phi_T\rangle \right).
\end{align}

Thus,

\begin{equation}
(18) \quad |||\xi_T\rangle \otimes |T\rangle|| = |||\xi_T\rangle|| = |||\phi_T\rangle - 0.5\Phi'_{\text{out}}|\phi_T\rangle - 0.5|\phi_{T-1}\rangle||.
\end{equation}

### 6.3. Analyzing low-energy states of $H$. Here we argue that if $|\psi\rangle$ is any state for which the energy $\langle \psi | H | \psi \rangle$ is sufficiently small, then our operation $R = R_V$, when applied to $|\psi\rangle\langle\psi|$, produces a state accepted with high probability by $V$. No corresponding result is needed or established in Kitaev’s original work [26], which analyzed the minimal eigenvalue of $H$, but not the structure of ground states themselves. Subsequent works, including [25, 29], have provided more detailed information about the low-energy subspaces of several local-Hamiltonian reductions (although these works do not immediately yield the conclusions we seek). We will make crucial use of results from [25, 29] in section 7.

We first describe the idea of our analysis. Suppose $|\psi\rangle$ is any unit vector for which $||H|\psi\rangle||$ is “very small.” We have

$$||H|\psi\rangle||^2 = \sum_i |||\xi_i\rangle \otimes |t\rangle||^2 = \sum_i |||\xi_i\rangle||^2,$$

so each $|\xi_i\rangle$ is a very small vector. If $t \in [T - 1]$, then (10) tells us that $|\phi_t\rangle = U_t^\dagger \ldots U_1^\dagger |\psi_0\rangle$ is nearly equal to the average of $|\phi_{t-1}\rangle$ and $|\phi_{t+1}\rangle$. For $t = 0$, (11) tells us that $|\phi_0\rangle$ is nearly the average of $\Pi_{\text{in}} |\phi_0\rangle$ and $|\phi_1\rangle$; and for $t = T$, (18) tells us that $|\phi_T\rangle$ is nearly the average of $\Phi'_{\text{out}}|\phi_T\rangle$ and $|\phi_{T-1}\rangle$. Thus, the sequence

\begin{equation}
(19) \quad \Pi_{\text{in}} |\phi_0\rangle, \ |\phi_0\rangle, |\phi_1\rangle, \ldots, |\phi_T\rangle, \ \Phi'_{\text{out}}|\phi_T\rangle
\end{equation}

is very nearly an arithmetic progression within the $N$-qubit Hilbert space of the circuit register.

Now there are essentially two possibilities. In the first, “good” case, the terms in this near-arithmetic progression are all nearly equal to $|\phi_0\rangle$, so that each $|\psi_t\rangle$ is nearly equal to $U_t \ldots U_t^\dagger |\phi_0\rangle$. Inspecting the definitions of $\Pi_{\text{in}}$ and $\Phi'_{\text{out}}$, we then find that $\Pi_{\text{in}} |\psi_0\rangle$ and $\Phi'_{\text{out}}|\psi_T\rangle$ are both $\approx 0$. This implies that $|\psi_0\rangle$, after normalization, is close to a legal input state (i.e., with the ancilla register in the all-zero state) causing the verifier $V$ to accept with high probability. Moreover, we may obtain a near-perfect copy of $|\psi_0\rangle$ by the operation $R_V$ defined earlier.

In the second, “bad” case, our near-arithmetic progression has some nontrivial step size, and its terms are close to being $(T + 3)$ equally spaced points along some
line in Hilbert space. Now in such an arrangement, it is an intuitive fact that the furthest of these points from the origin will be either the first or the last point along the line. Thus, either \( \Pi_{in} |\phi_0\rangle \) or \( \Phi'_\text{out} |\phi_T\rangle \) will have the largest norm from among the vectors in our sequence. However, one easily verifies that \( \Pi_{in} \) and \( \Phi'_\text{out} \) each have operator norm at most 1, so that \( ||\Pi_{in} |\phi_0\rangle|| \leq |||\phi_0||| \) and \( ||\Phi'_\text{out} |\phi_T\rangle|| \leq |||\phi_T||| \). So in fact the bad case cannot occur.

With this informal sketch in mind, we begin. Fix any \( \delta > 0 \) satisfying

\[
\delta < \frac{1}{8^8(T + 3)^{18}}.
\]

As discussed in section 6.1, we will assume that there is some unit vector \( |\psi\rangle = \sum_{t=0}^T |\psi_t\rangle \otimes |t\rangle \) such that

\[
\langle \psi | H | \psi \rangle \leq .02 \delta / T,
\]

and we will show that \( \mathbb{E}[V(\xi)] \geq 1 - \delta^{\Omega(1)} \), where \( \xi := R(\langle \psi | \psi \rangle) \). First, we claim that the vector \( H |\psi\rangle \) has small norm. To see this, first use the spectral theorem to write

\[
H = \sum_{\ell \in [2^N]} \lambda_\ell |\ell\rangle \langle \ell |,
\]

where \( \{|\ell\rangle\} \) is an orthonormal eigenbasis for \( H \) and \( \{\lambda_\ell = \lambda_\ell(H)\} \) are the corresponding eigenvalues. We have \( 0 \leq \lambda_1 \leq \cdots \leq \lambda_{2^N} \). Write \( |\psi\rangle = \sum_{\ell \in [2^N]} \beta_\ell |\ell\rangle \) with \( \beta_\ell \in \mathbb{C} \) and \( \sum_{\ell \in [2^N]} |\beta_\ell|^2 = 1 \). We have the expressions

\[
H |\psi\rangle = \sum_{\ell \in [2^N]} \beta_\ell \lambda_\ell |\ell\rangle, \quad ||H |\psi\rangle||^2 = \sum_{\ell \in [2^N]} |\beta_\ell|^2 \lambda_\ell^2, \quad \langle \psi | H |\psi\rangle = \sum_{\ell \in [2^N]} |\beta_\ell|^2 \lambda_\ell.
\]

Thus \( \|H |\psi\rangle\|^2 \leq \lambda_{2^N} \cdot \langle \psi | H |\psi\rangle = ||H|| \cdot \langle \psi | H |\psi\rangle \). We have the crude operator-norm bound \( ||H|| \leq 10T \), which follows by summing bounds on the norms of each term of \( H \). Thus,

\[
||H |\psi\rangle||^2 = \sum_{t=0}^T ||\xi_t||^2 \leq \delta,
\]

where we again define \( \{||\xi_t||\}_t \) by the relation \( H |\psi\rangle = \sum_t |\xi_t\rangle \otimes |t\rangle \).

For each \( t \in \{0, 1, \ldots, T\} \), let \( \delta_t := ||\xi_t||^2 \). Define

\[
|\Delta_0\rangle := |\phi_0\rangle - \Pi_{in}' |\phi_0\rangle
\]

as the difference between the first two terms in the sequence from (19). For \( t \in [T] \), define

\[
|\Delta_t\rangle := |\phi_t\rangle - |\phi_{t-1}\rangle,
\]

and define

\[
|\Delta_{T+1}\rangle := \Phi'_\text{out} |\phi_T\rangle - |\phi_T\rangle.
\]

By (11), we have

\[
\delta_0 = |||\xi_0|||^2
= ||.5(|\phi_0\rangle - \Pi_{in}' |\phi_0\rangle) - .5(|\phi_1\rangle - |\phi_0\rangle)||^2
= .25|||\Delta_0\rangle - |\Delta_1\rangle||^2.
\]
Similarly, by (10), for \( t \in [T - 1] \) we have
\[
\delta_t = |||\xi_t|||^2 \\
= ||.5(\phi_t - |\phi_{t-1}\rangle) - .5(|\phi_{t+1}\rangle - |\phi_t\rangle)||^2 \\
= .25|||\Delta_t)||^2 - |||\Delta_{t+1}|||^2.
\]
(22)

Finally, by (18) we have
\[
\delta_T = |||\xi_T|||^2 \\
= ||.5(\phi_T - |\phi_{T-1}\rangle) - .5(\Phi'_\text{out}(\phi_T) - |\phi_T\rangle)||^2 \\
= .25|||\Delta_T)||^2 - |||\Delta_{T+1}|||^2.
\]
(23)

Combining our work, we find that for each \( t \in \{0, 1, \ldots, T\} \) we have
\[
|||\Delta_{T+1} - |\Delta_T\rangle|| = 2\sqrt{\delta_t}.
\]
(24)

At this point, for notational convenience we define
\[
|\phi_{-1}\rangle := \Pi'_\text{in}|\phi_0\rangle, \quad |\phi_{T+1}\rangle := \Phi'_\text{out}|\phi_T\rangle.
\]
From the definitions of \( \Pi'_\text{in}, \Phi'_\text{out} \) one can verify that their operator norms are each at most 1, so that
\[
|||\phi_{-1}\rangle|| \leq |||\phi_0\rangle|| \leq |||\psi\rangle|| = 1
\]
and
\[
|||\phi_{T+1}\rangle|| \leq |||\phi_T\rangle|| \leq |||\psi\rangle|| = 1.
\]

By our definitions, for each \( t \in [T + 1] \) we have
\[
|\phi_t\rangle = |\phi_{-1}\rangle + \sum_{t'=0}^{t} |\Delta_{t'}\rangle \\
= |\phi_{-1}\rangle + \sum_{t'=0}^{t} \left( |\Delta_0\rangle + \sum_{t''=0}^{t'-1} (|\Delta_{t''+1}\rangle - |\Delta_{t''}\rangle) \right) \\
= |\phi_{-1}\rangle + (t + 1)|\Delta_0\rangle + \sum_{s=0}^{t-1} (t - s) \cdot (|\Delta_{s+1}\rangle - |\Delta_s\rangle).
\]

Using the triangle inequality and (24), we find that for \( t \in [T + 1] \),
\[
|||\phi_t\rangle - (|\phi_{-1}\rangle + (t + 1)|\Delta_0\rangle)|| = \left| \sum_{s=0}^{t-1} (t - s) \cdot (|\Delta_s\rangle - |\Delta_{s-1}\rangle) \right| \\
\leq 2 \cdot \sum_{s=0}^{t-1} (t - s) \sqrt{\delta_s}.
\]
(25)

In particular, this implies that
\[
|||\phi_T||| \leq |||\phi_{-1}||| + (T + 1)|\Delta_0||| + 2 \cdot \sum_{s=0}^{T-1} (T - s) \sqrt{\delta_s}
\]
and

\[ |||\phi_{t+1}||| \geq |||\phi_{-1}|| + (T + 2)||\Delta_0|| - 2 \cdot \sum_{s=0}^{T} (T - s + 1) \sqrt{\delta_s}. \]

To understand these bounds, consider the linear function \( \ell : \mathbb{R}^{T+1} \to \mathbb{R} \) given by

\[ \ell(x_0, \ldots, x_T) := \sum_{s=0}^{T} (T - s + 1)x_s. \]

It is a standard fact that the maximum value of \( \ell \) in the disk \( B_{0,r} = \{ \mathbf{x} \in \mathbb{R}^{T+1} : \sum_s x_s^2 \leq r^2 \} \) is attained at the point

\[ \mathbf{x}^* = (x_0^*, \ldots, x_t^*) := r \cdot \frac{\nabla \ell}{||\nabla \ell||}, \]

where the gradient function \( \nabla \ell \) is defined as

\[ \nabla \ell := \left( \frac{\partial \ell}{\partial x_0}, \ldots, \frac{\partial \ell}{\partial x_T} \right). \]

In our case, the gradient is the constant vector \( \nabla \ell = (T, T, T - 1, \ldots, 1) \).

