# The Power of Unentanglement (Extended Abstract)

Scott Aaronson* MIT

Salman Beigi MIT

Andrew Drucker MIT

Bill Fefferman University of Chicago

Peter Shor MIT

## Abstract

*The class* $\mathsf{QMA}(k)$*, introduced by Kobayashi et al., consists of all languages that can be verified using* $k$ *unentangled quantum proofs. Many of the simplest questions about this class have remained embarrassingly open: for example, can we give any evidence that* $k$ *quantum proofs are more powerful than one? Can we show any upper bound on* $\mathsf{QMA}(k)$*, besides the trivial* $\mathsf{NEXP}$*? Does* $\mathsf{QMA}(k) = \mathsf{QMA}(2)$ *for* $k \geq 2$*? Can* $\mathsf{QMA}(k)$ *protocols be amplified to exponentially small error?*

*In this paper, we make progress on all of the above questions.*

- *We give a protocol by which a verifier can be convinced that a* 3SAT *formula of size* $n$ *is satisfiable, with constant soundness, given* $\widetilde{O}(\sqrt{n})$ *unentangled quantum witnesses with* $O(\log n)$ *qubits each. Our protocol relies on Dinur's version of the PCP Theorem and is inherently non-relativizing.*
- *We show that assuming the famous Additivity Conjecture from quantum information theory, any* $\mathsf{QMA}(2)$ *protocol can be amplified to exponentially small error, and* $\mathsf{QMA}(k) = \mathsf{QMA}(2)$ *for all* $k \geq 2$.
- *We give evidence that* $\mathsf{QMA}(2) \subseteq \mathsf{PSPACE}$*, by showing that this would follow from "strong amplification" of* $\mathsf{QMA}(2)$ *protocols.*
- *We prove the nonexistence of "perfect disentanglers" for simulating multiple Merlins with one.*

## 1 Introduction

Quantum entanglement is often described as a complicated, hard-to-understand resource. But ironically, many questions in quantum computing are easiest to answer assuming unlimited entanglement, and become much more difficult if entanglement is *not* allowed! One way to understand this is that Hilbert space—the space of *all* quantum states—has extremely useful linear-algebraic properties, and when we restrict to the set of separable states we lose many of those properties. So for example, finding a quantum state that maximizes the probability of a given measurement outcome is just a principal eigenvector problem, but finding a separable state that does the same is $\mathsf{NP}$-hard [6].

These observations naturally give rise to a general question at the intersection of computational complexity and entanglement theory. Namely: supposing we had $k$ quantum proofs, could we use the promise that the proofs were unentangled to verify mathematical statements beyond what we could verify otherwise?

### 1.1 Background and Related Work

The class $\mathsf{QMA}$, or Quantum Merlin-Arthur, consists of all languages that admit a proof protocol in which Merlin sends Arthur a polynomial-size quantum state $|\psi\rangle$, and then Arthur decides whether to accept or reject in quantum polynomial time. This class was introduced by Kitaev [11] and Watrous [20] as a quantum analogue of $\mathsf{NP}$. By now we know a reasonable amount about $\mathsf{QMA}$: for example, it allows amplification of success probabilities [15], is contained in $\mathsf{PP}$ [15], and has natural complete promise problems [11]. (See Aharonov and Naveh [3] for a survey.)

In 2003, Kobayashi, Matsumoto, and Yamakami [12] defined a generalization of $\mathsf{QMA}$ called $\mathsf{QMA}(k)$. Here there are $k$ Merlins, who send Arthur $k$ quantum proofs $|\psi_1\rangle, \ldots, |\psi_k\rangle$ respectively that are guaranteed to be unentangled with each other. (Thus $\mathsf{QMA}(1) = \mathsf{QMA}$.) Notice that in the classical case, this generalization is completely uninteresting: we have $\mathsf{MA}(k) = \mathsf{MA}$ for all $k$, since we can always simulate $k$ Merlins by a single Merlin who sends Arthur a concatenation of the $k$ proofs. In the quantum case, however, a single Merlin could cheat by *entangling* the $k$ proofs, and we know of no general way to detect such entanglement.

When we try to understand $\mathsf{QMA}(k)$, we encounter at least four basic questions. First, do multiple quantum

*To whom correspondence should be addressed. Email: aaronson@csail.mit.edu.

proofs ever actually help? That is, can we find some sort of evidence that $\mathsf{QMA}(k) \neq \mathsf{QMA}(1)$ for some $k$? Second, can we show any nontrivial upper bound on the power of multiple quantum proofs? (The trivial upper bound is $\mathsf{QMA}(k) \subseteq \mathsf{NEXP}$, which follows by just guessing exponential-size classical descriptions of the $k$ quantum proofs.) Third, can $\mathsf{QMA}(k)$ protocols be amplified to exponentially small error? Fourth, are two Merlins the most we ever need? That is, does $\mathsf{QMA}(k) = \mathsf{QMA}(2)$ for all $k \geq 2$?

We know of three previous results that are relevant to the above questions.

First, in their original paper on $\mathsf{QMA}(k)$, Kobayashi et al. [12] proved that a positive answer to the third question implies a positive answer to the fourth. That is, if $\mathsf{QMA}(k)$ protocols can be amplified, then $\mathsf{QMA}(k) = \mathsf{QMA}(2)$ for all $k \geq 2$.

Second, Liu, Christandl, and Verstraete [14] gave a natural problem from quantum chemistry, called *pure state $N$-representability*, which is in $\mathsf{QMA}(2)$ but is not known to be in $\mathsf{QMA}$.

Third, Blier and Tapp [6] recently (and independently of us) gave an interesting $\mathsf{QMA}(2)$ protocol for an $\mathsf{NP}$-complete problem, namely 3-COLORING. In this protocol, Arthur verifies that an $n$-vertex graph $G$ is 3-colorable, using two unentangled witnesses with only $O(\log n)$ qubits each. There is a crucial caveat, though: if $G$ is *not* 3-colorable, then Arthur can only detect this with probability $\Omega\left(1/n^6\right)$ rather than constant probability.[1]

### 1.2 Our Results

In this paper, we present new results about all four problems listed previously. Our main results are as follows:

**Proving 3SAT With $\widetilde{O}(\sqrt{n})$ Qubits.** In Section 3, we give a protocol by which Arthur can verify that a 3SAT instance of size $n$ has a satisfying assignment, using $O(\sqrt{n}\operatorname{polylog} n)$ unentangled witnesses with $O(\log n)$ qubits each. Of course, this is a larger number of qubits than in the protocol of Blier and Tapp [6], but the point is that Arthur can detect cheating with *constant* probability. Our protocol relies on the PCP Theorem, and in particular, on the existence of PCP's of size $O(n \operatorname{polylog} n)$, which was recently shown by Dinur [9].

**Additivity Implies Amplification.** In Section 4, we reduce several open problems about $\mathsf{QMA}(k)$ to the famous *Additivity Conjecture* in quantum information theory. In particular, we show that the Additivity Conjecture implies that any $\mathsf{QMA}(k)$ protocol can be amplified to exponentially small error, and that the "$\mathsf{QMA}(k)$ hierarchy" collapses to $\mathsf{QMA}(2)$. Assuming the Additivity Conjecture, we also show that letting the Merlins have a limited amount of entanglement does not change the power of $\mathsf{QMA}(2)$, and neither does forcing their witnesses to be identical.

**Evidence That $\mathsf{QMA}(k) \subseteq \mathsf{PSPACE}$.** In Section 5, we give the first evidence for an upper bound on $\mathsf{QMA}(k)$ better than the trivial $\mathsf{NEXP}$. In particular, we show that $\mathsf{QMA}(k) \subseteq \mathsf{PSPACE}$, assuming what we call the *Strong Amplification Conjecture*: that it is possible to amplify $\mathsf{QMA}(k)$ protocols in such a way that one of the Merlin's Hilbert space dimensions remains smaller than the inverse of the error bound.

**Nonexistence of Perfect Disentanglers.** In Section 6, we rule out one possible approach to showing $\mathsf{QMA}(2) = \mathsf{QMA}$, by giving an extremely simple result that nevertheless seems new and might be of interest. Namely, given finite-dimensional Hilbert spaces $\mathcal{H}, \mathcal{K}$, there is no quantum operation mapping the set of all states in $\mathcal{H}$ to the set of all separable states in $\mathcal{K} \otimes \mathcal{K}$.

In the remainder of this introduction, we give some intuition behind each of these results.

### 1.3 Proving 3SAT With $\widetilde{O}(\sqrt{n})$ Qubits

Let $\varphi$ be a 3SAT instance with $n$ variables. Then how long a proof does Merlin need to send Arthur, to convince him that $\varphi$ is satisfiable? (As usual, Merlin is an omniscient prover and Arthur is a skeptical $\mathsf{BPP}$ verifier.)

Intuitively, it seems the answer should be about $n$ bits. Certainly, if sublinear-size proofs of satisfiability existed, then 3SAT would be in solvable in $2^{o(n)}$ time, since Arthur could just loop over all possible proofs until he found one that worked. Even in the quantum case, one can make a similar statement: if *quantum* proofs of satisfiability with $o(n)$ qubits existed, then 3SAT would have a $2^{o(n)}$-time quantum algorithm.[2]

On the other hand, suppose Arthur is given *several* quantum proofs, which are guaranteed to be unentangled with each other. Then the previous argument no longer seems to work.[3] And this at least raises the possibility that 3SAT might have sublinear proofs in this setting.

We will show that this possibility is realized:

**Theorem 1.** *Let $\varphi$ be a satisfiable 3SAT instance with $n$ variables and $m \geq n$ clauses. Then one can prove the satisfiability of $\varphi$, with perfect completeness and constant*

[1] Indeed, if the soundness gap were constant rather than $1/\operatorname{poly}(n)$, then Blier and Tapp's protocol could presumably be "scaled up by an exponential" to show $\mathsf{QMA}(2) = \mathsf{NEXP}$!

[2] For Arthur could first use the in-place amplification of Marriott and Watrous [15] to make his error probability exponentially small (without increasing the size of the quantum proof $|\psi\rangle$), and then use Grover search to find $|\psi\rangle$ in $2^{o(n)}$ time.

[3] A first reason is that it is unclear how to do in-place amplification of $\mathsf{QMA}(k)$ protocols. A second reason is that, even *assuming* amplification, it is unclear how to search efficiently among unentangled witnesses. In Section 5, we will show that the first reason is actually the crucial one.

*soundness, using $O\left(\sqrt{m}\operatorname{polylog} m\right)$ unentangled quantum proofs, each with $O\left(\log m\right)$ qubits.*

In particular, if $m = O\left(n\right)$,[4] then we get an almost-quadratic improvement over the witness size needed in the classical world (or that matter, in the quantum world with one prover).

