Electronic Colloquium on Computational Complexity, Report No. 51 (2008)

# The Power of Unentanglement

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 [8].

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?

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

ISSN 1433-8092

### 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 [13] and Watrous [26] 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 [18], is contained in $\mathsf{PP}$ [18], and has natural complete promise problems [13]. (See Aharonov and Naveh [3] for a survey.)

In 2003, Kobayashi, Matsumoto, and Yamakami [15] 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 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 $k \geq 2$?[1]

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. [15] proved that a positive answer to question (3) implies a positive answer to question (4). 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 [17] 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 [8] 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.[2]

### 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 [8], but the point is that Arthur can detect cheating with *constant*

[1] The third and fourth questions are motivated, in part, by an analogy to classical *multi-prover interactive proof systems*—where the Parallel Repetition Theorem of Raz [22] and the $\mathsf{MIP}(k) = \mathsf{MIP}(2)$ theorem of Ben-Or et al. [4] turned out to be crucial for understanding the class $\mathsf{MIP}$.

[2] 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 $\mathsf{QMA}(2) = \mathsf{NEXP}$!

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 [11].

**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 $\mathsf{QMA}(2)$ 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 proving 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.[3]

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

In particular, if $m = O(n)$,[5] 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).

[3] For Arthur could first use the in-place amplification of Marriott and Watrous [18] 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.

[4] 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.

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

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 [11], which increase the number of variables and clauses by at most an $O(\operatorname{polylog} n)$ 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

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

where $x_1, \ldots, x_n \in \{0,1\}^n$ is 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 $|\psi\rangle$ in a basis corresponding to the clauses of $\phi$. With constant probability, he will get an outcome of the form

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

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 $|\psi\rangle$ 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 $|\psi\rangle$ 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 $|\psi\rangle$, then he can check with constant soundness whether $|\psi\rangle$ is proper or far from any proper state. Indeed, even if Arthur is given $K = O(\sqrt{n})$ states $|\varphi_1\rangle, \ldots, |\varphi_K\rangle$ 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 $|\psi\rangle$. This then yields a protocol for 3SAT with constant soundness and $O(\sqrt{n})$ unentangled proofs of size $O(\log n)$—for Arthur can just choose randomly whether to perform the satisfiability test described above, or to check whether most of the $|\varphi_k\rangle$'s are close to some proper state $|\psi\rangle$.

To check that most of the states are at least close to *each other*, Arthur simply has to perform a "swap test" between (say) $|\varphi_1\rangle$ and a random other state $|\varphi_k\rangle$. So the problem is reduced to the following: assuming most of the $|\varphi_k\rangle$'s are close to $|\varphi_1\rangle$, how can Arthur decide whether $|\varphi_1\rangle$ 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 $|\varphi_k\rangle$ in an orthonormal basis that contains the vectors $|i\rangle + |j\rangle$ and $|i\rangle - |j\rangle$ 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 $\Omega(1)$ probability, Arthur will find a *collision*: that is, two outcomes of the form $|i\rangle \pm |j\rangle$, for the same edge $(i,j) \in \mathcal{M}$. So suppose this happens. Then if the $|\varphi_k\rangle$'s are all equal to a proper state $|\psi\rangle = \sum_{i=1}^{n} (-1)^{x_i} |i\rangle$, the two outcomes will clearly "agree": that is, they will either both be $|i\rangle + |j\rangle$ (if $x_i = x_j$) or both be $|i\rangle - |j\rangle$ (if $x_i \neq x_j$). On the other hand, suppose the $|\varphi_k\rangle$'s are far from any proper state. In that case, we show that the outcomes will "disagree" (that is, one will be $|i\rangle + |j\rangle$ and the other will be $|i\rangle - |j\rangle$) with $\Omega(1)$ probability.