Now recall that \( \sum_{s=0}^{T} \delta_s \leq \delta \). It follows that

\[ \sum_{s=0}^{T} (T - s + 1) \sqrt{\delta_s} = \ell(\sqrt{\delta_0}, \sqrt{\delta_1}, \ldots, \sqrt{\delta_T}) \]

\[ \leq \ell \left( \sqrt{\delta} \cdot \frac{\nabla \ell}{||\nabla \ell||} \right) \]

\[ = \frac{\sqrt{\delta}}{||\nabla \ell||} \cdot \ell(\nabla \ell) \]

\[ = \frac{\sqrt{\delta}}{\sqrt{\sum_{s=0}^{T} (T - s + 1)^2}} \cdot \left( \sum_{s=0}^{T} (T - s + 1)^2 \right) \]

\[ = \delta \cdot \sum_{s=0}^{T} (T - s + 1)^2 \]

\[ = \frac{\delta(T + 2)(T + 3)(2T + 3)}{6} \]

\[ < \sqrt{\frac{\delta(T + 3)^3}{3}}. \]

Combining this with (25), we find that for \( t \in [T + 1] \),

\[ |||\phi_t|| - (||\phi_{-1}|| + (t + 1)||\Delta_0||)\| \leq 2 \sqrt{\frac{\delta(T + 3)^3}{3}}. \]

In particular, we have

\[ |||\phi_T||| \leq |||\phi_{-1}|| + (T + 1)||\Delta_0|| + 2 \sqrt{\frac{\delta(T + 3)^3}{3}} \]
Next, we claim that the quantity
\[
Q_{\text{ext}} := |||\phi_{-1}|||^2 + |||\phi_{-1} + (T + 2)|\Delta_0|||^2
\]
is slightly larger than
\[
Q_{\text{int}} := |||\phi_0|||^2 + |||\phi_{-1} + (T + 1)|\Delta_0|||^2
\]
if \(|\Delta_0|| is of noticeable size. This will be a useful way to quantify our intuition that the largest point in an arithmetic progression should be one of the endpoints. Recall that \(|\phi_0|| = |\phi_{-1} + |\Delta_0||. We have
\[
Q_{\text{int}} = \left(\langle \phi_{-1} | \phi_{-1} \rangle + \langle \Delta_0 | \Delta_0 \rangle + \langle \phi_{-1} | \Delta_0 \rangle + \langle \phi_{-1} | \Delta_0 \rangle \right)
- \left(\langle \phi_{-1} | \phi_{-1} \rangle + (T + 1)^2 \langle \Delta_0 | \Delta_0 \rangle + (T + 1) \left(\langle \phi_{-1} | \Delta_0 \rangle + \langle \phi_{-1} | \Delta_0 \rangle \right) \right)
= 2|||\phi_{-1}|||^2 + (T^2 + 2T + 2)|||\Delta_0|||^2 + (T + 2) \left(\langle \phi_{-1} | \Delta_0 \rangle + \langle \phi_{-1} | \Delta_0 \rangle \right).
\]
By a similar calculation,
\[
Q_{\text{ext}} = 2|||\phi_{-1}|||^2 + (T^2 + 2T + 4)|||\Delta_0|||^2 + (T + 2) \left(\langle \phi_{-1} | \Delta_0 \rangle + \langle \phi_{-1} | \Delta_0 \rangle \right),
\]
so that
\[
Q_{\text{ext}} - Q_{\text{int}} = 2 \cdot |||\Delta_0|||^2.
\]
We next define
\[
Q'_{\text{ext}} := |||\phi_{-1}|||^2 + |||\phi_{T+1}|||^2
\]
and
\[
Q'_{\text{int}} := |||\phi_0|||^2 + |||\phi_T|||^2.
\]
Using (29), we have
\[
Q_{\text{ext}} - Q'_{\text{ext}} = |||\phi_{-1} + (T + 2)|\Delta_0|||^2 - |||\phi_{T+1}|||^2
\leq \left(|||\phi_{T+1}||| + 2\sqrt{\frac{\delta(T + 3)^3}{3}}\right)^2 - |||\phi_{T+1}|||^2
\leq \frac{4\delta(T + 3)^3}{3} + 4\sqrt{\frac{\delta(T + 3)^3}{3}}
\leq 8\sqrt{\frac{\delta(T + 3)^3}{3}},
\]
where in the last two steps we used the fact that \(|||\phi_{T+1}||| \leq 1\) and our smallness assumption on \(\delta.\)
Similarly, using (28),

\[
Q'_{\text{int}} - Q_{\text{int}} = |||\phi_T|||^2 - |||\phi_{t-1}|||^2 + (T + 1)|\Delta_0|^2 \\
\leq |||\phi_T|||^2 - \left(|||\phi_T||| - 2\sqrt{3}\frac{\delta(T + 3)^3}{3}\right)^2 \\
\leq 4\sqrt{\frac{3}{\delta(T + 3)^3}},
\]

where in the last step we used that \(|||\phi_T||| \leq 1.\)

Combining (30), (31), and (32), we compute that

\[
Q'_{\text{ext}} - Q'_{\text{int}} = (Q'_{\text{ext}} - Q_{\text{ext}}) + (Q_{\text{ext}} - Q_{\text{int}}) + (Q_{\text{int}} - Q'_{\text{int}}) \\
\geq -8\sqrt{\frac{3}{\delta(T + 3)^3}} + 2||\Delta_0||^2 - 4\sqrt{\frac{3}{\delta(T + 3)^3}} \\
= 2\left(||\Delta_0||^2 - \sqrt{12\delta(T + 3)^3}\right).
\]

On the other hand, recall that \(|||\phi_{t-1}||| \leq |||\phi_0|||\) and \(|||\phi_{T+1}||| \leq |||\phi_T|||\). Thus, \(Q'_{\text{ext}} - Q'_{\text{int}} \leq 0.\) With (33), this implies that

\[
|||\Delta_0||| \leq (12\delta)^{1/4} (T + 3)^{3/4}.
\]

Informally, this tells us that \(|\Delta_0\) is small, so that we are not in the bad case described earlier.

Now, (27) and (34) combine to show us that \(|\phi_1\), \(|\phi_2\), ..., \(|\phi_{T+1}\) are all close to \(|\phi_{t-1}\): for \(t \in [T + 1],\)

\[
|||\phi_t - |\phi_{t-1}||| \leq |||\phi_t - (|\phi_{t-1} + (t + 1)|\Delta_0))|| + |||(\phi_{t-1} + (t + 1)|\Delta_0)) - |\phi_{t-1}||| \\
\leq 2\sqrt{\frac{3}{\delta(T + 3)^3}} + (t + 1)||\Delta_0|| \\
\leq 2\sqrt{\frac{3}{\delta(T + 3)^3}} + (12\delta)^{1/4} (T + 3)^{7/4} \\
\leq 4\delta^{1/4} (T + 3)^{7/4},
\]

using our smallness assumption on \(\delta\) for the last step. This also implies that

\[
|||\phi_{T+1} - |\phi_T||| \leq 8\delta^{1/4}(T + 3)^{7/4}.
\]

Also, using (34) again, we have

\[
|||\phi_0 - |\phi_{t-1}||| = |||\Delta_0||| \leq 2\delta^{1/4}(T + 3)^{3/4}.
\]

Next we argue that for each \(t \in [0, T]\), the vector \(|\phi_t\) has norm close to \(T^{-1/2}.\) Recall that \(|||\phi_t||| = |||\psi_t|||\) for \(t \in [0, T],\) and that \(\sum_{t=0}^T |||\psi_t|||^2 = 1.\) Thus there is at least one value \(t^* \in [0, T]\) for which \(|||\phi_{t^*}||| \geq (T + 1)^{-1/2}.\) Then, using (35), for any \(t \in [0, T + 1]\) we have

\[
|||\phi_t||| \geq |||\phi_{t^*}||| - |||\phi_{t^*} - |\phi_{t-1}||| - |||\phi_{t-1} - |\phi_{t-1}||| \\
\geq (T + 1)^{-1/2} - 8\delta^{1/4}(T + 3)^{7/4} \\
\geq (1 - \delta^{1/4})(T + 1)^{-1/2},
\]
where in the last step we again used our smallness assumption on $\delta$. Similarly, using (35) and (37) we obtain

$$|||\phi_{-1}||| \geq (1 - \delta^{1/8})(T + 1)^{-1/2}.$$  

On the other hand, there is also a $t_\ast \in [0, T]$ for which $|||\phi_{t_\ast}||| \leq (T + 1)^{-1/2}$. By modifying the above arguments only slightly, we find that for each $t \in [-1, T]$,

$$(39) \quad |||\phi_t||| \leq (1 + \delta^{1/8})(T + 1)^{-1/2}.$$  

For $t \in [-1, T + 1]$, define the normalized vector

$$|\phi_t\rangle := \frac{|\phi_t\rangle}{|||\phi_t|||}.$$  

Also, for $t \in [0, T]$, similarly define

$$|\psi_t\rangle := \frac{|\psi_t\rangle}{|||\psi_t|||}.$$  

Next, we define $\gamma_t$ by the relation

$$|\phi_t\rangle = (1 + \gamma_t)\sqrt{T + 1} : |\phi_t\rangle;$$  

by (38)–(39), we have $\gamma_t \in [-\delta^{1/8}, +\delta^{1/8}]$. Then for $t \in [0, T + 1]$ we have

$$|||\phi_t\rangle - |\phi_{-1}\rangle||| \leq \sqrt{T + 1} \cdot (|||\phi_t\rangle - |\phi_{-1}\rangle||| + |\gamma_t| \cdot |||\phi_t||| + |\gamma_{-1}| \cdot |||\phi_{-1}|||)$$

$$\leq \sqrt{T + 1} \cdot (4\delta^{1/4}(T + 3)^{7/4} + 2\delta^{1/8}(1 + \delta^{1/8})(T + 1)^{-1/2})$$

$$\leq 4\delta^{1/4}(T + 3)^{9/4} + 3\delta^{1/8}$$

(40)