We now explain the intuition behind Theorem 1. The first step in our protocol is to reduce 3SAT to a more convenient problem called 2-OUT-OF-4-SAT, where every clause has exactly four literals, and is satisfied if and only if exactly two of the literals are. We also want our 2-OUT-OF-4-SAT instance to be a PCP: that is, either it should be satisfiable, or else at most a $1-\varepsilon$ fraction of clauses should be satisfiable for some constant $\varepsilon > 0$. Finally we want the instance to be *balanced*, meaning that every variable occurs in at most a constant number of clauses. Fortunately, we can get all of this via known classical reductions, including the "tight" PCP Theorem of Dinur [9], which increase the number of variables and clauses by at most an $O\left(\operatorname{polylog} n\right)$ factor.

So suppose Arthur has applied these reductions, to obtain a balanced 2-OUT-OF-4-SAT PCP instance $\phi$ with $n$ variables. And now suppose Merlin sends Arthur a $\log n$-qubit quantum state of the form

$$\left|\psi\right\rangle = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(-1\right)^{x_i}\left|i\right\rangle,$$

where $x_1,\ldots,x_n \in \{0,1\}^n$ is the claimed satisfying assignment for $\phi$. (We call a state having the above form a *proper* state.) Then we show that Arthur can check the veracity of $x_1,\ldots,x_n$ with perfect completeness and constant soundness. To do so, Arthur simply measures $\left|\psi\right\rangle$ in a basis corresponding to the clauses of $\phi$. With constant probability, he will get an outcome of the form

$$\left(-1\right)^{x_i}\left|i\right\rangle + \left(-1\right)^{x_j}\left|j\right\rangle + \left(-1\right)^{x_k}\left|k\right\rangle + \left(-1\right)^{x_\ell}\left|\ell\right\rangle$$

where $(i,j,k,\ell)$ is a randomly chosen clause of $\phi$. Assuming this occurs, Arthur can perform a measurement that accepts with certainty if $x_i + x_j + x_k + x_\ell = 2$ and rejects with constant probability otherwise.

Thus, if only Arthur could somehow assume $\left|\psi\right\rangle$ was proper, we would have a $\log n$-qubit witness for 3SAT! The problem, of course, is that Arthur has no way of knowing whether Merlin has cheated and given him an improper state. For example, what if Merlin concentrates the amplitude of $\left|\psi\right\rangle$ on some small subset of basis states, and simply omits the other basis states?

Our key technical contribution is to show that, if Arthur gets not one but $O(\sqrt{n})$ copies of $\left|\psi\right\rangle$, then he can check with constant soundness whether $\left|\psi\right\rangle$ is proper or far from any proper state. Indeed, even if Arthur is given $K = O(\sqrt{n})$ states $\left|\varphi_1\right\rangle,\ldots,\left|\varphi_K\right\rangle$ which are not necessarily identical, so long as the states are not entangled with each other Arthur can check with constant soundness whether most of them are close to some proper state $\left|\psi\right\rangle$. This then yields a protocol for 3SAT with constant soundness and $O(\sqrt{n})$ unentangled proofs of size $O\left(\log n\right)$—for Arthur can just choose randomly whether to perform the satisfiability test described above, or to check whether most of the $\left|\varphi_k\right\rangle$'s are close to some proper state $\left|\psi\right\rangle$.

To check that most of the states are at least close to *each other*, Arthur simply has to perform a "swap test" between (say) $\left|\varphi_1\right\rangle$ and a random other state $\left|\varphi_k\right\rangle$. So the problem is reduced to the following: assuming most of the $\left|\varphi_k\right\rangle$'s are close to $\left|\varphi_1\right\rangle$, how can Arthur decide whether $\left|\varphi_1\right\rangle$ is proper or far from any proper state?

In our protocol, Arthur does this by first choosing a matching $\mathcal{M}$ on the set $\{1,\ldots,n\}$ uniformly at random. He then measures each state $\left|\varphi_k\right\rangle$ in an orthonormal basis that contains the vectors $\left|i\right\rangle + \left|j\right\rangle$ and $\left|i\right\rangle - \left|j\right\rangle$ for every edge $(i,j) \in \mathcal{M}$.

Let us think about what happens when Arthur does this. Since he is performing $O(\sqrt{n})$ measurements on almost-identical states, and since each measurement has $n$ possible outcomes, by using a suitable generalization of the Birthday Paradox, one can prove that with $\Omega\left(1\right)$ probability, Arthur will find a *collision*: that is, two outcomes of the form $\left|i\right\rangle \pm \left|j\right\rangle$, for the same edge $(i,j) \in \mathcal{M}$. So suppose this happens. Then if the $\left|\varphi_k\right\rangle$'s are all equal to a proper state $\left|\psi\right\rangle = \sum_{i=1}^{n}\left(-1\right)^{x_i}\left|i\right\rangle$, the two outcomes will clearly "agree": that is, they will either both be $\left|i\right\rangle+\left|j\right\rangle$ (if $x_i = x_j$) or both be $\left|i\right\rangle - \left|j\right\rangle$ (if $x_i \neq x_j$). On the other hand, suppose the $\left|\varphi_k\right\rangle$'s are far from any proper state. In that case, we show that the outcomes will "disagree" (that is, one will be $\left|i\right\rangle + \left|j\right\rangle$ and the other will be $\left|i\right\rangle - \left|j\right\rangle$) with $\Omega\left(1\right)$ probability.

To understand why, consider that there are two ways for a state $\left|\varphi\right\rangle = \sum_{i=1}^{n}\alpha_i\left|i\right\rangle$ to be far from proper. First, the probability distribution $\left(\left|\alpha_1\right|^2,\ldots,\left|\alpha_n\right|^2\right)$, which corresponds to measuring $\left|\varphi\right\rangle$ in the standard basis, could be far from the uniform distribution. Second, the $\alpha_i$'s could be roughly equal in magnitude, but they could have complex phases that cause $\left|\varphi\right\rangle$ to be far from any state involving positive and negative real amplitudes only. In either case, though, if Arthur measures according to a random matching $\mathcal{M}$, then with high probability he will obtain an outcome $\alpha_i\left|i\right\rangle + \alpha_j\left|j\right\rangle$ that is not close to either $\left|i\right\rangle+\left|j\right\rangle$ or $\left|i\right\rangle-\left|j\right\rangle$.

As one would imagine, making all of these claims quantitative and proving them requires a good deal of work.

The reason we need the assumption of unentanglement is that without it, cheating Merlins might correlate their states in such a way that a swap test between any two states passes

[4] Note that setting $m = O\left(n\right)$ is fairly common in the study of 3SAT, and indeed, the "hardest" random 3SAT instances are believed to occur around $m \approx 4.25n$.

with certainty, and yet no collisions are ever observed. As we point out in Section 3.5, it seems unlikely that the assumption of unentanglement can be removed, since this would lead to a $2^{\widetilde{O}(\sqrt{n})}$-time classical algorithm for 3SAT. On the other hand, we believe it should be possible to improve our protocol to one involving only *two* unentangled proofs. This is a problem we leave to future work.

## 1.4 Additivity Implies Amplification

In the one-prover case, it is easy to amplify a QMA protocol with constant error to a protocol with exponentially small error. Merlin simply sends Arthur $m = \operatorname{poly}(n)$ copies of his proof; then Arthur checks each of the copies and outputs the majority answer. To show that this works, the key observation is that *Merlin cannot gain anything by entangling the $m$ proofs*. Indeed, because of the convexity of Arthur's acceptance probability, Merlin might as well send Arthur an unentangled state $|\psi\rangle^{\otimes m}$, in which case the completeness and soundness errors will decrease exponentially with $m$ by the usual Chernoff bound.

Now suppose we try the same argument for $\mathsf{QMA}(2)$. If we ask each Merlin to send $m$ copies of his state, each Merlin might cheat by instead sending an entangled state on $m$ registers. And in that case, as soon as Arthur checks the first copy (consisting of one register from Merlin$_A$ and one from Merlin$_B$), *his doing so might create entanglement in the remaining copies where there was none before!* This is because of a counterintuitive phenomenon called *entanglement swapping* [21], by which two quantum systems that have never interacted in the past can nevertheless become entangled, provided those systems are entangled with *other* systems on which an entangling measurement is performed.

Let us give a small illustration of this phenomenon. Suppose that each "proof" is a single qubit, and that Arthur asks for two proofs from each Merlin (thus, 4 qubits in total). Then if Merlin$_A$ is dishonest, he might send Arthur the entangled state $|\psi_A\rangle = |00\rangle + |11\rangle$, and likewise Merlin$_B$ might send Arthur $|\psi_B\rangle = |00\rangle + |11\rangle$ (omitting normalization). Now suppose Arthur measures the qubits $|\psi_A\rangle_{(1)}$ and $|\psi_B\rangle_{(1)}$ in the "Bell basis," consisting of the four entangled states $|00\rangle + |11\rangle$, $|00\rangle - |11\rangle$, $|01\rangle + |10\rangle$, and $|01\rangle - |10\rangle$. Then conditioned on the outcome of this measurement, it is not hard to see that the joint state of $|\psi_A\rangle_{(2)}$ and $|\psi_B\rangle_{(2)}$ will also be entangled.[5]

Of course, as soon as the remaining $m-1$ copies become entangled, we lose our soundness guarantee and the proof of amplification fails.

Nevertheless, there is a natural amplification procedure that seems like it *ought* to be robust against such "pathological" behavior. Suppose Arthur chooses the number of copies $m$ to be very large, say $n^{10}$ (much larger than the number of copies he is actually going to check), and suppose that each copy he *does* check is chosen uniformly at random. Then whatever entanglement Arthur produces during the checking process ought be "spread out" among the copies, so that with high probability, every copy that Arthur actually encounters is close to separable.

It follows, from the "finite quantum de Finetti theorem" of König and Renner [13], that if the number of copies were large enough then the above argument would work. Unfortunately, the number of copies needs to be exponential in $n$ for that theorem to apply.

We will show that the argument works with $\operatorname{poly}(n)$ copies, provided one can formalize terms like "spread out" and "close to separable" using a suitable measure of entanglement. The only problem, then, is that a measure of entanglement with the properties we need is not yet known to exist! Informally, we need an entanglement measure $E$ that

(i) is *superadditive* (meaning it "spreads itself out" among registers), and

(ii) is *faithful* (meaning if $E(\rho)$ is polynomially small then $\rho$ is polynomially close to a separable state in trace distance).

Among existing entanglement measures, there is one—the *entanglement of formation* $E_F$, introduced by Bennett et al. [5]—that is known to satisfy (ii), and is conjectured to satisfy (i).[6] This conjecture is known to be equivalent to the Additivity Conjecture from quantum information theory [18].

Our first result says that, if the Additivity Conjecture holds, then any $\mathsf{QMA}(2)$ protocol can be amplified to exponentially small error. We also prove that any $\mathsf{QMA}(k)$ protocol with constant soundness can be simulated by a $\mathsf{QMA}(2)$ protocol with $\Omega(1/k)$ soundness. Combining these two results, we find that if the Additivity Conjecture holds, then $\mathsf{QMA}(k) = \mathsf{QMA}(2)$ for all $k \geq 2$.

Two other interesting consequences we get are the following. First, assuming the Additivity Conjecture, two Merlins who share $h(n)$ ebits of entanglement can simulate two unentangled Merlins, for every fixed polynomial $h$. In other words, a bounded amount of entanglement gives the Merlins no additional power to cheat. Second, again assuming the Additivity Conjecture, $k$ Merlins who all send copies of the same witness yield the same computational power as $k$ Merlins who can send different witnesses.

[5] Indeed, this example can be seen as a special case of *quantum teleportation* [4]: Arthur uses the entanglement between Merlin$_A$'s left and right registers, as well as between Merlin$_B$'s left and right registers, to teleport an entangled state into the two right registers by acting only on the two left registers.

[6] There is also another measure—the *squashed entanglement* $E_{sq}$, introduced by Christandl and Winter [8]—that is known to satisfy (i), but unfortunately can be shown *not* to satisfy (ii).

### 1.5 Evidence That $\mathsf{QMA}(k) \subseteq \mathsf{PSPACE}$

It is well-known that $\mathsf{QMA} \subseteq \mathsf{PP}$ [15]. On the other hand, the only known upper bound for $\mathsf{QMA}(2)$ is the trivial $\mathsf{NEXP}$, and improving this (even to $\mathsf{QMA}(2) \subseteq \mathsf{EXP}$) has been an open problem for several years. In this paper we show that $\mathsf{QMA}(2) \subseteq \mathsf{PSPACE}$, assuming what we call the Strong Amplification Conjecture: that is possible to amplify any $\mathsf{QMA}(k)$ protocol, in such a way that one of the Merlin's Hilbert space dimensions remains small compared to the inverse of the error bound. Note that, since strong amplification *also* implies $\mathsf{QMA}(k) = \mathsf{QMA}(2)$ for all $k \geq 2$, we then get $\mathsf{QMA}(k) \subseteq \mathsf{PSPACE}$ as well.

Our proof is based on an idea called "de-Merlinization," which was previously used by Aaronson [1] to show $\mathsf{QMA/qpoly} \subseteq \mathsf{PSPACE/poly}$. We show that if strong amplification holds, then Arthur can "partially de-Merlinize" any $\mathsf{QMA}(2)$ protocol—that is, remove one of the Merlins from the picture—at the cost of an exponential increase in running time. We then have $\mathsf{QMA}(2) \subseteq \mathsf{QMA}_{\mathsf{PSPACE}}$, where $\mathsf{QMA}_{\mathsf{PSPACE}}$ is the version of $\mathsf{QMA}$ where Arthur runs in quantum polynomial *space* instead of quantum polynomial time. But it follows from results of Watrous [19] that $\mathsf{QMA}_{\mathsf{PSPACE}} = \mathsf{BQPSPACE} = \mathsf{PSPACE}$.

### 1.6 Nonexistence of Perfect Disentanglers

While we now have a few examples where multiple quantum proofs seem to help—such as the 3SAT protocol of this paper, and the pure state $N$-representability problem [14]—we still have no "complexity-theoretic" evidence that $\mathsf{QMA}(2) \neq \mathsf{QMA}$. Indeed, even proving an oracle separation between $\mathsf{QMA}(2)$ and $\mathsf{QMA}$ seems extremely difficult.

Thus, let us consider the other direction and try to prove these classes are the same. Potentially the first approach would be to equip Arthur with a *disentangler*: that is, a quantum operation that would convert Merlin's (possibly-entangled) witness into a separable witness, and thereby let Arthur simulate a $\mathsf{QMA}(2)$ protocol in $\mathsf{QMA}$. In this paper we take a first step in the study of disentanglers, by proving that in finite-dimensional Hilbert spaces, there is no operation that produces all and only the separable states as output.

Note that, if we are willing to settle for there being an output *close* to every separable state, then a disentangler does exist: for example, take as input a classical description of the separable state $\sigma$ to be prepared, measure that description in the computational basis, and then prepare $\sigma$.[7] The key problem is that the input Hilbert space needs to be exponentially larger than the output Hilbert space. Watrous (personal communication) has conjectured that this exponentiality is an unavoidable feature of any approximate disentangler; proving or disproving this remains one of the central open problems about $\mathsf{QMA}(2)$.

[7] This argument also shows that our result fails if the input Hilbert space is infinite-dimensional—for then one could give an infinitely-precise description of $\sigma$.

## 2 Preliminaries

In this section, we first define the complexity class $\mathsf{QMA}(k, a, b)$, or Quantum Merlin-Arthur with $k$ unentangled witnesses and error bounds $a, b$, and state some basic facts and conjectures about this class. We then survey some concepts from quantum information theory we will need, including trace distance and the swap test.

### 2.1 Multiple-Prover QMA

**Definition 2.** *A language $L$ is in $\mathsf{QMA}(k, a, b)$ if there exists a polynomial-time quantum algorithm $Q$ such that for all inputs $x \in \{0,1\}^n$:*

*(i) If $x \in L$ then there exist witnesses $|\psi_1\rangle, \ldots, |\psi_k\rangle$, with $\mathrm{poly}(n)$ qubits each, such that $Q$ accepts with probability at least $b$ given $|x\rangle \otimes |\psi_1\rangle \otimes \cdots \otimes |\psi_k\rangle$.*

*(ii) If $x \notin L$ then $Q$ accepts with probability at most $a$ given $|x\rangle \otimes |\psi_1\rangle \otimes \cdots \otimes |\psi_k\rangle$, for all $|\psi_1\rangle, \ldots, |\psi_k\rangle$.*

*As a convention, we also define* $\mathsf{QMA}(k) := \mathsf{QMA}(k, 1/3, 2/3)$*, and* $\mathsf{QMA} := \mathsf{QMA}(1)$.[8]

The above definition makes sense for all integers $k$ from 1 up to $\mathrm{poly}(n)$, and nonnegative real functions $2^{-\mathrm{poly}(n)} \leq a(n) < b(n) \leq 1 - 2^{-\mathrm{poly}(n)}$.

In the one-prover case, we know that $\mathsf{QMA}(1, 1/3, 2/3) = \mathsf{QMA}\left(1, 2^{-p(n)}, 1 - 2^{-p(n)}\right)$ for all polynomials $p$ (see [15] for example). This is what justifies the convention $\mathsf{QMA}(1) := \mathsf{QMA}(1, 1/3, 2/3)$. By contrast, we do not yet know whether the convention $\mathsf{QMA}(k) := \mathsf{QMA}(k, 1/3, 2/3)$ is justified for $k \geq 2$. That it *is* justified is the content of the following conjecture:

**Conjecture 3** (Amplification)**.** *For all $k$, all $a < b$ with $b - a = \Omega(1/\mathrm{poly}(n))$, and all polynomials $p$,* $\mathsf{QMA}(k, a, b) = \mathsf{QMA}\left(k, 2^{-p(n)}, 1 - 2^{-p(n)}\right)$.

One is tempted to make an even stronger conjecture: that the entire hierarchy of $\mathsf{QMA}(k, a, b)$'s we have defined collapses to just two complexity classes, namely $\mathsf{QMA}$ and $\mathsf{QMA}(2)$.

**Conjecture 4** (Collapse)**.** *For all $k \geq 2$, all $a < b$ with $b - a = \Omega(1/\mathrm{poly}(n))$, and all polynomials $p$,* $\mathsf{QMA}(k, a, b) = \mathsf{QMA}\left(2, 2^{-p(n)}, 1 - 2^{-p(n)}\right)$.

[8] For purposes of definition, we assume we have fixed a specific machine model (e.g., a universal set of quantum gates)—though if the Amplification Conjecture to be discussed shortly holds, then this choice will turn out not to matter.

The main progress so far on these conjectures has been due to Kobayashi et al. [12], who showed that the Amplification and Collapse Conjectures are actually equivalent:

**Theorem 5** (**[12]**). *Conjecture 3 implies Conjecture 4.*

Let us observe that one can make the *completeness* error (though not the soundness error) exponentially small, using a simple argument based on Markov's inequality. We will need this observation in Section 4.

**Lemma 6.**
$\mathsf{QMA}(k,a,b) \subseteq \mathsf{QMA}\left(k, 1-(b-a), 1-2^{-p(n)}\right)$ *for all $k$, all $a<b<1$, and all polynomials $p$.*

*Proof.* We use the following protocol. Each Merlin provides $m = C \cdot \frac{p(n)}{(b-a)^2}$ registers for some constant $C$. Then Arthur runs his verification procedure $m$ times in parallel, once with each $k$-tuple of registers, and accepts if and only if at least a $d$ fraction of invocations accept, for some $d$ slightly less than $b$.

To show completeness, we use a Chernoff bound. Assuming the Merlins are honest, each one simply provides $m$ copies of his witness. Then on each invocation, Arthur accepts with independent probability at least $b$. So assuming we chose a sufficiently large constant $C$, the probability that Arthur accepts less than $dm$ times is at most $2^{-p(n)}$.

To show soundness, we use Markov's inequality. The expected number of accepting invocations is at most $am$ (by linearity, this is true even if the registers are entangled). Hence the probability that this number exceeds $dm$ is at most $a/d$, which we can ensure is less than $1-(b-a)$ by choosing $d$ sufficiently close to $b$ (and using the fact that $b<1$). $\square$

## 2.2 Quantum Information

We now review some quantum information concepts that we will need. For more details see Nielsen and Chuang [16].