This in particular implies

$$(41) \quad 4\delta^{1/8} \geq |||U_T \ldots U_1(|\phi_T\rangle - |\phi_{-1}\rangle)||| = |||\hat{\psi}_T - U_T \ldots U_1 \left(\Pi_{in}^\dagger |\psi_0\rangle / |||\Pi_{in}|\psi_0|||\right)|||.$$  

Next, expanding the definitions of terms in (36), we find that

$$8\delta^{1/4}(T + 3)^{7/4} \geq |||\Phi_{out}'|\phi_T\rangle - |\phi_T\rangle|||$$

$$= |||\Phi_{out}'|\phi_T\rangle|||$$

$$= |||U_T^1 \ldots U_T^2 |\Pi_{out} U_T \ldots U_1 (U_T^1 \ldots U_T^2 \hat{\psi}_T)\rangle|||$$

$$= |||U_T^1 \ldots U_T^2 |\Pi_{out} \hat{\psi}_T\rangle|||$$

$$= |||\Pi_{out} \hat{\psi}_T\rangle|||.$$  

(42)

Now (38), applied with $t := T$, tells us that $|||\hat{\psi}_T\rangle||| = |||\phi_T\rangle||| \geq (1 - \delta^{1/8})(T + 1)^{-1/2}$. Combining this with (42), we have

$$(43) \quad |||\Pi_{out} |\psi_T\rangle\rangle||| \leq 2\sqrt{T + 1} \cdot 8\delta^{1/4}(T + 3)^{7/4} \leq 16\delta^{1/4}(T + 3)^{9/4}.$$  

\( \Pi_{\text{out}} \) is an orthogonal projection and has operator norm 1. This fact, combined with (41) and (43), allows us to infer that
\[
\left\| \Pi_{\text{out}} \left( U_T \ldots U_1 \left( \frac{\Pi'_{\text{in}}}{\| \Pi'_{\text{in}} \|} \right) \right) \right\| \leq 4\delta^{1/8} + 16\delta^{1/4}(T + 3)^{9/4} \leq 8\delta^{1/8}.
\]

This shows that the quantum state \( \left\| \phi_{-1} \right\| = \left( \frac{\Pi'_{\text{in}}}{\| \Pi'_{\text{in}} \|} \right) \), when set as the initial state of the circuit register of the verifier \( V \), causes \( V \) to accept with probability \( \geq 1 - \delta^{O(1)} \).
Also, \( \left\| \phi_{-1} \right\| \) lies in the kernel of the orthogonal projector \( \Pi_{\text{in}} = I_N - \Pi'_{\text{in}} \), so its final \( N - m \) qubits are in the all-zero state.

Finally, we claim that being given the state \( \left\| \psi \right\| \) allows us to recover a close approximation to \( \left\| \phi_{-1} \right\| \) by applying our quantum operation \( R = R_V \). This procedure first measures the clock register. If \( t \in [T] \) is observed, the post-measurement circuit register state is \( \left\| \psi_t \right\| ; \) the transformation
\[
\left\| \psi_t \right\| \rightarrow U_t^\dagger \ldots U_1^\dagger \left\| \psi_t \right\| = \left\| \tilde{\phi}_t \right\|
\]
is then performed. (If the value \( t = 0 \), the post-measurement state on the circuit register is \( \left\| \psi_0 \right\| = \left\| \tilde{\phi}_0 \right\| \) and \( R \) applies none of these unitaries.) Equation (40) tells us that the resulting state \( \left\| \tilde{\phi}_t \right\| \) on the circuit register is \( 4\delta^{1/8} \)-close to the desired state \( \left\| \phi_{-1} \right\| \). Thus, the reduced state of \( \left\| \tilde{\phi}_t \right\| \) on the \( m \)-qubit proof register (which \( R \) outputs) causes \( V \) to accept with probability \( \geq 1 - \delta^{O(1)} \). We have established the variant of Theorem 24 which requires \( H \) only to be \( O(\log T) \)-local.

### 6.4. Reduction to locality 5.
Following [26, 7], we now describe a small alteration of the above \( O(\log T) \)-local reduction that produces a 5-local Hamiltonian.

#### 6.4.1. The modified reduction.

The Hilbert space used still consists of an \( N \)-qubit along with a clock register. This time, however, the clock register consists of \( T \) qubits; informally, its “intended purpose” is to store a time-index \( t \in [0, T] \) by the unary encoding \( |1^{T-t} \rangle \). A clock-register basis state of this form is called valid; basis states not of this form are said to be invalid and will be penalized by our Hamiltonian. For \( t \in [0, T - 2] \) and bits \( a, b, c, a', b', c' \), we let
\[ |a'b'c'\rangle \left\langle abc \right|_{\text{clk}(t)} \]
denote the 3-local operator \( |a'b'c'\rangle \left\langle abc \right|_{\text{clk}(t)} \) applied to the \( t \)-th, \( (t+1) \)-th, and \( (t+2) \)-th clock register qubits. Similarly, \( |a'b'\rangle \left\langle ab \right|_{\text{clk}(t)} \) denotes \( |a'b'\rangle \left\langle ab \right| \) applied to the \( t \)-th and \( (t+1) \)-th clock qubits.

We modify the Hamiltonian \( H = H_V : \mathcal{B}^{\otimes(N+D)} \rightarrow \mathcal{B}^{\otimes(N+D)} \) from our previous work to produce a new Hamiltonian \( H' = H'_V \) acting on the new Hilbert space \( \mathcal{B}^{\otimes(N+T)} \). First, in each tensor term appearing in \( H_{\text{in}}, H_{\text{out}} \), we replace the clock-register projectors \( |0\rangle\langle 0|, |T\rangle\langle T| \) with \( |00\rangle\langle 00|_{\text{clk}(0)}, |11\rangle\langle 11|_{\text{clk}(T-1)} \), respectively, to get modified operators \( H'_{\text{in}}, H'_{\text{out}} \) acting on our new Hilbert space:
\[
H'_{\text{in}} := \frac{1}{2} \sum_{i=m+1}^{N} |i\rangle\langle i|_1 \otimes |00\rangle\langle 00|_{\text{clk}(0)}, \quad H'_{\text{out}} := \frac{1}{2} |00\rangle\langle 00|_1 \otimes |11\rangle\langle 11|_{\text{clk}(T-1)}.
\]

Similarly, we define \( H'_{\text{prop}} := \sum_{i=1}^{T} H'_{\text{prop},t} \) as follows. In each tensor-product term defining \( H_{\text{prop},t} \), if \( t \in [2, T-1] \), then we replace the clock-register projectors
\[
|t\rangle\langle t|, |t-1\rangle\langle t-1|, |t\rangle\langle t-1|, |t-1\rangle\langle t|
\]
with, respectively,

\[ |110⟩⟨110|, |100⟩⟨100|, |110⟩⟨100|, |100⟩⟨110|, |100⟩⟨100|. \]

to obtain \( H'_{\text{prop}} \). Finally, we introduce a new “clock term” \( H'_{\text{clk}} := \sum_{t=1}^{T} I_N \otimes \langle 01 | \langle 01 | \rangle_{\text{clk}(t-1)} \), penalizing invalid clock-register states. We let \( H' := H'_{\text{in}} + H'_{\text{out}} + H'_{\text{prop}} + H'_{\text{clk}} \). The operator norms of the individual 5-local terms of \( H' \) are \( \Theta(1) \), satisfying the norm requirement in Theorem 24’s statement.

The modified quantum operation \( R' \) is defined in close analogy to \( R \) from our previous reduction. The only difference is that when \( R' \) first measures the clock register (now on \( T \) qubits), a measurement outcome \( 10^T t \) is interpreted as the time-index \( t \), and an outcome not of this form is interpreted (arbitrarily) as seeing the time-index \( t = 0 \).

6.4.2. The analysis. Following previous works, we make several observations about \( H' \). First, \( H' \) is PSD by the same argument as for \( H \), and its operator norm still satisfies the crude upper-bound \( \| H' \| \leq 10T \) used previously. Next, define the subspace \( S_{\text{val}} \subseteq B^{\otimes(N+T)} \) as all vectors which place amplitude 0 on invalid clock-register basis states. Note that \( H'(S_{\text{val}}) \subseteq S_{\text{val}} \), and therefore (as \( H' \) is Hermitian) also \( H'(S_{\text{val}}^\perp) \subseteq S_{\text{val}}^\perp \).

Let \( L : S_{\text{val}} \rightarrow B^{\otimes(N+D)} \) be the linear mapping defined on basis states by

\[ L(|x⟩ \otimes 1^t0^T) := |x⟩ \otimes |t⟩ \quad \text{for} \quad x \in \{0,1\}^N, \quad t \in \{0,T\}. \]

Then we observe that for any \( |φ⟩ \in S_{\text{val}} \), we have the relation

\[ H'(|φ⟩) = H(L(|φ⟩)). \]

Moreover, \( L \) is surjective; it follows that \( \lambda_1(H') \leq \lambda_1(H) \). We claim, however, that for any \( |φ⟩ \) in the orthogonal complement \( S_{\text{val}}^\perp \) (consisting of vectors which place zero amplitude on valid clock-register states), we have \( ⟨φ|H'|φ⟩ \geq 1 \). To see this, just note that \( ⟨φ|H'_{\text{clk}}|φ⟩ \geq 1 \), and that \( ⟨φ|(H'_{\text{in}} + H'_{\text{out}} + H'_{\text{prop}})|φ⟩ \geq 0 \) (since each of the three inner summands is PSD). Thus \( S_{\text{val}} \) is spanned by eigenvalues of \( H' \), all of which are \( \geq 1 \).

Following the discussion at the end of section 6.1, let us once more assume that \( \max \mathbb{E}[V(ξ)] \geq 1 - γ \), where \( γ = Θ(\delta/T) \) is sufficiently small that \( \lambda_1(H) < .0001\delta/T \), where \( δ \) is as in (20). Let \( |φ⟩ \in B^{\otimes(N+T)} \) be any unit vector satisfying \( ⟨φ|H'|φ⟩ < .002δ/T \). (Some such \( |φ⟩ \) must exist, since \( \lambda_1(H') \leq \lambda_1(H) \).) Decompose \( |φ⟩ = α|φ⟩_{\text{val}} + β|φ⟩_{\text{invalid}} \) into its components in \( S_{\text{val}}, S_{\text{val}}^\perp \), respectively (where \( |φ⟩_{\text{val}}, |φ⟩_{\text{invalid}} \) are normalized). \( H'|φ⟩_{\text{invalid}} \) is contained in \( S_{\text{val}}^\perp \) and has inner product at least 1 with \( |φ⟩_{\text{invalid}} \), so we must have \( |β|^2 \leq .002δ/T \). Thus, if we define the unit vector \( |φ'⟩ := \frac{α}{|α|}|φ⟩_{\text{val}} \in S_{\text{val}} \), we have

\[ |||φ⟩ - |φ'⟩|| \leq O(\sqrt{δ/T}). \]

\( |φ'⟩ \) also satisfies \( ⟨φ'|H'|φ'⟩ \leq \frac{1}{\sqrt{T}} · ⟨φ|H'|φ⟩ < .02δ/T \). Equation (44) and our analysis of the Hamiltonian \( H \) from previous sections then imply that the state \( ξ := R(L(|φ⟩⟨φ'|)) \) satisfies \( \mathbb{E}[V(ξ)] \geq 1 - δ^{Ω(1)} \). Now observe that, by our definition of \( R' \), the state \( R'(|φ⟩⟨φ'|) \) is identically distributed to \( ξ \) (over the randomness in the measurement of the clock register). Thus \( \mathbb{E}[V(R'(|φ⟩⟨φ'|))] \geq 1 - δ^{Ω(1)} \). Combining this with (45), we conclude that \( \mathbb{E}[V(R'(|φ⟩⟨φ|))] \geq 1 - δ^{Ω(1)} \). This proves Theorem 24.
7. Reduction to 2-local Hamiltonians.