Given two mixed states $\rho$ and $\sigma$, their *trace distance* is $\|\rho-\sigma\|_{\text{tr}} := \frac{1}{2}\sum_{i=1}^{n}|\lambda_i|$, where $(\lambda_1,\ldots,\lambda_n)$ are the eigenvalues of $\rho-\sigma$. We will sometimes say $\sigma$ is *$\varepsilon$-close* to $\rho$ if $\|\rho-\sigma\|_{\text{tr}} \le \varepsilon$, and *$\varepsilon$-far* otherwise. The importance of trace distance comes from the following fact:

**Proposition 7.** *Suppose $\sigma$ is $\varepsilon$-close to $\rho$. Then any measurement that accepts $\rho$ with probability $p$, accepts $\sigma$ with probability at most $p+\varepsilon$.*

Given a pure state $|\psi\rangle$ and a mixed state $\rho$, their *squared fidelity* $\langle\psi|\rho|\psi\rangle$ is the probability of obtaining $|\psi\rangle$ as the result of a projective measurement on $\rho$. Squared fidelity behaves nicely under tensor products:

**Proposition 8.** *Given a $k$-partite state $\rho^{A_1A_2\cdots A_k}$, suppose there are pure states $|\psi_1\rangle,\ldots,|\psi_k\rangle$ such that $\langle\psi_i|\rho^{A_i}|\psi_i\rangle \ge 1-\varepsilon_i$ for all $i$. Let $|\Psi\rangle := |\psi_1\rangle\otimes\cdots\otimes|\psi_k\rangle$ and $\varepsilon := \varepsilon_1+\cdots+\varepsilon_k$. Then $\langle\Psi|\rho^{A_1A_2\cdots A_k}|\Psi\rangle \ge 1-\varepsilon$.*

Trace distance and squared fidelity are related to each other as follows:

**Proposition 9.** $\langle\psi|\rho|\psi\rangle + \|\rho-|\psi\rangle\langle\psi|\|_{\text{tr}}^2 \le 1$ *for all $\rho$ and $|\psi\rangle$.*

Given a product state $\rho\otimes\sigma$, the *swap test* is a quantum operation that measures the overlap between $\rho$ and $\sigma$. The test accepts with probability $\frac{1+\text{tr}(\rho\sigma)}{2}$ and rejects otherwise. The swap test can also reveal information about the purity of a state, as follows:

**Proposition 10.** *Suppose $\langle\psi|\rho|\psi\rangle < 1-\varepsilon$ for all pure states $|\psi\rangle$. Then a swap test between $\rho$ and any other state rejects with probability greater than $\varepsilon/2$.*

# 3 Proving 3SAT With $\widetilde{O}(\sqrt{n})$ Qubits

We now present our protocol for proving the satisfiability of a 3SAT instance, using $\widetilde{O}(\sqrt{n})$ unentangled quantum proofs with $O(\log n)$ qubits each. For ease of presentation, the protocol will be broken into four steps: first, classical reductions from the 3SAT problem to a different NP-complete problem that we will actually use; second, a protocol for the special case where the witness is "proper"; third, a protocol for the case where the Merlins send Arthur $\widetilde{O}(\sqrt{n})$ witnesses, which are not necessarily proper but which are guaranteed to be identical; and fourth, a protocol for the general case. We end in Section 3.5 with some general observations about our protocol and the prospects for improving it further.

## 3.1 Classical Reductions

It will be convenient to work not with 3SAT but with a related problem called 2-OUT-OF-4-SAT, in which every clause has exactly four literals, and is satisfied if and only if exactly two of the literals are. We will also need our 2-OUT-OF-4-SAT instance to be a PCP, and to have every variable appear in at most $O(1)$ clauses. The following lemma shows how to get everything we want with only a polylogarithmic blowup in the number of variables and clauses.

**Lemma 11.** *There exists a polynomial-time Karp reduction that maps a 3SAT instance $\varphi$ to a 2-OUT-OF-4-SAT instance $\phi$, and that has the following properties:*

*(i) If $\varphi$ has $n$ variables and $m \ge n$ clauses, then $\phi$ has $O(m \operatorname{polylog} m)$ variables and $O(m \operatorname{polylog} m)$ clauses.*

*(ii) Every variable of $\phi$ occurs in at most $c$ clauses, for some constant $c$.*

*(iii) The reduction is a PCP (meaning that satisfiable instances map to satisfiable instances, while unsatisfiable instances map to instances that are $\varepsilon$-far from satisfiable for some constant $\varepsilon > 0$).*

*Proof.* Given a 3SAT instance $\varphi$, we first amplify its soundness gap to a constant using the celebrated method of Dinur [9]. Next we use a reduction due to Papadimitriou and Yannakakis [17], which makes every variable occur in exactly 29 clauses, while weakening the soundness gap by only a constant factor. Finally we use a gadget due to Khanna et al. [10], which converts from 3SAT to 2-OUT-OF-4-SAT, while decreasing the soundness gap and increasing the number of clauses per variable by at most constant factors. Note that the reduction of Dinur [9] incurs only a polylogarithmic blowup in the number of variables and clauses, while the other two reductions incur a constant blowup. □

### 3.2 The Proper State Case

Suppose Arthur has applied Lemma 11, to obtain a balanced 2-OUT-OF-4-SAT instance $\phi$ with $N = O(m \operatorname{polylog} m)$ variables, $M = O(m \operatorname{polylog} m)$ clauses, and a constant soundness gap $\varepsilon > 0$. And now suppose Merlin sends Arthur a $\log N$-qubit state of the form

$$|\psi\rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (-1)^{x_i} |i\rangle ,$$

where $x_1, \ldots, x_N \in \{0,1\}^N$ is a claimed satisfying assignment for $\phi$. Call a state having the above form (for some Boolean $x_i$'s) a *proper* state. Then we claim the following:

**Lemma 12.** *Assuming $|\psi\rangle$ is proper, Arthur can check whether $\phi$ is satisfiable with perfect completeness and constant soundness.*

*Proof.* To perform the check, Arthur uses the following *Satisfiability Test*. First he partitions the clauses of $\phi$ into a constant number of blocks $B_1, \ldots, B_s$, such that within each block, no two clauses share a variable. Such a partition clearly exists by the assumption that $\phi$ is balanced, and furthermore can be found efficiently (e.g., using a greedy algorithm). Next he chooses one of the blocks $B_r$ uniformly at random, and measures $|\psi\rangle$ in an orthonormal basis with one projector for each clause in $B_r$. Because a single block in the partition of clauses does not necessarily cover all the variables, it is possible that the measurement result will not correspond to any clause in $B_r$, in which case Arthur accepts. However, suppose that the measurement yields the following reduced state, for some random clause $C_{ijk\ell} := (i, j, k, \ell)$ in $B_r$:

$$|\psi_{ijkl}\rangle := (-1)^{x_i} |i\rangle + (-1)^{x_j} |j\rangle + (-1)^{x_k} |k\rangle + (-1)^{x_\ell} |\ell\rangle .$$

Notice that, of the 16 possible assignments to the variables $(x_i, x_j, x_k, x_\ell)$, six of them satisfy $C_{ijk\ell}$, and those six lead to three states $|\psi_{ijk\ell}\rangle$ that are orthogonal to one another (as well as the negations of those states, which are essentially the same). It follows that Arthur can perform a projective measurement on $|\psi_{ijk\ell}\rangle$, which accepts with probability 1 if $C_{ijk\ell}$ is satisfied, and rejects with constant probability if $C_{ijk\ell}$ is unsatisfied.

Furthermore, because the number of blocks $B_r$ is a constant, each of the $M$ clauses of $\phi$ is checked in this test with probability $\Omega(1/M)$. And we know that, if $x_1, \ldots, x_N$ is *not* a satisfying assignment for $\phi$, then a constant fraction of the clauses will be unsatisfied. Putting everything together, we find that if $\phi$ is satisfiable, then the Satisfiability Test accepts $|\psi\rangle$ with probability 1; while if $\phi$ is unsatisfiable, then it rejects with constant probability. □

### 3.3 The Symmetric Case

Thus, the problem we need to solve is "merely" how to force Merlin to send a proper state. For example, how can Arthur prevent a cheating Merlin from concentrating the amplitude of $|\psi\rangle$ on some subset of basis states for which the Satisfiability Test accepts, and omitting the other basis states?

To solve this problem, Arthur is going to need more Merlins. In particular, let us suppose there are $K = \Theta(\sqrt{N})$ unentangled Merlins, who send Arthur $\log N$-qubit states $|\varphi_1\rangle, \ldots, |\varphi_K\rangle$ respectively. By convexity, we can assume without loss of generality that these states are pure. For the time being, we also assume that the states are identical; that is, $|\varphi_i\rangle = |\varphi\rangle$ for all $i \in [K]$. Given these states, Arthur performs one of the following two tests, each with probability $1/2$:

**Satisfiability Test:** *Arthur chooses any copy of $|\varphi\rangle$, and performs the Satisfiability Test described in Section 3.2.*

**Uniformity Test:** *Arthur chooses a matching $\mathcal{M}$ on $[N]$ uniformly at random. He then measures each copy of $|\varphi\rangle$ in an orthonormal basis, which contains the vectors $|i\rangle + |j\rangle, |i\rangle - |j\rangle$ for every edge $(i, j) \in \mathcal{M}$. If for some $(i, j) \in \mathcal{M}$, the two outcomes $|i\rangle + |j\rangle$ and $|i\rangle - |j\rangle$ both occur among the $K$ measurement outcomes, then Arthur rejects. Otherwise he accepts.*

It is clear that the above protocol has perfect completeness. For if $\phi$ is satisfiable, then the Merlins can just send $K$ copies of a proper state $|\psi\rangle$ corresponding to a satisfying assignment for $\phi$. In that case, both tests will accept with probability 1. Our goal is to prove the following:

**Theorem 13.** *The protocol has constant soundness (again, assuming the $|\varphi_i\rangle$'s are all identical).*

To prove Theorem 13, we need to show that if $\phi$ is unsatisfiable, then one of the two tests rejects with constant probability. There are two cases. First suppose $|\varphi\rangle$ is $\varepsilon$-close in trace distance to some proper state $|\psi\rangle$. Then provided we choose $\varepsilon > 0$ sufficiently small, Lemma 12, combined with Proposition 7, already implies that the Satisfiability Test rejects with constant probability. So our task reduces to proving the following:

**Claim 14.** *Suppose $|\varphi\rangle$ is $\varepsilon$-far in trace distance from any proper state $|\psi\rangle$, for some $\varepsilon > 0$. Then the Uniformity Test rejects with some constant probability $\delta(\varepsilon) > 0$.*

In analyzing the Uniformity Test, we say that Arthur *finds a collision* if he obtains two measurement outcomes of the form $|i\rangle \pm |j\rangle$ for the same $(i, j)$ pair, and that he *finds a disagreement* if one of the outcomes is $|i\rangle + |j\rangle$ and the other is $|i\rangle - |j\rangle$. Of course, finding a disagreement is what causes him to reject.

The first step, though, is to lower-bound the probability that Arthur finds a collision. Let $|\varphi\rangle = \alpha_1 |1\rangle + \cdots + \alpha_N |N\rangle$. Then for every copy of $|\varphi\rangle$ and every edge $(i, j) \in \mathcal{M}$, Arthur measures an outcome of the form $|i\rangle \pm |j\rangle$ with probability $|\alpha_i|^2 + |\alpha_j|^2$, and these outcomes are independent from one copy to the next. We are interested in the probability that, for some $(i, j)$ pair, Arthur measures $|i\rangle \pm |j\rangle$ more than once. But this is just the famous Birthday Paradox, with $K = \Theta(\sqrt{N})$ "people" (the copies of $|\varphi\rangle$) and $N/2$ "days" (the edges in $\mathcal{M}$). The one twist is that the distribution over birthdays need not be uniform. However, a result of Bloom and Knight [7] shows that the Birthday Paradox occurs in the nonuniform case as well.

Therefore Arthur finds a collision with constant probability. The hard part is to show that he finds a *disagreement* with constant probability. Here, of course, we have to use the fact that $|\varphi\rangle$ is $\varepsilon$-far from proper.

For now, let us restrict attention to two copies of $|\varphi\rangle$. For each edge $(i, j) \in \mathcal{M}$, define the "disagreement probability"

$$p_{ij} = \frac{|\alpha_i + \alpha_j|^2 |\alpha_i - \alpha_j|^2}{2\left(|\alpha_i|^2 + |\alpha_j|^2\right)^2}$$

to be the probability that, conditioned on measuring two outcomes of the form $|i\rangle \pm |j\rangle$, one of the outcomes is $|i\rangle + |j\rangle$ and the other one is $|i\rangle - |j\rangle$. Also, say an edge $(i, j) \in \mathcal{M}$ is *$c$-unbalanced with respect to* $|\varphi\rangle$ if $p_{ij} \geq c$, and let $\mathcal{S}_c \subseteq \mathcal{M}$ be the set of $c$-unbalanced edges. Say a set of edges $\mathcal{S} \subseteq \mathcal{M}$ is *$d$-large with respect to* $|\varphi\rangle$ if

$$\sum_{(i,j)\in\mathcal{S}} \left(|\alpha_i|^2 + |\alpha_j|^2\right) \geq d.$$

> **The 3SAT Protocol**
>
> Given $|\varphi_1\rangle, \ldots, |\varphi_K\rangle$, Arthur performs one of the following three tests, each with probability $1/3$.
>
> **Satisfiability Test:** Arthur applies the Satisfiability Test, described in Section 3.2, to $|\varphi_1\rangle$.
>
> **Symmetry Test:** Arthur chooses an index $k \in \{2, \ldots, K\}$ uniformly at random, performs a swap test between $|\varphi_1\rangle$ and $|\varphi_k\rangle$, and accepts if and only if the swap test accepts.
>
> **Uniformity Test:** Arthur chooses a matching $\mathcal{M}$ on $[N]$ uniformly at random. He then measures each $|\varphi_k\rangle$ in an orthonormal basis, which contains the vectors
>
> $$\frac{|i\rangle + |j\rangle}{\sqrt{2}}, \frac{|i\rangle - |j\rangle}{\sqrt{2}}$$
>
> for every edge $(i, j) \in \mathcal{M}$. If for some $(i, j) \in \mathcal{M}$, the two outcomes $\frac{|i\rangle+|j\rangle}{\sqrt{2}}$ and $\frac{|i\rangle-|j\rangle}{\sqrt{2}}$ both occur among the $K$ measurement outcomes, then Arthur rejects. Otherwise he accepts.

Then the key fact is the following:

**Theorem 15.** *Suppose $|\varphi\rangle$ is $\varepsilon$-far in trace distance from any proper state. Then $\mathcal{S}_c$ is $d$-large with respect to $|\varphi\rangle$ with probability at least $1/3$ over the choice of $\mathcal{M}$, for some constants $c$ and $d$ depending on $\varepsilon$.*

The proof of Theorem 15 is deferred to the full version.

Assuming Theorem 15, we can complete the proof of Claim 14, and hence of Theorem 13. The idea is this: when Arthur performs the Uniformity Test, simply discard all measurement outcomes that are not of the form $|i\rangle \pm |j\rangle$ for some $(i, j) \in \mathcal{S}_c$. Assuming $\mathcal{S}_c$ is $d$-large—which it is with constant probability by Theorem 15—with overwhelming probability that still leaves $\Theta(\sqrt{N})$ "good" measurement outcomes. Then by the Birthday Paradox, with constant probability there will be a collision among these good outcomes. And by the definition of $\mathcal{S}_c$, any such collision will also be a disagreement with constant probability, thereby causing Arthur to reject.

## 3.4 The General Case

Of course, in general the states $|\varphi_1\rangle, \ldots, |\varphi_K\rangle$ sent by the $K = \Theta(\sqrt{N})$ Merlins need not be identical. To deal with this, we now give our final protocol (see box), which removes the symmetry restriction.

It is clear that the protocol has perfect completeness, and thus the problem is to show soundness: that is, if $\phi$ is unsatisfiable, then one of the three tests rejects with constant probability. There are three cases.

The first case is that $|\varphi_1\rangle$ is $\varepsilon$-close to some proper state $|\psi\rangle$. Then as before, the Satisfiability Test will reject with constant probability, provided we choose $\varepsilon$ sufficiently small.

The second case is that $|\langle\varphi_1|\varphi_k\rangle| < 1-\delta$ for at least a $\gamma$ fraction of indices $k \in \{2,\ldots,K\}$. In that case it is clear that the Symmetry Test will reject with probability at least $\gamma\delta/2$.

The third case is that $|\langle\varphi_1|\varphi_k\rangle| \geq 1-\delta$ for more than a $1-\gamma$ fraction of indices $k \in \{2,\ldots,K\}$, but nevertheless $|\varphi_1\rangle$ is $\varepsilon$-far from any proper state. In this case we need to generalize the results of the previous section, to show that the Uniformity Test will still reject with constant probability (dependent on $\varepsilon$, $\delta$, and $\gamma$).

The details of this generalization are deferred to the full version. Here, we will just mention one key ingredient, which is to generalize the Birthday Paradox further, to the case where the birthday distributions are not only nonuniform but can also differ from each other by small amounts. In particular we want the following:

**Theorem 16.** *Let $X_1,\ldots,X_K$ be independent random variables over $[N]$, and let $\mathcal{D}_i$ be the distribution over $X_i$. Suppose $K \geq 6\sqrt{N}$ and $\|\mathcal{D}_i - \mathcal{D}_j\| \leq 1/10$ for all $i,j$. Then*

$$\Pr[\exists i,j : X_i = X_j] \geq \frac{1}{2}.$$

In the full version, we present a proof of Theorem 16 based on the second moment method. (Indeed, our proof works even if the $X_i$'s are only 4-wise independent.)

The bottom line is that we get a protocol with perfect completeness, constant soundness, and $\widetilde{O}(\sqrt{m})$ unentangled witnesses with $O(\log m)$ qubits each.

As a final remark, we can amplify the soundness error to $1/p(m)$ for any desired polynomial $p$. To do so, we simply multiply the number of Merlins by a further $\operatorname{polylog} m$ factor, and repeat the whole protocol $\operatorname{polylog} m$ times.

### 3.5 General Observations

We conclude this subsection by making four general observations about Theorem 1.

First, we strongly believe that our protocol can be improved to one involving two provers, one of whom sends $O(\log m)$ qubits and the other of whom sends $O(\sqrt{m}\operatorname{polylog} m)$ qubits. Specifically, if all but one of the witnesses in our protocol are entangled with one another, in a way that breaks the protocol's soundness, we believe Arthur should be able to use the remaining witness to detect this. This is a problem we leave to future work.

Second, our protocol made essential use of the PCP Theorem, in the strong version proved by Dinur [9]. One might wonder whether Theorem 1 could also be proved in a "black-box" fashion, without exploiting anything about the structure of 3SAT. The following simple theorem, proved in the full version, shows that the answer is no—and that indeed, in the black-box setting, there is essentially no savings at all over the classical witness size.

**Theorem 17.** *Let $f : \{0,1\}^n \rightarrow \{0,1\}$ be a black-box function. Then any $\mathsf{QMA}^f(k)$ protocol to convince Arthur that there exists an $x$ such that $f(x)=1$, with soundness gap $\Omega(1/\operatorname{poly}(n))$, must involve $n - O(\log n)$ qubits sent by the Merlins.*

Third, notice that our protocol does not let Arthur *find* a satisfying assignment for $\varphi$; it only convinces him that such an assignment exists. If there were a way to modify our protocol to let Arthur recover an assignment, this would have a spectacular consequence for quantum algorithms. Namely, by running Arthur's verification procedure with the $\widetilde{O}(\sqrt{m})$-qubit maximally mixed state in place of the witnesses, we could find a satisfying assignment for $\varphi$ with probability $2^{-\widetilde{O}(\sqrt{m})}$, with no help from any Merlins. But this would yield a $2^{\widetilde{O}(\sqrt{m})}$-time quantum algorithm for 3SAT—and in particular, a $2^{\widetilde{O}(\sqrt{n})}$-time algorithm in the "critical regime" $m = O(n)$!

Fourth, one of course wonders whether our $\widetilde{O}(\sqrt{m})$-qubit protocol is optimal. In Section 5, we will give evidence that *some* polynomial dependence on $m$ is necessary. In particular, it will follow from our results there that, assuming the Strong Amplification Conjecture, there are no unentangled witnesses of size $n^{o(1)}$ for any $\mathsf{NP}$-complete problem, which can be verified by an $n^{o(1)}$-time quantum algorithm, unless $\mathsf{NP} \subseteq \mathsf{DTIME}(2^{n^{o(1)}})$.

## 4 Additivity Implies Amplification

In this section we show how to amplify any $\mathsf{QMA}(k)$ protocol to exponentially small error, and to simulate $k$ provers with two, assuming the Additivity Conjecture.

### 4.1 Entanglement of Formation

The analysis of our amplification protocol will involve showing that Arthur cannot create "too much" entanglement during his verification procedure. To make this precise, we need some way to measure the entanglement of mixed states. Fortunately, this is one of the most studied topics in quantum information theory. One particular entanglement measure—the *entanglement of formation* $E_F$ defined by Bennett et al. [5]—will be particularly useful for us.

**Definition 18.** *Given a bipartite state $\rho^{AB}$, the entanglement of formation $E_F(\rho^{AB})$ is the minimum of $\sum_i p_i E(|\psi_i\rangle)$ over all decompositions $\rho^{AB} = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, where $E(|\psi_i\rangle)$ is the entanglement entropy* *of $|\psi_i\rangle$ (see Nielsen and Chuang [16] for a more detailed definition).*

Intuitively, $E_F$ measures the minimum number of entangled pairs $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ that are needed to prepare $\rho^{AB}$.