7.1. Goals of the section and proof of Theorem 23. In this section, we complete the proof of Theorem 23. The following definition will be of central importance. Informally speaking, it gives a notion of “witness-preserving reductions” between two problems in QMA, where the “witnesses” here are quantum states. (The precise definition given here is specific to the setting of local Hamiltonian problems.)

**Definition 25.** Let \( k > k' > 1 \) be integers. A \((k,k')\)-approximate ground-space-preserving reduction (AGPR) is a (classical, deterministic) algorithm \( A \) of the following form. \( A \) takes as input a tuple \((H,W,\beta)\), where \( H \) is a description of a \( k \)-local Hamiltonian \( H = \sum_{i \in [s]} H_i \) acting on some number \( n \) of qubits; \( W \geq 1 \) is an integer; and \( \beta \in (0,1) \) is an accuracy parameter. The \( s \geq n \) terms \( H_1, \ldots, H_s \) are each expected to have operator norm \( \|H_i\| \) in the range \([W^{-1}, W] \)—if not, \( A \) may behave arbitrarily. \( A \) runs in time \( \text{poly}(s,W,1/\beta) \) and outputs a pair \((H',R)\), where the following hold:

- \( H' \) is a \( k' \)-local Hamiltonian acting on some number \( n' \leq \text{poly}(s,W,1/\beta) \) of qubits. Each term in the expression for \( H' \) has operator norm in the range \([1/W',W'] \) for some \( W' \leq \text{poly}(s,W,1/\beta) \).
- \( R \) is a quantum operation involving one or more measurements that maps a pure \( n' \)-qubit state \(|\psi\rangle \) to a pure \( n \)-qubit state under every possible set of measurement outcomes. (The resulting pure state depends on the outcomes.) \( R \) is implemented by a quantum circuit of size \( \text{poly}(s,W,1/\beta) \).

Let \( \lambda_1, \lambda_i \in \mathbb{R} \) denote the minimal eigenvalues of \( H, H' \), respectively, the pair \((H',R)\) are required to obey the following property: there is a \( \delta \leq \beta^{\Omega(1)} \cdot \text{poly}(W,s) \) such that, if \(|\psi\rangle \in \mathcal{B}^n \) is any pure state such that

\[ \langle \psi | H' | \psi \rangle < \lambda_1 + \beta, \]

then the state \(|\phi\rangle \) outputted by \( R(|\psi\rangle) \) satisfies

\[ \langle \phi | H | \phi \rangle < \lambda_1 + \delta \]

with probability at least \( 1 - \delta \) over the randomness in \( R \).

We will prove the following.

**Theorem 26.** For each of \( k \in \{5,4,3\} \), there exists a \((k,k-1)\)-AGPR.

In fact, in the reductions we construct are able to take \( \delta \leq O(\beta) \) in Definition 25, although this is not crucial to our work. We defer the proof of Theorem 26 to subsequent sections. AGPRs also compose nicely, as we prove next.

**Lemma 27.** Let \( k > k' > k'' > 1 \) be integers. Suppose there exists a \((k,k')\)-AGPR, call it \( A \), and a \((k',k'')\)-AGPR \( A' \). Then there also exists a \((k,k'')\)-AGPR.

**Proof.** We will compose \( A \) and \( A' \) with suitably chosen parameters. At the outset we note that, by the polynomial slack factor allowed in Definition 25, we may assume that \( \beta \leq \frac{1}{D(s+W)} \) for some fixed constant \( D > 1 \). We will indicate where this assumption is used.

Consider the reduction \( A^* \) which takes as input a \( k \)-local Hamiltonian \( H^{(k)} \) (of \( s \) terms, acting on \( n \) qubits); a bound \( W \) as in the definition; and an \( \beta > 0 \). \( A^* \)}
works as follows. First, we choose \( \gamma := \beta^c \) with \( c > 0 \) a small value to be determined later. We apply our \((k, k')\)-AGPR \( A \) to \((H^{(k)}, W, \gamma)\) to obtain a pair \((H^{(k')}, R)\) each acting on \( n' \leq \text{poly}(s, W, 1/\gamma) \) qubits with \( H^{(k')} \) expressed by \( s' \leq \text{poly}(s, W, 1/\gamma) \) terms. By subdividing terms if necessary, we can assume \( s' \geq n' \). Associated with \( H^{(k')} \) is a second norm-bounding value \( W' \leq \text{poly}(s, W, 1/\gamma) \) as in Definition 25. Let \( \delta \leq \gamma^{\Omega(1)} \cdot \text{poly}(s, W) \) be as in the guarantee for the pair \((H^{(k')}, H^{(k')})\).

Next, we apply our \((k', k'')\)-AGPR \( A' \) to \((H^{(k'')}, W', \beta)\). We get a pair \((H^{(k'')}, R')\) each acting on \( n'' \leq \text{poly}(s', W', 1/\beta) \) qubits. Let \( \delta' \leq \beta^{\Omega(1)} \cdot \text{poly}(s', W') \leq \beta^{R(1)} \cdot \text{poly}(s, W, 1/\gamma) \) be the value in the associated guarantee for the pair \((H^{(k')}, H^{(k'')})\).

We have \( \delta' \leq C(s + W)^C \beta^{1/C} / \gamma^C \) for some constant \( C > 1 \) (independent of our choice for \( \gamma \)). We choose \( \gamma := \beta^{1/(3C^2)} \). It follows that \( \delta' \leq C(s + W)^C \beta^2 / (3C) \).

Now using our aforementioned slack, we require that \( \beta \) is a sufficiently small inverse-polynomial in \((s + W)\) that the above also implies \( \delta' \leq \gamma \).

Our reduction \( A' \) outputs \( H^{(k'')} \) and the composed reduction \( R^* := R \circ R' \), which (by the assumed properties of \( R, R' \)) maps pure \( n'' \) qubit-states to pure \( n'' \)-qubit states, and is implemented by a circuit of size \( \text{poly}(s, W, 1/\beta) \). \( H^{(k'')} \) is \( k'' \)-local as needed and is expressed by \( s'' \leq \text{poly}(s, W, 1/\beta) \) terms whose operator norms are each in \([1/W', W'']\) for some \( W'' \leq \text{poly}(s, W, 1/\beta) \).

Now suppose \(|\psi\rangle \in B^{\otimes n''} \) is any state satisfying \( \langle \psi | H^{(k'')} | \psi \rangle < \lambda_1(H^{(k'')}) + \beta \). Let \( |\phi\rangle := R'(|\phi\rangle) \), where \(|\phi\rangle\) is determined by the measurement outcomes in \( R' \). By the AGPR property of \( R' \), with probability at least \( 1 - \delta' \) over \( R' \) we have \( \langle \phi | H^{(k')} | \phi \rangle < \lambda_1(H^{(k)}) + \delta' \leq \lambda_1(H^{(k)}) + \gamma \). Condition on this event, and let \( |\nu\rangle := R(|\phi\rangle) \).

Then with probability at least \( 1 - \delta \) over \( R \), we have \( \langle \nu | H^{(k)} | \nu \rangle < \lambda_1(H^{(k)}) + \gamma \leq \lambda_1(H^{(k)}) + \beta^{R(1)} \). Thus our reduction \( R^* \) satisfies the desired AGPR guarantee for the value \( \delta^* := \delta + \gamma \leq \beta^{\Omega(1)} \cdot \text{poly}(s, W) \). \( \square \)

Theorem 23 now follows readily from our assembled results.

**Proof of Theorem 23.** Let \( V(\xi) \) be a verifier circuit as in Theorem 23’s statement, and let \( \varepsilon > 0 \) be given such that \( \max_{\xi} E[V(\xi)] \geq 1 - \varepsilon \). We apply Theorem 24 to \( V \) to obtain a 5-local Hamiltonian \( \tilde{H} \) on \( N^* = O(T) \) qubits with \( s \leq \text{poly}(T) \) terms of operator norm in the range \([W^{-1}, W]\) for some \( W \leq \text{poly}(T) \), and a quantum operation \( R \).

Next, it follows from the combination of Theorem 26 and Lemma 27 (applied twice) that there exists a \((5, 2)\)-AGPR \( A \). We apply \( A \) to \((H, W, \beta) \) with \( \beta \geq \varepsilon^{O(1)} / \text{poly}(T) \) a small value to be determined later. We obtain a 2-local Hamiltonian \( \tilde{H} \) and associated quantum operation \( R' \) (both acting on \( B^{\otimes N'} \) for some \( N' \leq \text{poly}(s, 1/\beta) \leq \text{poly}(T, 1/\varepsilon) \)), and a termwise operator norm bound \( W' \leq \text{poly}(T, 1/\varepsilon) \) for \( \tilde{H} \).

For the Hamiltonian \( H_{V, \varepsilon} \), we choose the Hilbert space \( B^{\otimes N'} \) and let \( H_{V, \varepsilon} := \tilde{H} \).

For the operation \( R_{V, \varepsilon} \), we take the composed measurement \( R_{V, \varepsilon} := R \circ R' \). The efficient constructibility claims in Theorem 23 are satisfied for our choice, by the efficiency properties of Theorem 24 and Definition 26 and the requirement \( \beta \geq \varepsilon^{O(1)} / \text{poly}(T) \).

Similarly, the termwise operator-norm bound in Theorem 23 is satisfied.

Now let \(|\psi\rangle \in B^{\otimes N'} \) be any ground state of \( \tilde{H}^* = H_{V, \varepsilon} \). Let \(|\phi\rangle := R'(|\psi\rangle) \in B^{\otimes N'} \) be the pure state determined by the measurement outcomes in \( R' \) applied to \(|\psi\rangle \). By the AGPR property of \( R' \) for some \( \delta \leq \beta^{\Omega(1)} \cdot \text{poly}(T) \), we have \( \text{Pr}_{R'}(|\phi| H|\phi\rangle < \lambda_1(H) + \delta) \geq 1 - \delta. \) We choose \( \beta \geq \varepsilon^{O(1)} / \text{poly}(T) \) sufficiently small so that \( \delta \leq \varepsilon \).