Almost by definition, $E_F$ satisfies *convexity*: for all $\rho^{AB}$ and $\sigma^{AB}$,

$$E_F\left(\alpha\rho^{AB} + \beta\sigma^{AB}\right) \leq \alpha E_F\left(\rho^{AB}\right) + \beta E_F\left(\sigma^{AB}\right).$$

It is also easy to see that $E_F\left(\rho^{AB}\right) = 0$ if and only if $\rho^{AB}$ is separable. In this paper, we will need two further properties of $E_F$. The first property is what we called "faithfulness" in Section 1.4.

**Lemma 19.** *Suppose $E_F(\rho^{AB}) \leq \varepsilon$. Then there exists a separable state that is $\sqrt{2\varepsilon}$-close to $\rho^{AB}$ in trace distance.*

The second property is that $E_F$ cannot increase by much by acting on few qubits.

**Lemma 20.** *Suppose $\sigma^{AB}$ is obtained from $\rho^{AB}$ by acting on at most $n$ qubits from each register. Then $E_F\left(\sigma^{AB}\right) \leq E_F\left(\rho^{AB}\right) + 2n$.*

*Proof.* Let $\tau^{AB}$ be $\rho^{AB}$ tensored with $2n$ EPR pairs. Then clearly $E_F\left(\tau^{AB}\right) \leq E_F\left(\rho^{AB}\right) + 2n$. Furthermore, it is not hard to see that $\sigma^{AB}$ can be obtained from $\tau^{AB}$ using local operations and classical communication, as follows. First teleport $n$ qubits from the $A$ register to the $B$ register (using $n$ EPR pairs), then apply the requisite superoperator, then teleport $n$ qubits from the $B$ register back to the $A$ register (using another $n$ EPR pairs). Hence $E_F\left(\sigma^{AB}\right) \leq E_F\left(\tau^{AB}\right)$, and the lemma follows. □

Given an entanglement measure $E$, we call $E$ *superadditive* if for any state $\rho^{AA',BB'}$ on four registers,

$$E(\rho^{AA',BB'}) \geq E\left(\rho^{AB}\right) + E(\rho^{A'B'}).$$

As mentioned earlier, the analysis of our $\mathsf{QMA}(k)$ amplification protocol will rely on the following central conjecture from quantum information theory:

**Conjecture 21** (Additivity Conjecture)**.** *$E_F$ is superadditive.*

Shor [18] showed that Conjecture 21 is equivalent to several other additivity conjectures in quantum information theory, including the additivity of the Holevo capacity for quantum channels.

## 4.2 The Two-Prover Case

We now show that the Additivity Conjecture implies the $\mathsf{QMA}(2)$ Amplification Conjecture.

**Theorem 22.** *Assume the Additivity Conjecture. Then $\mathsf{QMA}(2,a,b) = \mathsf{QMA}\left(2, 2^{-p(n)}, 1-2^{-p(n)}\right)$ for all $b - a = \Omega(1/\operatorname{poly}(n))$ and all polynomials $p$.*

*Proof.* Let $L$ be a language in $\mathsf{QMA}(2,a,b)$; then we need to show $L \in \mathsf{QMA}\left(2, 2^{-p(n)}, 1-2^{-p(n)}\right)$. Let $Q$ be Arthur's verification algorithm in the original $\mathsf{QMA}(2,a,b)$ protocol, and let the original Merlins' messages have $r(n)$ qubits each for some polynomial $r$. Also, let $T(n)$ be a number of repetitions of $Q$ that suffices to amplify it to error probability $2^{-p(n)}$, assuming no entanglement among Merlin$_A$'s or Merlin$_B$'s registers. By a standard Chernoff bound, we can take $T(n) := C \cdot p(n)/(b-a)^2$ for some constant $C$.

Our amplified protocol is the following.

(1) Arthur asks Merlin$_A$ and Merlin$_B$ to supply $q(n)$ copies each of their respective witnesses, where $q(n) := 128T(n)r(n)/(b-a)^2$. Denote by $\rho^{A_1A_2\cdots A_{q(n)}}$ and $\rho^{B_1B_2\cdots B_{q(n)}}$ the $q(n)r(n)$-qubit states that Arthur actually receives.

(2) For all $t := 1$ to $T(n)$, Arthur chooses registers $A_j$ and $B_k$ uniformly and independently from among those not already chosen, and runs $Q$ on the state $\rho^{A_jB_k}$.

(3) Arthur accepts if at least $\frac{a+b}{2}T(n)$ of the $T(n)$ invocations of $Q$ accepted, and rejects otherwise.

We need to show two things about this protocol, completeness and soundness.

**Completeness:** If the Merlins are honest, they can simply send $|\psi_A\rangle^{\otimes q(n)}$ and $|\psi_B\rangle^{\otimes q(n)}$ respectively, where $|\psi_A\rangle \otimes |\psi_B\rangle$ is a witness that $Q$ accepts with probability at least $b$. Then by assumption, Arthur will accept with probability at least $1 - 2^{-p(n)}$.

**Soundness:** As usual, this is the interesting part. Our central claim is the following:

*At every one of the $T(n)$ iterations, Arthur can be considered to be running $Q$ on a bipartite state $\rho^{AB}$ that is $\varepsilon$-close to a separable state, where $\varepsilon := \sqrt{8T(n)r(n)/q(n)}$.*

Let us first see why soundness follows from the above claim. Suppose $x \notin L$. Then $Q$ accepts every separable state with probability at most $a$. By Proposition 7, then, $Q$ also accepts every state that is $\varepsilon$-close to separable with probability at most $a + \varepsilon$. But

$$\varepsilon = \sqrt{\frac{8T(n)r(n)}{q(n)}} \leq \frac{b-a}{4}.$$

So every invocation of $Q$ accepts with probability at most $a + \frac{b-a}{4}$. Therefore, provided we choose a sufficiently large constant $C$ when defining $T(n)$, Arthur will accept with probability at most $2^{-p(n)}$ by a Chernoff bound.

We now prove the claim. By Lemma 20, the entanglement of formation between Merlin$_A$'s registers and Merlin$_B$'s registers can be at most $2r(n)$ after the first iteration, at most $4r(n)$ after the second iteration, and so on. Hence

$$E_F\left(\rho^{A_1A_2\cdots A_{q(n)},B_1B_2\cdots B_{q(n)}}\right) \leq 2T(n)\,r(n)$$

throughout. Also, let $S_A$ and $S_B$ be the sets of $A$-registers and $B$-registers respectively that Arthur has not yet chosen. Then $|S_A| = |S_B| = q(n) - T(n)$. Assuming the Additivity Conjecture, we therefore have

$$\begin{aligned}&\sum_{A_j\in S_A, B_k\in S_B} E_F\left(\rho^{A_jB_k}\right)\\ &\leq (q(n) - T(n))\,E_F\left(\rho^{A_1A_2\cdots A_{q(n)},B_1B_2\cdots B_{q(n)}}\right)\\ &\leq 2T(n)\,r(n)\,(q(n) - T(n)).\end{aligned}$$

So if we define

$$\sigma := \frac{1}{|S_A|\,|S_B|}\sum_{A_j\in S_A, B_k\in S_B}\rho^{A_jB_k},$$

then the convexity of $E_F$ implies that

$$\begin{aligned}E_F(\sigma) &\leq \frac{1}{|S_A|\,|S_B|}\sum_{A_j\in S_A, B_k\in S_B} E_F\left(\rho^{A_jB_k}\right)\\ &\leq \frac{2T(n)\,r(n)}{q(n) - T(n)}\\ &\leq \frac{4T(n)\,r(n)}{q(n)},\end{aligned}$$

using the fact that $T(n) \leq q(n)/2$. By Lemma 19, this means that $\sigma$ is $\sqrt{8T(n)\,r(n)/q(n)}$-close to a separable state, as claimed. □

### 4.3 The $k$-Prover Case

Recall that Kobayashi et al. [12] showed that amplification of $\mathsf{QMA}(k)$ protocols implies $\mathsf{QMA}(k) = \mathsf{QMA}(2)$ for all $k \geq 2$. Now that we have shown that "additivity implies amplification," one might think it would follow that additivity implies collapse of $\mathsf{QMA}(k)$ to $\mathsf{QMA}(2)$. Unfortunately, the result of Kobayashi et al. requires amplification for all $\mathsf{QMA}(k)$, while we have only shown that additivity implies amplification for $\mathsf{QMA}(2)$. In this section we solve the problem by strengthening Kobayashi et al.'s result. In particular, we will show that *any* $\mathsf{QMA}(k)$ *protocol with constant soundness can be simulated by a* $\mathsf{QMA}(2)$ *protocol with soundness* $\Omega(1/k)$. Combined with Theorem 22, this will then imply that $\mathsf{QMA}(k) = \mathsf{QMA}(2)$ for all $k \geq 2$ assuming the Additivity Conjecture.

**Theorem 23.** $\mathsf{QMA}(k, a, b) \subseteq \mathsf{QMA}\left(2, 1 - \frac{(b-a)^2}{8k}, 1 - 2^{-n}\right)$.

*Proof.* We will show that for all $k$ and all $\delta = \Omega(1/\operatorname{poly}(n))$,

$$\mathsf{QMA}\left(k, 1-\delta, 1-2^{-n}\right) \subseteq \mathsf{QMA}\left(2, 1 - \frac{\delta^2}{8k}, 1 - 2^{-n}\right).$$

This will suffice to prove the theorem, since Lemma 6 implies that for all $k$ and all $a, b$, we have $\mathsf{QMA}(k, a, b) \subseteq \mathsf{QMA}(k, 1 - (b-a), 1 - 2^{-n})$.

Our protocol is as follows. Merlin$_A$ and Merlin$_B$ send $k$-partite states $\rho^{A_1A_2\cdots A_k}$ and $\rho^{B_1B_2\cdots B_k}$ respectively. Given these states, Arthur performs one of the following two tests, each with probability $1/2$:

(1) Choose $i \in [k]$ uniformly at random, perform a swap test between $\rho^{A_i}$ and $\rho^{B_i}$, and accept if and only if the swap test accepts.

(2) Simulate the $\mathsf{QMA}(k, 1-\delta, 1-2^{-n})$ protocol, using $\rho^{A_1A_2\cdots A_k}$ in place of the $k$ witness registers.

We first show completeness of the above protocol. If the Merlins are honest, they can both simply send $k$ unentangled accepting witnesses for the $\mathsf{QMA}(k)$ protocol being simulated. In that case step (1) accepts with probability 1, while step (2) accepts with probability at least $1 - 2^{-n}$.