Consider conditioning on any outcome to \(|\phi\rangle\) above such that \( \langle \phi | H | \phi \rangle < \lambda_1(H) + \delta \leq \lambda_1(H) + \varepsilon \). It follows from the guarantee in Theorem 24 that for \( \xi := R(|\phi\rangle | \phi\rangle) \) the verifier satisfies \( E[V(\xi)] \geq 1 - \varepsilon^{R(1)} \cdot \text{poly}(T) \). Thus, under no conditioning on \(|\phi\rangle\)
we have

$$E[V(\xi)] \geq 1 - \varepsilon^{O(1)} \cdot \text{poly}(T) - \delta \geq 1 - \varepsilon^{O(1)} \cdot \text{poly}(T).$$

This proves Theorem 23. □

7.2. Proof of Theorem 26. In our proof of Theorem 26, we use the perturbative gadgets and analysis ideas of Oliveira and Terhal [29], who build upon work of Kempe, Kitaev, and Regev [25]. Our main effort will be to show that, for any $k \geq 4$, there exists a $(k, \lceil k/2 \rceil + 1)$-AGPR. This will imply Theorem 26 in the cases $k = 5, 4$. Then, a slightly different reduction from [29] gives a $(3, 2)$-AGPR; this will complete the proof.

7.3. The locality-halving reduction.

7.3.1. The initial setup. Fix a constant $k \geq 4$. As the input to our $(k, \lceil k/2 \rceil + 1)$-AGPR, we are given a tuple $(H_{\text{targ}}, W, \beta)$, where $H_{\text{targ}}$ (which we will call the “target Hamiltonian”) is a $k$-local Hamiltonian expressed as the sum of some number $s$ of $k$-local terms over an $n$-qubit Hilbert space $\mathcal{H}_{\text{comp}} \cong \mathbb{B}^\otimes n$. All $k$-local terms of $H$ have operator norms $||H_i|| \in [W^{-1}, W]$.

By standard preprocessing steps, we can and will assume the following:

- $H_{\text{targ}}$ is a sum of $s'$ terms of form $H_i = H_{i,1}H_{i,2} \ldots H_{i,k}$, where each $H_{i,a}$ is 1-local\(^{1}\) and $||H_{i,a}|| \leq \text{poly}(s + W)$, and $H_{i,1}, \ldots, H_{i,k}$ act on distinct qubits (hence they commute). In what follows we write $s$ in place of $s'$.

- For each $i$, we assume $\min(||H_{i,1}H_{i,2} \ldots H_{i,\lceil k/2 \rceil}||, ||H_{i,\lceil k/2 \rceil + 1} \ldots H_{i,k}||) \in [1, K]$ for some $K \leq \text{poly}(W)$. (The lower bound is easily achieved by scaling $H_{\text{targ}}$ by a poly($W$) factor.)

To satisfy Definition 25, we will create a $\lceil k/2 \rceil + 1$-local derived Hamiltonian $H' = \tilde{H}$ on the larger Hilbert space $\mathcal{H} = \mathcal{H}_{\text{comp}} \otimes \mathcal{H}_{\text{anc}}$. We refer to $\mathcal{H}_{\text{comp}}, \mathcal{H}_{\text{anc}}$ as the computational and ancilla registers, respectively. For our quantum operation $R$ as in Definition 25, we will take the operation which simply measures the ancilla register in the standard basis.

7.3.2. Further preprocessing. First, we replace $H_{\text{targ}}$ with $H_{\text{targ}}^* := H_{\text{targ}} - M \cdot I$ for some $0 < M \leq \text{poly}(s + W)$ chosen large enough to ensure that $\lambda_1(H_{\text{targ}}^*)$ is less than $-1$. For any $j, |\psi\rangle$ we have

$$\lambda_j(H_{\text{targ}}^*) = \lambda_j(H_{\text{targ}}) - M \quad \text{and} \quad \langle \psi | H_{\text{targ}}^* | \psi \rangle = \langle \psi | H_{\text{targ}} | \psi \rangle - M.$$

$M \cdot I$ can be implemented 1-locally, so $H_{\text{targ}}^*$ is $k$-local.

In the remainder of our work, we will use $H_{\text{targ}}$ to denote $H_{\text{targ}}^*$, so that $\lambda_1(H_{\text{targ}})$ is now assumed to be less than $-1$.

7.3.3. The components of $H_{\text{targ}}$. For each $i \in [s]$, write $H_i = A_iB_i$, where we have grouped the $k$ factors of $H_i$ into a $\lceil k/2 \rceil$-local part $A_i$ and a $\lceil k/2 \rceil$-local part $B_i$. By our assumption, $||A_i||, ||B_i|| \geq 1$.

Exploiting cancellations and the fact that $A_i, B_i$ commute, we may write

$$H_i = (A_i^2 + B_i^2)/2 - (-A_i + B_i)^2/2 = -(A_i + B_i)^2/2 + H_i, \text{else},$$

\(^{1}\)Here, each $H_{i,a}$ denotes an operator over all of $\mathcal{H}_{\text{comp}}$, which is the tensor product $H_{i,a} = Y_{i,a} \otimes I_{\text{rest}}$ of an operator $Y_{i,a}$ on the Hilbert space of a single qubit with the identity operator $I_{\text{rest}}$ on the other $n - 1$ qubits.
where $H_{i,\text{else}}$ is a sum of $\lfloor k/2 \rfloor$-local and $\lceil k/2 \rceil$-local terms. Let

\begin{equation}
H_{\text{else}} := \sum_{i \in [s]} H_{i,\text{else}}.
\end{equation}

**7.3.4. The ancilla register.** For each index $i \in [s]$ corresponding to a term in $H_{\text{targ}}$, we introduce an ancilla qubit that we refer to as $w(i)$. Thus $\mathcal{H}_{\text{anc}}$ consists of $s$ qubits. For a Hamiltonian $E$ acting on the space of a single qubit, we use $E_{w(i)}$ to denote the application of $E$ to $w(i)$ (tensored with the identity on the rest of $\mathcal{H}_{\text{anc}}$). Similarly, for a Hamiltonian $F$ on $s$ qubits we use $F_{w}$ to indicate the operator on $\mathcal{H}_{\text{anc}}$ which applies $F$ to the ordered qubit-set $(w(1), \ldots, w(s))$.

**7.3.5. The derived Hamiltonian $\tilde{H}$.** The construction takes a parameter $0 < \Delta \leq \text{poly}(s, W, 1/\beta)$, to be chosen later as a sufficiently large value. We will take

\begin{equation}
\tilde{H} = H_0 + V,
\end{equation}

where

\begin{equation}
H_0 := \Delta \sum_{i \in [s]} |1\rangle\langle 1|_{w(i)},
\end{equation}

and where

\begin{equation}
V := H_{\text{else}} + \sqrt{\Delta/2} \cdot \sum_{i \in [s]} (-A_i + B_i) \otimes X_{w(i)}.
\end{equation}

Here, $X_{w(i)}$ is the Pauli $X$ operator applied to $w(i)$.

When we choose a large value $\Delta$, we will have $||H_0|| \gg ||V||$. In the analytical framework of [25, 29], $H_0$ is referred to as the “unperturbed” reference Hamiltonian; $V$ as the “perturbation” operator, regarded as “small”; and $\tilde{H}$ as the “perturbed” Hamiltonian, thought of as a slightly deformed version of $H_0$.

**7.4. Some tools for the analysis.**

**7.4.1. The effective Hamiltonian.** For future use we define

\begin{equation}
H_{\text{eff}} = H_{\text{targ}} \otimes |0\rangle\langle 0|_{w}.
\end{equation}

We will show that $\tilde{H}$ “behaves like” $H_{\text{eff}}$ in an appropriate sense; hence $H_{\text{eff}}$ is referred to as the “effective Hamiltonian” for $H$.

The eigenvalues of $H_{\text{eff}} = H_{\text{targ}} \otimes |0\rangle\langle 0|_{w}$ are the same as those of $H_{\text{targ}}$, along with 0. The introduction of this “unwanted” 0 eigenvalue is why we initially applied a global shift to $H_{\text{targ}}$ to assume its eigenvalues are negative, to ensure that the “lowest-energy part” of $H_{\text{targ}}$ is preserved. In particular, we have

\begin{equation}
\lambda_1(H_{\text{eff}}) = \lambda_1(H_{\text{targ}}) < -1, \quad 1 < ||H_{\text{eff}}|| < \text{poly}(s + W).
\end{equation}

**7.4.2. The eigenspaces of $H_0$, their projectors, and some notation.** In our analysis, we will use the derived Hamiltonian $H_0$ as a “reference” with which we decompose our Hilbert space $\mathcal{H} = \mathcal{H}_{\text{comp}} \otimes \mathcal{H}_{\text{anc}}$. First, it is obvious from the construction that $H_0$ has only nonnegative eigenvalues, including 0 and $\Delta$, and with no eigenvalues in $(0, \Delta)$. We define the subspaces

\begin{equation}
\mathcal{L}_-, \mathcal{L}_+ \leq \mathcal{H},
\end{equation}

\begin{equation}
\lambda_1(H_{\text{eff}}) = \lambda_1(H_{\text{targ}}) < -1, \quad 1 < ||H_{\text{eff}}|| < \text{poly}(s + W).
\end{equation}
where $L_-$ is the 0 eigenspace of $H_0$, and $L_+ := L_-^\perp$. We define $\Pi_-$, $\Pi_+$ as the projectors onto $L_-$ and $L_+$; we have the expressions

$$
\Pi_- = |0^s\rangle\langle 0^s|, \quad \Pi_+ = \sum_{x \in \{0,1\}^s \setminus 0^s} |x\rangle\langle x|.
$$

Now for any operator $A$ on $\mathcal{H}$, following [25, 29] we define

$$
A_{++} := \Pi_+ A \Pi_+, \quad A_{--} := \Pi_- A \Pi_-, \quad A_+ := \Pi_+ A \Pi_-, \quad A_- := \Pi_- A \Pi_+.
$$

Also define

$$
A_+ := A_{++}, \quad A_- := A_{--}.
$$

The $A_+$ notation will be used when $A(L_+) \subseteq L_+$, and similarly for $A_-, L_-$. 

### 7.4.3. Some perturbation theory definitions.

We will not introduce perturbation theory, only some definitions used here. The terms we introduce will be defined with reference to the “unperturbed” derived Hamiltonian $H_0$, explicitly and through the notation $A_{\pm\pm}$ introduced previously. In one definition we will also make reference to the perturbation operator $V$.