We now show soundness. Suppose any set of unentangled witnesses causes the $\mathsf{QMA}(k)$ protocol to reject with probability at least $\delta$. Then we need to show that any pair of witnesses $\rho^{A_1A_2\cdots A_k}$ and $\rho^{B_1B_2\cdots B_k}$ causes the $\mathsf{QMA}(2)$ protocol to reject with probability at least $\frac{\delta^2}{8k}$. We consider two cases.

First suppose $\rho^{A_1A_2\cdots A_k}$ is $\varepsilon$-close in trace distance to some separable pure state $|\Psi\rangle$. Then by Proposition 7, step (2) rejects with probability at least $\delta - \varepsilon$.

Next suppose $\rho^{A_1A_2\cdots A_k}$ is $\varepsilon$-far in trace distance from any separable pure state. Then by Proposition 9, we have $\left\langle\Psi|\rho^{A_1A_2\cdots A_k}|\Psi\right\rangle < 1 - \varepsilon^2$ for all separable pure states $|\Psi\rangle$. So taking the contrapositive of Proposition 8, for all pure states $|\psi_1\rangle, \ldots, |\psi_k\rangle$ we have

$$\sum_{i=1}^{k}\left(1 - \left\langle\psi_i|\rho^{A_i}|\psi_i\right\rangle\right) > \varepsilon^2.$$

Hence step (1) rejects with probability greater than $\frac{\varepsilon^2}{2k}$ by Proposition 10.

Setting $\varepsilon = \delta/2$, we thus find that the protocol rejects with probability at least $\frac{\delta^2}{8k}$. □

Combining Theorem 23 with Theorem 22 now yields the following:

**Corollary 24.** *The Additivity Conjecture implies the Collapse Conjecture, that* $\mathsf{QMA}(k) = \mathsf{QMA}(2)$ *for all* $k \geq 2$.

### 4.4 Limited Entanglement

Let us mention another interesting result that can be obtained by the same techniques as in Theorem 22. Define the complexity class $\mathsf{QMA}(2; h)$ to be the same as $\mathsf{QMA}(2)$, except that now, instead of being completely unentangled, the two Merlins are allowed to share $h$ EPR pairs $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Assuming the Additivity Conjecture, we show that limited entanglement gives the Merlins no more power to cheat than no entanglement at all:

**Theorem 25.** *The Additivity Conjecture implies* $\mathsf{QMA}(2) \subseteq \mathsf{QMA}(2; h(n))$ *for every fixed polynomial* $h$.

*Proof Sketch.* To simulate a $\mathsf{QMA}(2)$ protocol in $\mathsf{QMA}(2; h(n))$, we use the amplified protocol exactly as in Theorem 22, except that instead of asking the Merlins for $O(T(n) r(n))$ witnesses each, Arthur asks them for $O(T(n) r(n) + h(n))$ witnesses. The only observation we need to make is that the proof of Theorem 22 still goes through if, in addition to the entanglement that Arthur creates in the course of his verification, there is *also* some fixed amount of entanglement to start. □

It is an interesting question whether the converse holds: that is, whether $\mathsf{QMA}(2; h(n)) \subseteq \mathsf{QMA}(2)$.

### 4.5 Symmetric $\mathsf{QMA}(k)$

Define the complexity class $\mathsf{SymQMA}(k, a, b)$ the same way as $\mathsf{QMA}(k, a, b)$, except that now we are promised that the $k$ witnesses are all identical (in both the completeness and soundness cases). We saw in Section 3.3 that symmetric $\mathsf{QMA}(k)$ protocols are sometimes easier to analyze than non-symmetric ones. However, in the full version we show that assuming the Additivity Conjecture, $\mathsf{QMA}(k)$ and $\mathsf{SymQMA}(k)$ are actually equivalent.

The first step is to show they are (unconditionally) equivalent up to a loss in error bounds.

**Lemma 26.** $\mathsf{QMA}(k, a, b) \subseteq \mathsf{SymQMA}(k, a, b) \subseteq \mathsf{QMA}\left(k, 1 - \frac{(b-a)^2}{8k}, 1 - 2^{-n}\right)$.

Combining Lemma 26 with Theorem 23, we immediately get the following.

**Theorem 27.** *The Additivity Conjecture implies* $\mathsf{SymQMA}(k) = \mathsf{QMA}(k) = \mathsf{QMA}(2)$ *for all* $k \geq 2$.

## 5 Evidence That $\mathsf{QMA}(k) \subseteq \mathsf{PSPACE}$

It is obvious that $\mathsf{QMA}(k) \subseteq \mathsf{NEXP}$: simply guess exponentially-long classical descriptions of the $k$ quantum proofs. Yet this trivial upper bound is still the best we know. In this section, we will show the nontrivial upper bound $\mathsf{QMA}(k) \subseteq \mathsf{PSPACE}$, assuming the following conjecture.

**Conjecture 28** (Strong Amplification)**.** *Every language in* $\mathsf{QMA}(2)$ *admits a protocol with completeness* $1 - 2^{-n}$ *and soundness* $2^{-2s(n)}$*, where* $s(n)$ *is the number of qubits sent by Merlin*$_B$.

Let us say a few words about why Conjecture 28 might be true. In studying probabilistic complexity classes, one typically assumes amplification theorems will hold unless there is some clear obstruction to them. In the case of $\mathsf{QMA}(2)$ amplification where both of the witnesses remain small, there really is such an obstruction: namely, it will follow from results in this section that such in-place amplification would imply $\mathsf{NP} \subseteq \mathsf{DTIME}(n^{\text{polylog}\, n})$. On the other hand, we know of no similar obstruction in the case where one witness remains small, but the other could grow by a polynomial factor depending on the desired error bound.

We now turn to proving that Conjecture 28 implies $\mathsf{QMA}(k) \subseteq \mathsf{PSPACE}$ for all $k$. We know from Kobayashi et al. [12] that even the ordinary amplification conjecture implies $\mathsf{QMA}(k) = \mathsf{QMA}(2)$ for all $k \geq 2$. Therefore, our task reduces to showing that Conjecture 28 implies $\mathsf{QMA}(2) \subseteq \mathsf{PSPACE}$.

We will need the following lemma of Aaronson [1].

**Lemma 29** (**[1]**)**.** *Let $M$ be a 2-outcome POVM on a bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. Also, let $\{|1\rangle, \ldots, |d\rangle\}$ be any orthonormal basis for $\mathcal{H}_B$, and for all $j \in \{1, \ldots, d\}$ let $M_j$ be the POVM on $\mathcal{H}_A$ induced by applying $M$ to $\mathcal{H}_A \otimes |j\rangle$. Suppose that there exists a product state $\rho \otimes \sigma$ in $\mathcal{H}_A \otimes \mathcal{H}_B$ such that $M$ yields outcome 1 with probability at least $p > 0$ when applied to $\rho \otimes \sigma$. Then, if we apply $M_{j_1}, \ldots M_{j_T}$ in sequence to $\rho$, where $j_1, \ldots j_T$ are drawn uniformly and independently from $\{1, \ldots, d\}$, and $T \geq d/p^2$, the probability that at least one of these measurements yields outcome 1 is at least $\left(p - \sqrt{d/T}\right)^2$.*

Let $\mathsf{QMA}_{\mathsf{PSPACE}}$ be the same as $\mathsf{QMA}$, except that Arthur can run in quantum polynomial space.

**Lemma 30.** *Conjecture 28 implies* $\mathsf{QMA}(2) \subseteq \mathsf{QMA}_{\mathsf{PSPACE}}$.

*Proof.* Let $L$ be a language in $\mathsf{QMA}(2)$. By Conjecture 28, there is a protocol for $L$ in which the completeness and soundness bounds are $1 - 2^{-n}$ and $2^{-ns(n)}$, respectively, and Merlin$_B$'s message is over $s(n)$ qubits. Let $M$ be the two-outcome POVM induced by Arthur's verification procedure. As in Lemma 29, Arthur can receive just the message of Merlin$_A$, guess a classical basis state in place of Merlin$_B$'s message, apply $M$, repeat this process $T$ times, and finally take the OR of the outcomes as his answer.

More precisely, we set $d := 2^{s(n)}$ and $T := 2^{2s(n)-2}$. Then if $x \in L$, Arthur accepts with probability at least $(1 - 2^{-n} - \sqrt{d/T})^2 > 2/3$ by Lemma 29. If $x \notin L$, on the other hand, then in each step Arthur's probability of acceptance is at most $2^{-2s(n)}$. So by the union bound, his total probability of acceptance after taking the OR is at most $T2^{-2s(n)} < 1/3$. $\square$

**Lemma 31.** $\mathsf{QMA}_{\mathsf{PSPACE}} = \mathsf{PSPACE}$.

*Proof Sketch.* Let $L$ be a language in $\mathsf{QMA}_{\mathsf{PSPACE}}$. Then $L$ has a protocol in which Arthur receives a witness with $p(n)$ qubits (for some polynomial $p$), and then decides whether to accept or reject it in quantum polynomial space. Hence there exists a positive Hermitian matrix $A$, of size $2^{p(n)} \times 2^{p(n)}$, such that if $x \in L$ then the largest eigenvalue of $A$ is at least $2/3$, while if $x \notin L$ then the largest eigenvalue is at most $1/3$. Furthermore, $A$ is equal to the product of exponentially many efficiently-computable matrices. So computing $\mathrm{Tr}(A^{2p(n)})$ is just an exponential-size linear algebra problem, which can be solved in $\mathsf{PSPACE}$. On the other hand $\mathrm{Tr}(A^{2p(n)})$ depends on the largest eigenvalue of $A$, and is greater than $(2/3)^{2p(n)}$ if $x \in L$, and less than $2^{p(n)}/3^{2p(n)}$ if $x \notin L$. Hence we can decide $L$ in $\mathsf{PSPACE}$, and $\mathsf{QMA}_{\mathsf{PSPACE}} \subseteq \mathsf{PSPACE}$. Since $\mathsf{PSPACE} \subseteq \mathsf{QMA}_{\mathsf{PSPACE}}$ is obvious we are done. $\square$

Combining Lemma 30 with Lemma 31 now yields the main result.

**Theorem 32.** *Conjecture 28 implies* $\mathsf{QMA}(2) \subseteq \mathsf{PSPACE}$.

Or if we "scale down by an exponential," Conjecture 28 implies that

$$\mathsf{QMA}_{\log}(2) \subseteq \mathsf{DSPACE}(\mathrm{polylog}\, n) \subseteq \mathsf{DTIME}(n^{\mathrm{polylog}\, n})$$

where $\mathsf{QMA}_{\log}(2)$ is the same as $\mathsf{QMA}(2)$ except that the witnesses have size $O(\log n)$ and are verified in time $\mathrm{polylog}\, n$. Assuming Conjecture 28, this means in particular that the 3-COLORING protocol of Blier and Tapp [6] cannot be amplified to constant soundness, unless $\mathsf{NP} \subseteq \mathsf{DTIME}(n^{\mathrm{polylog}\, n})$.