We define three functions,

$$
G, \tilde{G}, \Sigma_-
$$

each of which takes as input a value $z \in \mathbb{C}$ and outputs an operator over $H$; the definitions involve matrix inversion and for some values $z$ the output may be undefined. We define $\tilde{G}$, the *resolvent* of $\tilde{H}$, by

$$
\tilde{G}(z) := (zI - \tilde{H})^{-1}.
$$

Define the *self-energy* $\Sigma_-(z)$ by

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

### 7.4.4. The perturbation theorems.

Here we state a result from [29] that expresses the sense in which $\tilde{H}$ approximates $H_{\text{eff}}$. First we introduce one piece of helpful notation. For an operator $A$ over Hilbert space $\mathcal{H}$ and a subspace $S \subseteq \mathcal{H}$, we will use

$$
\|A\|_S := \max_{|\psi\rangle \in S \setminus \{0\}} \frac{\|A|\psi\rangle\|}{\|\psi\|}
$$

to denote the ($\ell_2$) operator norm of $A$ with inputs restricted to $S$.

**Theorem 28 (special case of [29, Theorem A.1]).** Say we are given Hamiltonians $H_0, \tilde{H}, V, H_{\text{eff}}$ and real values $\Delta > b > 0$, satisfying the following assumptions:

1. $\tilde{H} = H_0 + V$.
2. $\|V\| < \Delta/2$.
3. $H_0$ has the eigenvalues $\{0, \Delta\}$ with $L_-, L_+$ defined as above relative to $H_0$, and with operators $A_{\pm\pm}$ defined relative to these subspaces.

12 Here, in [29, Theorem A.1], we are fixing the setting $\lambda_* := \Delta/2$, as per the discussion in [29, pp. 19–20].
4. All eigenvalues of $H_{\text{eff}}$ are contained in $[-b, b]$.\footnote{We are setting $a := -b$ in Theorem A.1 of [29].}
5. $H_{\text{eff}} = \Pi_{+} H_{\text{eff}} \Pi_{-}$.

Next, fix $r, \varepsilon > 0$, and let $D_r := \{ z \in \mathbb{C} : |z| \leq r \}$ be the disk of radius $r$ in the complex plane, centered at the origin. Assume that

\[(60)\quad b + \varepsilon < r < \Delta/2.\]

Now our central assumption is that for all $z \in D_r$, the resolvent $\Sigma_{-}(z)$ is a good approximation to $H_{\text{eff}}$:

\[(60)\quad ||\Sigma_{-}(z) - H_{\text{eff}}|| \leq \varepsilon.\]

Let

\[(61)\quad \tilde{S} \leq \mathcal{H}\]

denote the “low-energy subspace” of $\tilde{H}$, namely, the subspace generated by the eigenvectors of $\tilde{H}$ whose eigenvalues are less than $\Delta/2$. Then $\tilde{S}$ has dimension at least 1. Moreover, it holds that $H_{\text{eff}}$ is well-approximated by $\tilde{H}$ on $\tilde{S}$:

\[(62)\quad ||\tilde{H} - H_{\text{eff}}||_{\tilde{S}} \leq 3(||H_{\text{eff}}|| + \varepsilon) ||\mathcal{V}|| + \frac{r(r + z_0)\varepsilon}{(r - b)(r - b - \varepsilon)}||\mathcal{V}||.\]

We will also use the following theorem from [25] relating the spectrum of $\tilde{H}$ to that of $H_{\text{eff}}$.

**Theorem 29** (special case of [25, Theorem 3]; see also [29, Theorem 7]). Under the same assumptions as in Theorem 28, we have the following. For every index $j$ for which $\lambda_j(\tilde{H}) < \Delta/2$ (in particular, this must include $j = 1$), we have

\[(63)\quad |\lambda_j(\tilde{H}) - \lambda_j(H_{\text{eff}})| \leq \varepsilon.\]

Theorem 29 is also used in the proof of Theorem 28.

**7.5. Application of the perturbation theorems.** For the construction of $H_0, \tilde{H}, \mathcal{V}$ described in section 7.3, it is immediate that conditions 1, 3, and 5 in Theorem 28 are satisfied. Condition 2, asking that $||\mathcal{V}|| < \Delta/2$, is satisfied for sufficiently large $\Delta \leq \text{poly}(s, W, 1/\beta)$; this follows by crudely bounding the norms of all terms used to define $\mathcal{V}$, using our initial norm-bound assumptions on $H_{\text{targ}}$.

As noted, the eigenvalues of $H_{\text{eff}}$ are the same as those of $H_{\text{targ}}$, along with 0. Thus we have $||H_{\text{eff}}|| \leq \text{poly}(s + W)$, independent of $\Delta$, and if we take $b := ||H_{\text{eff}}||$, condition 4 in Theorem 28 is satisfied.

Now, to satisfy the last requirement of that theorem, (60), we first set $r := 2b + \varepsilon$ with

\[\varepsilon := \beta/20.\]

(Recall that $\beta > 0$ is an input parameter to our desired AGPR.) Thus $D_r$ is a disk of radius $2||H_{\text{eff}}|| + \varepsilon$ in the complex plane, centered at the origin.

Our key tool is a bound shown in [29, p. 11, equation (25)]: for $|z| < \Delta$,

\[(64)\quad \Sigma_{-}(z) = \left( H_{\text{else}} + \frac{\Delta}{2(z - \Delta)} \sum_{i \in [s]} (-A_i + B_i)^2 \right) \otimes |0^s\rangle \langle 0^s| + O\left( \frac{||\mathcal{V}||^3}{(z - \Delta)^2} \right).\]
Note that for \( \Delta \gg z \) the left-hand term approaches \( H_{\text{eff}} \) (as defined in (52)), and the right-hand error term approaches 0. Indeed, following the discussion in [29, pp. 11, 20], by taking a sufficiently large \( \Delta \leq \text{poly}(s + W)/\varepsilon^2 \) we obtain

\[
|\Sigma_-(z) - H_{\text{eff}}| \leq \varepsilon \quad \forall z \in D_r.
\]

Thus all requirements of Theorem 28 are satisfied for our settings, and we conclude that

\[
||\tilde{H} - H_{\text{eff}}||_S \leq \frac{3(||H_{\text{eff}}|| + \varepsilon)||V||}{\Delta - ||H_{\text{eff}}|| - \varepsilon} + \frac{r(r + z_0)\varepsilon}{(r - b)(r - b - \varepsilon)}
\]

\[
\leq \varepsilon + 4\varepsilon = 5\varepsilon,
\]

with the last inequality valid if we choose \( \Delta \) large enough compared to \( ||H_{\text{eff}}|| \). For future work, we also stipulate that \( \Delta \) be chosen large enough to satisfy

\[
\frac{1}{\Delta} \leq \frac{\varepsilon}{2||H_{\text{eff}}||}.
\]

All this only requires \( \Delta \leq \text{poly}(s, W, 1/\beta) \).

Under the same settings to our parameters, it is immediate that we also obtain the conclusions of Theorem 29. In particular, using (53) we have

\[
|\lambda_1(\tilde{H}) - \lambda_1(H_{\text{targ}})| = |\lambda_1(\tilde{H}) - \lambda_1(H_{\text{eff}})| \leq \varepsilon.
\]

For future work, we note that \( \tilde{S} \) is a proper subspace of \( \mathcal{H} \), since \( ||\mathcal{H}|| \geq ||H_0|| - ||V|| \geq \Delta - \Delta/2 \).

### 7.5.1. Consequences for nearly minimal energy states.

Let us now consider any nearly minimal energy state \( |\psi\rangle \in \mathcal{H} \) for the Hamiltonian \( \tilde{H} \), satisfying

\[
\langle \psi|\tilde{H}|\psi\rangle < \lambda_1(\tilde{H}) + \beta < \lambda_1(H_{\text{eff}}) + \beta + \varepsilon.
\]

We will upper-bound \( \langle \psi|H_{\text{eff}}|\psi\rangle \) to show that \( |\psi\rangle \) is also nearly minimal energy for this second Hamiltonian.

A small complication for our analysis is that \( |\psi\rangle \) may not lie within \( \tilde{S} \). Decompose \( |\psi\rangle \) as

\[
|\psi\rangle = \alpha_1|\psi_\tilde{S}\rangle + \alpha_2|\psi_{\tilde{S}^\perp}\rangle,
\]

according to its components in \( \tilde{S} \) and its orthogonal complement \( \tilde{S}^\perp \) (so, we have \( |\alpha_1|^2 + |\alpha_2|^2 = 1 \) and \( \langle \psi_\tilde{S}|\psi_{\tilde{S}^\perp}\rangle = 0 \)). Recall that both of these spaces have dimension at least 1. We assume that \( a \) is real and positive; this assumption is without loss of generality, by applying a phase factor \( \overline{\alpha}_1/|\alpha_1| \) to the state if necessary, and just simplifies our expressions slightly.

By the definition of \( \tilde{S}^\perp \), we see that it is spanned by eigenvectors of \( \tilde{H} \) with eigenvalues \( \geq \Delta/2 \). Thus,

\[
\langle \psi|\tilde{H}|\psi\rangle = |\alpha_1|^2\langle \psi_\tilde{S}|\tilde{H}|\psi_\tilde{S}\rangle + |\alpha_2|^2\langle \psi_{\tilde{S}^\perp}|\tilde{H}|\psi_{\tilde{S}^\perp}\rangle
\]

\[
\geq \lambda_1(\tilde{H}) + |\alpha_2|^2\Delta/2.
\]
Combining this with (70) and (69), we find
\begin{equation}
|\alpha_2|^2 \leq \frac{2\beta}{\Delta} \leq \frac{\varepsilon}{\|H_{\text{eff}}\|},
\end{equation}
where the last step follows from our prior largeness requirement on $\Delta$ in (68). It also follows that $|\alpha_1 - 1|^2 \leq \sqrt{1 - \varepsilon}/\|H_{\text{eff}}\| \leq 1 - \varepsilon/\|H_{\text{eff}}\| - 1 \leq \varepsilon^2/\|H_{\text{eff}}\|^2$ (using here that $\alpha_1 \in \mathbb{R}^+$. For analysis purposes, define the (nonnormalized) state
\begin{equation}
|v\rangle := (\alpha_1 - 1)|\psi_S\rangle + \alpha_2|\psi_{S^+}\rangle.
\end{equation}
We have
\begin{equation}
\|\|v\|\|^2 = \langle v|v \rangle = |\alpha_1 - 1|^2 + |\alpha_2|^2 \leq \frac{2\varepsilon}{\|H_{\text{eff}}\|}.
\end{equation}
Now note that, using the definition of $|v\rangle$ and (76), we have
\begin{align}
\langle \psi|H_{\text{eff}}|\psi \rangle &= \langle \psi_S|H_{\text{eff}}|\psi_S \rangle + \langle v|H_{\text{eff}}|v \rangle \\
&\leq \langle \psi_S|H_{\text{eff}}|\psi_S \rangle + \|H_{\text{eff}}\| \cdot \|\|v\|\|^2 \\
&\leq \langle \psi_S|H_{\text{eff}}|\psi_S \rangle + 4\varepsilon.
\end{align}
Next, applying (67) and the fact that $|\psi_S\rangle \in \tilde{S}$, we obtain
\begin{align}
\langle \tilde{S}|H_{\text{eff}}|\tilde{S} \rangle &\leq \left( \langle \tilde{S}|\tilde{H}|\tilde{S} \rangle + \|\tilde{H} - H_{\text{eff}}\| \cdot \|\|\tilde{S}\|\|^2 \right) + \varepsilon \\
&\leq \langle \tilde{S}|\tilde{H}|\tilde{S} \rangle + 6\varepsilon \\
&\leq \langle \psi|H_{\text{eff}}|\psi \rangle + 6\varepsilon \\
&\leq \lambda_1(H_{\text{eff}}) + 6\varepsilon + \beta.
\end{align}
(In the third inequality, we used the definition of $\tilde{S}$ as a low-energy subspace for $\tilde{H}$, and the fact that $|\psi_S\rangle$ is the component of $|\psi\rangle$ in $\tilde{S}$. In the last step, we used (70).) Combining (79) and (83), we conclude that
\begin{equation}
\langle \psi|H_{\text{eff}}|\psi \rangle \leq \lambda_1(H_{\text{eff}}) + 10\varepsilon + \beta < \lambda_1(H_{\text{eff}}) + 2\beta.
\end{equation}
Thus $|\psi\rangle$ is also nearly minimal energy for $H_{\text{eff}} = H_{\text{targ}} \otimes |0^r\rangle\langle 0^r|$.  

### 7.5.2. Obtaining a nearly minimal energy state for $H_{\text{targ}}$. 
Recall that $L_-$ is the subspace of $\mathcal{H}$ in which the ancilla qubits are all-zero. Any computational basis state in which the ancillas are not all-zero vanishes under the action of $H_{\text{eff}}$. For our state $|\psi\rangle$ as above, write
\begin{equation}
|\psi\rangle = |w\rangle|\psi_-\rangle + |z\rangle|\psi_+\rangle,
\end{equation}
where $|\psi_-\rangle \in L_-,|\psi_+\rangle \in L_+$ are unit vectors. Re-expressing our inner product in this basis, we have $\langle \psi|H_{\text{eff}}|\psi \rangle \geq |w|^2 \cdot \lambda_1(H_{\text{eff}}) + 0$, so by (84), and using the facts that $\lambda_1(H_{\text{eff}}) < -1$ and $10\varepsilon + \beta < 1$, we have
\begin{equation}
|w|^2 \geq 1 - \frac{10\varepsilon + \beta}{|\lambda_1(H_{\text{eff}})|} > 1 - 2\beta.
\end{equation}
Recall that the quantum operation $R$ measures the ancilla register of $|\psi\rangle$. By the above, with probability $> 1 - 2\beta$ this measurement yields the all-zero outcome, and
the post-measurement state is $|\psi_-\rangle$. Identifying $\mathcal{L}_-$ with the Hilbert space $\mathcal{H}_{\text{comp}}$, on which $H_{\text{targ}}$ acts, we have

$$
(87) \quad \langle \psi_- | H_{\text{targ}} | \psi_- \rangle = \frac{1}{|w|^2} \langle \psi | H_{\text{eff}} | \psi \rangle
$$

$$
(88) \quad \leq \lambda_1 (H_{\text{eff}}) + 10\varepsilon + \beta
$$

$$
(89) \quad = \lambda_1 (H_{\text{targ}}) + 10\varepsilon + \beta
$$

$$
(90) \quad < \lambda_1 (H_{\text{targ}}) + 2\beta
$$

using (53) in the penultimate step. Thus $(H', R)$ have the required AGPR properties (where we may take $\delta := 2\beta$ in Definition 25). We have proved Theorem 26 for the cases $k = 5, 4$.

7.6. The 3-local-to-2-local reduction. Given a 3-local target Hamiltonian $H_{\text{targ}}$, we can use a different gadget construction in [29, pp. 11–12]. The construction uses the same (1-local) unperturbed Hamiltonian $H_0 := \Delta \sum_{i \in [s]} |1\rangle \langle 1|_{w(i)}$ and the same effective Hamiltonian $H_{\text{eff}} := H_{\text{targ}} \otimes |0\rangle \langle 0|_{\text{comp}}$, with a different perturbation Hamiltonian $V$ (this time 2-local), which again satisfies $\|V\| < \Delta/2$ for sufficiently large $\Delta \leq \text{poly}(s, W, 1/\beta)$. As described in [29], for large enough $\Delta \leq \text{poly}(s, W, 1/\beta)$ one can ensure $\|\Sigma_-(z) - H_{\text{eff}}\| \leq \varepsilon$ for $\varepsilon := \beta/20$ and for $z$ in a disk of appropriately chosen radius. This allows us to apply Theorems 28 and 29 in the same fashion as before. This yields the required $(3, 2)$-AGPR, completing the proof of Theorem 26.

8. Further implications for quantum complexity theory. In this section, we use the $\text{BQP}/\text{poly} = \text{YQP}^*/\text{poly}$ theorem to harvest two more results about quantum complexity classes. The first is an “exchange theorem” stating that $\text{QCMA}/\text{poly} \subseteq \text{QMA}/\text{poly}$: in other words, one can always simulate quantum advice together with a classical witness by classical advice together with a quantum witness. This is a straightforward generalization of Theorem 20. The second result is a quantum Karp–Lipton theorem, which states that if $\text{NP} \subseteq \text{BQP}/\text{poly}$ (that is, NP-complete problems are efficiently solvable by quantum computers with quantum advice), then $\Pi^P_2 \subseteq \text{QMA}^{\text{PromiseQMA}}$, which one can think of as “almost as bad” as a collapse of the polynomial hierarchy. This result makes essential use of Theorem 20 and is a good illustration of how that theorem can be applied in quantum complexity theory.

Theorem 30 (exchange theorem). $\text{QCMA}/\text{poly} \subseteq \text{QMA}/\text{poly}$.

Proof. The proof is almost the same as that of Theorem 20. Fix any $L \in \text{QCMA}/\text{poly}$. Then there exists a polynomial-time quantum verifier $Q$, a family of polynomial-size advice states $\{\rho_n\}_n$, and a polynomial $p$ such that for all inputs $x \in \{0, 1\}^n$:

- $x \in L \implies \exists w \in \{0, 1\}^p(n) \mathbb{E}[Q(x, w, \rho_n)] \geq 2/3$.
- $x \notin L \implies \forall w \in \{0, 1\}^p(n) \mathbb{E}[Q(x, w, \rho_n)] \leq 1/3$.

Now consider the following promise problem: given $x$ and $w$ as input (regarded as two parts of the classical input string), as well as a constant $c \in [0, 1]$, decide whether $\mathbb{E}[Q(x, w, \rho_n)]$ is at most $c - 1/10$ or at least $c + 1/10$, promised that one of these is the case. (Equivalently, estimate the probability within an additive error $\pm 1/10$.) This problem is clearly in $\text{PromiseBQP}/\text{poly}$, since we can take $\rho_n$ as the advice. So by Theorem 20, the problem is in $\text{PromiseYQP}^*/\text{poly}$ as well, as witnessed by an input-oblivious advice-testing algorithm $Y((x, w), \sigma, a)$ and a classical advice string family $\{a_n\}_{n > 0}$. (By slight abuse of index notation, the advice string $a_n$ is taken to possess the correctness guarantee in Theorem 20 for inputs $(x, w) \in \{0, 1\}^n \times \{0, 1\}^{p(n)}$ obeying the promise.)
Fig. 2. Containments among complexity classes related to quantum proofs and advice, in light of this paper’s results. The containments \( \text{QMA/qpoly} \subseteq \text{PSPACE/poly} \) and \( \text{QCMA/qpoly} \subseteq \text{PP/poly} \) were shown previously by Aaronson [4]. This paper shows that \( \text{BQP/qpoly} \subseteq \text{QMA/poly} \), and indeed \( \text{BQP/qpoly} = \text{YQP/poly} \). It also shows that \( \text{QCMA/qpoly} \subseteq \text{QMA/poly} \).

Our \( \text{QMA/poly} \) verifier takes the Promise\( \text{YQP}^{\ast}/\text{poly} \) advice string \( a_n \) as its trusted classical advice, and a state of the form \( \sigma \otimes |w\rangle\langle w| \) as its untrusted witness state. It acts as follows:

1. Execute \( Y((x, w), \sigma, a_n) \), rejecting if the advice-testing bit \( b_{\text{adv}} = 0 \).
2. If \( b_{\text{adv}} = 1 \), measure the bit \( b_{\text{out}} \) from the same execution of \( Y \) and output this bit.

The protocol is polynomial-time, since \( Y \) is a polynomial-time quantum algorithm, and the completeness and soundness properties follow directly from the guarantees of Theorem 20.

Indeed, let \( \text{YQ-QCMA} \) denote the complexity class where a BQP verifier receives a classical untrusted witness that depends on the input, as well as an untrusted quantum witness that depends only on the input size \( n \). Then we can characterize QCMA/qpoly as equal to \( \text{YQ-QCMA/poly} \), similarly to how we characterized BQP/qpoly as equal to YQP/poly. Figure 2 shows the structure of the quantum complexity classes under discussion.

We now use Theorem 20 to prove an analogue of the Karp–Lipton theorem for quantum advice.

Recall that a promise problem is a pair \( \Pi = (\Pi_{\text{yes}}, \Pi_{\text{no}}) \) of disjoint subsets of \( \{0,1\}^* \). We say that a language \( A \) solves \( \Pi \) if for all \( x \in \Pi_{\text{yes}} \cup \Pi_{\text{no}} \), we have \( x \in A \iff x \in \Pi_{\text{yes}} \). We say that a language \( L \) is in \( \text{QMA}^{\Pi} \) if there is a single QMA verifier \( V^A \) with oracle access that witnesses the membership \( L \in \text{QMA}^A \) for any language \( A \) solving \( \Pi \). We let \( \text{QMA}^\text{PromiseQMA} := \bigcup_{\Pi \in \text{PromiseQMA}} \text{QMA}^{\Pi} \). This model of oracle access to promise problems, in which the machine may query strings...
violating the promise II (and for which the oracle may give arbitrary responses), is fairly standard; see, e.g., [15].

**Theorem 31** (quantum Karp-Lipton theorem). If $\mathsf{NP} \subseteq \mathsf{BQP}/\text{qpoly}$, then $\Pi_2^P \subseteq \mathsf{QMA}^{\text{PromiseQMA}}$.