Theorem 32 can also be seen as giving a *quasipolynomial-time approximation algorithm for an* $\mathsf{NP}$*-hard optimization problem*: namely, the problem of finding the separable state $|\psi_A\rangle |\psi_B\rangle$ that maximizes the expectation value of a given observable.[9] (Of course, such an algorithm would require a strong amplification procedure as a subroutine.) We now state the connection more precisely.

[9] We know that this problem is $\mathsf{NP}$-hard (and indeed, hard to approximate to within a $\Omega(1/N^6)$ additive term) by the result of Blier and Tapp [6].

**Theorem 33.** *Let $M$ be a measurement on a bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, and let $p(M)$ be the maximum, over all separable states $|\psi_A\rangle |\psi_B\rangle$, of the probability that $M$ accepts $|\psi_A\rangle |\psi_B\rangle$. Also, let $N = (\dim \mathcal{H}_A)(\dim \mathcal{H}_B)$ and $\varepsilon > 0$. Then assuming Conjecture 28, there exists a deterministic algorithm that takes $M$ as input, approximates $p(M)$ to within additive error $\varepsilon$, and runs in time $N^{\mathrm{polylog}\, N/\mathrm{poly}(\varepsilon)}$.*

# 6 Nonexistence of Perfect Disentanglers

**Definition 34.** *Let $\mathcal{H}$ and $\mathcal{K}$ be two finite-dimensional Hilbert spaces. Then given a superoperator $\Phi : \mathcal{H} \to \mathcal{K} \otimes \mathcal{K}$, we say $\Phi$ is an $(\varepsilon, \delta)$-disentangler if*

*(i) $\Phi(\rho)$ is $\varepsilon$-close to a separable state for every $\rho$, and*

*(ii) for every separable state $\sigma$, there exists a $\rho$ such that $\Phi(\rho)$ is $\delta$-close to $\sigma$.*

As pointed out in Section 1.6, if for sufficiently small constants $\varepsilon, \delta$ there exists an $(\varepsilon, \delta)$-disentangler with $\log \dim \mathcal{H} = O(\mathrm{poly}(\log \dim \mathcal{K}))$—and if, moreover, that disentangler can be implemented in quantum polynomial time—then $\mathsf{QMA}(2) = \mathsf{QMA}$.

Watrous (personal communication) has proposed the following fundamental conjecture.

**Conjecture 35** (Watrous). *For all constants $\varepsilon, \delta < 1$, any $(\varepsilon, \delta)$-disentangler requires $\dim \mathcal{H} = 2^{\Omega(\dim \mathcal{K})}$.*

A proof of Conjecture 35 would be an important piece of formal evidence that $\mathsf{QMA}(2) \neq \mathsf{QMA}$, and might even lead to a "quantum oracle separation" (as defined by Aaronson and Kuperberg [2]) between the two classes.

In the full version we show that, at least in the case $\varepsilon = \delta = 0$, no disentangler exists in *any* finite dimension. This result would be false if we let either $\varepsilon$ or $\delta$ be nonzero.

**Theorem 36.** *Let $\Phi : \mathcal{H} \to \mathcal{K} \otimes \mathcal{K}$ be any superoperator whose image is the set of separable states. Then $\dim \mathcal{K} \geq 2$ implies $\dim \mathcal{H} = \infty$.*

# 7 Open Problems

## 7.1 The Power of Multiple Merlins

The power of $\mathsf{QMA}(2)$ and related classes is still poorly understood. Can we find a "classical" problem (for example, a group-theoretic problem like those of Watrous [20]) that is in $\mathsf{QMA}(2)$ but not obviously in $\mathsf{QMA}$? Can we find a natural $\mathsf{QMA}(k)$-complete promise problem?

Regarding our 3SAT protocol, can we reduce the number of provers to two? Can we reduce the number of qubits below $\widetilde{O}(\sqrt{n})$, or alternatively, give evidence against this possibility? For example, can we show that $\Omega(\sqrt{n})$ witnesses are information-theoretically required for the Uniformity Test? Finally, can we show unconditionally that $\mathsf{QMA}(2) \subseteq \mathsf{EXP}$?

A long-shot possibility would be to give a quantum algorithm to *find* the unentangled witnesses in the 3SAT protocol, in as much time as it would take were the witnesses entangled. This would yield a $2^{\widetilde{O}(\sqrt{n})}$-time quantum algorithm for 3SAT.

### 7.2 Amplification and Other Complexity Issues

In defining $\mathsf{QMA}(k)$, does it matter if the amplitudes are reals or complex numbers? For BQP and QMA, it is not hard to show that this distinction is irrelevant. Interestingly, though, the usual equivalence proofs break down for $\mathsf{QMA}(k)$.

Can we show directly (i.e., without proving the full Additivity Conjecture) that $\mathsf{QMA}(k) = \mathsf{QMA}(2)$, or that $\mathsf{QMA}(2)$ protocols can be amplified?

Can we prove Conjecture 35: that there is no $(\varepsilon, \delta)$-disentangler with $\mathrm{poly}(n)$ qubits and $\varepsilon, \delta > 0$? Can we at least rule out such a disentangler when either $\varepsilon > 0$ *or* $\delta > 0$? Related to that, can we give a quantum oracle $U$ (as defined by Aaronson and Kuperberg [2]) such that $\mathsf{QMA}^U \neq \mathsf{QMA}^U(2)$? Can we at least show that Conjecture 35 would imply the existence of such an oracle?

### 7.3 $\mathsf{QMA}(k)$ With Unentangled Measurements

Recall that our 3SAT protocol involved three tests: Satisfiability, Symmetry, and Uniformity. Suppose we are willing to settle for completeness $1-\varepsilon$ rather than 1, and suppose we modify the Uniformity Test so that Arthur rejects on not seeing enough collisions. Then can the Symmetry Test be omitted? If so, then the resulting protocol would have the extremely interesting property of making no entangled measurements, yet nevertheless depending crucially on the absence of entanglement among the witnesses.

More generally, define $\mathsf{BellQMA}(k)$ to be the subclass of $\mathsf{QMA}(k)$ in which Arthur is restricted to making a separate measurement on each witness $|\varphi_i\rangle$, with no entanglement between the measurements. (The name arises because Arthur is essentially restricted to performing a "Bell experiment.") What is the power of this class? Does $\mathsf{BellQMA}(k) = \mathsf{QMA}(k)$? Does $\mathsf{BellQMA}(k) = \mathsf{BellQMA}(2)$ for all $k \geq 2$? Note that it is trivial to show amplification for $\mathsf{BellQMA}(k)$. This is because, without entangling measurements, the entanglement-swapping problem described in Section 1.4 can never arise.

## Acknowledgments

We thank Norbert Schuch for pointing out that Theorem 22 can be based on the Additivity Conjecture rather than a conjecture about squashed entanglement; Madhu Sudan for pointers on the PCP Theorem; John Watrous for posing Conjecture 35; an anonymous reviewer for greatly simplifying the proof of Lemma 20 and other help; and Patrick Hayden, Ryan O'Donnell, Alain Tapp, and Andreas Winter for helpful discussions.

## References

[1] S. Aaronson. QMA/qpoly is contained in PSPACE/poly: de-Merlinizing quantum protocols. In *Proc. IEEE Conference on Computational Complexity*, pages 261–273, 2006. quant-ph/0510230.

[2] S. Aaronson and G. Kuperberg. Quantum versus classical proofs and advice. *Theory of Computing*, 3(7):129–157, 2007. Previous version in Proceedings of CCC 2007. quant-ph/0604056.

[3] D. Aharonov and T. Naveh. Quantum NP - a survey. quant-ph/0210077, 2002.

[4] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. Wootters. Teleporting an unknown quantum state by dual classical and EPR channels. *Phys. Rev. Lett.*, 70:1895–1898, 1993.

[5] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, 54:3824–3851, 1996. quant-ph/9604024.

[6] H. Blier and A. Tapp. All languages in NP have very short quantum proofs. arXiv:0709.0738, 2007.

[7] D. M. Bloom and W. Knight. A birthday problem. *American Mathematical Monthly*, 80(10):1141–1142, December 1973.

[8] M. Christandl and A. Winter. "Squashed entanglement" - an additive entanglement measure. *J. Math. Phys.*, 45(3):829–840, 2004. quant-ph/0308088.

[9] I. Dinur. The PCP theorem by gap amplification. *J. ACM*, 54(3):12, 2007.

[10] S. Khanna, M. Sudan, L. Trevisan, and D. P. Williamson. The approximability of constraint satisfaction problems. *SIAM J. Comput.*, 30(6):1863–1920, 2000.

[11] A. Kitaev, A. Shen, and M. N. Vyalyi. *Classical and Quantum Computation*. American Mathematical Society, 2002.

[12] H. Kobayashi, K. Matsumoto, and T. Yamakami. Quantum Merlin-Arthur proof systems: are multiple Merlins more helpful to Arthur? In *ISAAC*, pages 189–198, 2003. quant-ph/0306051.

[13] R. König and R. Renner. A de Finetti representation for finite symmetric quantum states. *J. Math. Phys.*, 46(122108), 2005. quant-ph/0410229.

[14] Y.-K. Liu, M. Christandl, and F. Verstraete. N-representability is QMA-complete. *Phys. Rev. Lett.*, 98(110503), 2007. quant-ph/0609125.

[15] C. Marriott and J. Watrous. Quantum Arthur-Merlin games. *Computational Complexity*, 14(2):122–152, 2005.

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

[17] C. H. Papadimitriou and M. H. Yannakakis. Optimization, approximation, and complexity classes. *J. Comput. Sys. Sci.*, 43(3):425–440, 1991.

[18] P. W. Shor. Equivalence of additivity questions in quantum information theory. *Communications in Mathematical Physics*, 246(3):453–472, 2004. quant-ph/0305035.

[19] J. Watrous. Space-bounded quantum complexity. *J. Comput. Sys. Sci.*, 59(2):281–326, 1999.

[20] J. Watrous. Succinct quantum proofs for properties of finite groups. In *Proc. IEEE FOCS*, pages 537–546, 2000. cs.CC/0009002.

[21] M. Zukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert. Event-ready-detectors: Bell experiment via entanglement swapping. *Phys. Rev. Lett.*, 71:4287–4290, 1993.