In this result we use the model of oracle access to a promise problem which allows the algorithm to query inputs not obeying the promise; in such cases this allows the oracle to answer such queries arbitrarily. This model is fairly standard; see, e.g, [15].

Previously, Aaronson [3] showed that if $\mathsf{PP} \subseteq \mathsf{BQP}/\text{qpoly}$, then the counting hierarchy $\text{CH}$ collapses. However, he had been unable to show that $\mathsf{NP} \subseteq \mathsf{BQP}/\text{qpoly}$ would have unlikely consequences in the uniform world.

**Proof of Theorem 31.** By Theorem 20, the hypothesis implies $\mathsf{NP} \subseteq \mathsf{YQP}/\text{poly} = \mathsf{YQP}^*/\text{poly}$. So let $Y$ be a $\mathsf{YQP}^*/\text{poly}$ algorithm for $\mathsf{SAT}$, which takes an input $x \in \{0,1\}^n$ (representing a CNF formula), a trusted classical nonuniform advice string $a \in \{0,1\}^{\ell(n)}$ for some $\ell(n) \leq \text{poly}(n)$, and an untrusted advice state $\rho$ on $q(n) \leq \text{poly}(n)$ qubits. By inspecting the proof of Theorem 15, we see that the completeness and soundness parameters $.9, 1$ in Definition 19 can easily be strengthened to $(1 - e^{-n}, n^{-100})$; we assume that this holds for $Y$. Let $\{a_n\}_{n \geq 0}$ be the associated family of classical advice strings of length $\ell(n)$.

Now consider an arbitrary language $L \in \Pi_2^P$. As such, $L$ is defined by a deterministic polynomial-time predicate $R(x,y,z)$:

$$x \in L \iff \forall y \exists z : \ R(x,y,z) = 1,$$

where we expect $|y| = |z| = p(n)$ for some $p(n) \leq \text{poly}(n)$ on inputs $x \in \{0,1\}^n$.

Using $Y$ along with Cook’s theorem applied to the predicate $R$, we can construct a polynomial-time input-oblivious advice-testing algorithm $Y'\big(x,y,\rho,a\big)$ producing output bits $b_{\text{adv}},b_{\text{out}}$ (we use the notation $Y'_{\text{adv}},Y'_{\text{out}}(x,y,\rho,a)$ to denote the values of these two bits in an execution of $Y'$ on $(x,y,\rho,a)$, noting that $\mathbb{E}[b_{\text{adv}}]$ depends only on $\rho,a$), which has the following properties:

(P1) There exists a $\rho$ such that $\mathbb{E}[Y'_{\text{adv}}(x,y,\rho,a_n)] \geq 1 - 2^{-n}$ for all $x,y$.

(P2) For any $\rho$, if $\mathbb{E}[Y'_{\text{adv}}(x,y,\rho,a_n)] \geq n^{-3}$, then we have

$$\mathbb{E}[Y'_{\text{out}}(x,y,\rho,a_n) | b_{\text{adv}} = 1] \geq 1 - 1/(n \cdot p(n))$$

if there exists a $z$ such that $R(x,y,z)$ holds, and

$$\mathbb{E}[Y'_{\text{out}}(x,y,\rho,a_n) | b_{\text{adv}} = 1] \leq 1/(n \cdot p(n))$$

otherwise.

Using the standard search-to-decision reduction for $\mathsf{SAT}$, we can then strengthen property (P2) to the following for a polynomial-time quantum algorithm $Y'' = Y''(x,y,\rho,a)$ outputting a bit $b_{\text{adv}}$ (denoted $Y''_{\text{adv}}(x,y,\rho,a)$) along with a string $z \in \{0,1\}^{\ell(n)}$. Here as before, the bit $b_{\text{adv}}$ has expectation determined by $\rho,a$ alone.

The algorithm $Y''$ satisfies the following:

(P1') There exists a $\rho$ such that $\mathbb{E}[Y''_{\text{adv}}(x,y,\rho,a_n)] \geq 1 - 2^{-n}$ for all $x,y$.

(P2') For all $x,y$ pairs for which some $z$ satisfies $R(x,y,z) = 1$, and for all states $\rho$, we have the following. If $\mathbb{E}[Y''_{\text{adv}}(x,y,\rho,a_n)] \geq .01$, and if we condition on

---

14This reduction requires repeated use of the advice state $\rho$ to obtain the bits of a lexicographically first such $z$; these measurements may alter $\rho$. This is not a serious obstacle, however, by the principle that a measurement whose outcome is nearly information-theoretically certain has a small expected effect on the measured state.
\[
[b_{\text{adv}} = 1] \text{ in this execution, then with probability at least } .99, Y''(x, y, \rho, a_n) \text{ outputs } z \text{ such that } R(x, y, z) = 1.
\]

Now let \( U(x, y, \rho, a) \) be a quantum algorithm outputting a single bit, and expecting \( y, \rho, a \) of size determined by \( n = |x| \) exactly as with \( Y'' \). The algorithm \( U \) executes \( Y''(x, y, \rho, a) \) and does one of the following, both with equal probability:

- outputs \( \neg b_{\text{adv}} \);
- outputs \( 1 \) if and only if the string \( z \) outputted by \( Y'' \) satisfies \( R(x, y, z) = 1 \).

\( U \) is polynomial-time, and we claim that

\[(A1) \quad x \in L \implies \exists a, \rho : [E[Y''_{\text{adv}}(x, y, \rho, a)] \geq 9/10] \land [\text{for all } \sigma, y : E[U(a, \sigma, x, y)] \geq 1/5];
\]

\[(A2) \quad x \notin L \implies \forall a, \rho : [E[Y''_{\text{adv}}(x, y, \rho, a)] \leq 2/3] \lor [\exists \sigma, y : E[U(a, \sigma, x, y)] \leq 1/6].
\]

With reference to the machine \( U \), we define the promise problem \( \Pi = (\Pi_{\text{yes}}, \Pi_{\text{no}}) \) by

\[
\Pi_{\text{yes}} = \{(x, a) \in \{0,1\}^{n+\ell(n)} : \exists \rho, y \text{ such that } E[U(x, y, \rho, a)] \leq 1/6\},
\]

\[
\Pi_{\text{no}} = \{(x, a) \in \{0,1\}^{n+\ell(n)} : \forall \rho, y \text{ we have } E[U(x, y, \rho, a)] \geq 1/5\}
\]

and note that \( \Pi \in \text{PromiseQMA} \) by standard techniques. Also, it is clear that \( A1 \) and \( A2 \) together imply \( L \in \text{QMA}^1 \subseteq \text{QMA}^{\text{PromiseQMA}} \). (The crucial point here is that \( U \) does not take the existentially quantified advice state \( \rho \) as input in our query to \( \Pi \)—and therefore, the QMA machine does not need to pass a quantum state to the \text{PromiseQMA} oracle, which would be illegal. This is why we needed the \text{BQP/qpoly} = \text{QQP}/\text{poly} result here. Note also that in the case where \( x \in L \), our claim gives no control over the relevant acceptance probabilities of \( Q_1 \) and \( U \) for settings to \( a \) other than the “correct” setting; this necessitates the use of a \text{PromiseQMA} oracle—which is allowed to behave arbitrarily on inputs not obeying the promise \( \Pi \)—rather than \text{a PromiseQMA oracle}.)

We now prove \( A1 \) and \( A2 \). First suppose \( x \in L \). Then there exists an advice string \( a_n \) with the following properties:

\[(B1) \quad \text{There exists a } a_n \text{ such that } E[Y''_{\text{adv}}(x, y, \rho_n, a_n)] \geq 9/10 \text{ for all } y. \quad \text{(By } P1').\]

\[(B2) \quad \text{For all } \sigma, y \text{ pairs, either } E[Y''_{\text{adv}}(x, y, \rho, a_n)] \leq 1/2 \text{, or for the string } z \text{ outputted by this execution of } Y'' \text{, we have } Pr[R(x, y, z) \text{ holds}] \geq (5) \cdot (.99) > 2/5. \quad \text{(By } P2')\]

By \( B2 \), we have for all \( \sigma, y \) \( E[U(a, \sigma, x, y)] \geq 1/5 \). This proves \( A1 \).

Next suppose \( x \notin L \). Then given an advice string \( a \), suppose there exists a pair \( \rho, y \) such that \( E[Y''_{\text{adv}}(x, y, \rho, a)] > 2/3. \) (Then this relation holds for all \( y \), since \( E[b_{\text{adv}}] \) is a function of \( \rho, a \) alone.) Set \( \sigma := \rho \), and choose \( y \) for which there is no \( z \) such that \( R(x, y, z) \) holds. Then for the random string \( z \) as produced by \( Y''(x, y, \sigma, a) \) we have \( Pr[R(x, y, z) = 1] = 0 \), since \( x \notin L \).

It follows from the above that \( Pr[U(a, \sigma, x, y) \text{ accepts}] < \frac{1}{2}(1/3 + 0) = 1/6 \). This proves \( A2 \) and completes the proof of the theorem.

9. **Open problems.** One open problem is simply to find more applications of the majority-certificates lemma, which seems likely to have uses outside of quantum complexity theory. Can we improve the parameters of the majority-certificates lemma (the size of the certificates or the number \( O(n) \) of certificates), or alternatively, show that the current parameters are essentially optimal? Also, can we prove the real-valued majority-certificates lemma with an error tolerance \( \alpha \) that depends only on the desired accuracy \( \varepsilon \) of the final approximation, not on \( n \) or the fat-shattering dimension of \( S \)?
On the quantum complexity side, we mention several questions. First, in Theorem 22, is the polynomial blowup in the number of qubits unavoidable? Could one hope for a way to simulate an \( n \)-qubit advice state by the ground state of \( n \)-qubit local Hamiltonian, or would that have implausible complexity consequences? Second, can we use the ideas in this paper to prove any upper bound on the class QMA/qpoly better than the \( \text{PSPACE}/\text{poly} \) upper bound shown by Aaronson [4]? Third, if \( \text{NP} \subset \text{BQP}/\text{qpoly} \), then does QMA\(^{\text{PromiseQMA}}\) contain not just \( \Pi_2^p \) but the entire polynomial hierarchy? Finally, is \( \text{BQP}/\text{qpoly} = \text{BQP}/\text{poly} \)?

Acknowledgments. We thank Sanjeev Arora, Kai-Min Chung, Avinatan Hassidim, Ashwin Nayak, Roberto Oliveira, Thomas Vidick, John Watrous, and Colin Zheng for helpful comments and discussions, and the anonymous reviewers for their comments. We are particularly grateful to a reviewer who pointed out the need for further analysis of the local Hamiltonian reductions used in our work.

REFERENCES

A FULL CHARACTERIZATION OF QUANTUM ADVICE