To understand why, consider that there are two ways for a state $|\varphi\rangle = \sum_{i=1}^{n} \alpha_i |i\rangle$ to be far from proper. First, the probability distribution $\left(|\alpha_1|^2, \ldots, |\alpha_n|^2\right)$, which corresponds to measuring $|\varphi\rangle$ 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 $|\varphi\rangle$ 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 $\alpha_i |i\rangle + \alpha_j |j\rangle$ that is not close to either $|i\rangle + |j\rangle$ or $|i\rangle - |j\rangle$.

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 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 $\mathsf{QMA}$ protocol with constant error to a protocol with exponentially small error. Merlin simply sends Arthur $m = \text{poly}(n)$ 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* [27], 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.[6]

[6] Indeed, this example can be seen as a special case of *quantum teleportation* [6]: 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

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 [16], 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 $\text{poly}(n)$ 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. [7]—that is known to satisfy (ii), and is conjectured to satisfy (i).[7] This conjecture is known to be equivalent to the Additivity Conjecture from quantum information theory [23].

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.

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

It is well-known that $\mathsf{QMA} \subseteq \mathsf{PP}$ [18]. 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.

state into the two right registers by acting only on the two left registers.

[7] There is also another measure—the *squashed entanglement* $E_{sq}$, introduced by Christandl and Winter [10]—that is known to satisfy (i), but unfortunately can be shown *not* to satisfy (ii).

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 [25] that $\mathsf{QMA}_{\mathsf{PSPACE}} = \mathsf{BQPSPACE} = \mathsf{PSPACE}$.

Assuming strong amplification, we also get that the "maximum separable eigenvalue"[8] of an $n \times n$ Hermitian operator can be approximated to within an additive constant in $\mathsf{DTIME}(n^{\operatorname{polylog} n})$.

### 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 [17]—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: simply take as input a classical description of the separable state $\sigma$ to be prepared, measure that description in the computational basis, and then prepare $\sigma$.[9]

Likewise, if we are willing to settle for every output being close to a separable state, then a disentangler also exists. For some sufficiently large $N$, take as input a quantum state on registers $R_0, R_1, \ldots, R_N$, choose an index $i \in [N]$ uniformly at random, and output the joint state of $R_0$ and $R_i$ (discarding everything else). It follows, from the finite quantum de Finetti theorem [16], that with high probability this state will be close to separable.

The key problem with both of these approaches is that *the input Hilbert space needs to be exponentially larger than the output Hilbert space.* In the case of the "de Finetti approach," this follows from considering the maximally antisymmetric state

$$\frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} (-1)^{\operatorname{sgn}(\sigma)} \left|\sigma(1)\right\rangle \cdots \left|\sigma(N)\right\rangle,$$

which has the properties that

(i) there are exponentially many registers (as a function of $n = \log N$, the size of a given register), but

(ii) the reduced state of any two registers is far from a separable state.

[8] That is, the probability that a given measurement $M$ on an $n$-dimensional bipartite Hilbert space accepts, maximized over all separable states in the Hilbert space (rather than *all* states, which would correspond to taking the maximum eigenvalue of $M$).

[9] 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 $\sigma$.

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 $\mathsf{QMA}(2)$.

### 1.7 Table of Contents

The rest of the paper is organized as follows.

## 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, superoperators, 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* 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* $\mathsf{QMA} := \mathsf{QMA}(1)$*.*[10]

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

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 [18] 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). *$\mathsf{QMA}(k, a, b) = \mathsf{QMA}\left(k, 2^{-p(n)}, 1 - 2^{-p(n)}\right)$ for all $k$, all $a < b$ with $b - a = \Omega(1/\operatorname{poly}(n))$, and all polynomials $p$.*

[10] 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.

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

The main progress so far on these conjectures has been due to Kobayashi et al. [15], who showed that the Amplification and Collapse Conjectures are actually equivalent:

**Theorem 5** ([**15**])**.** *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 $b < 1$). □

## 2.2 Quantum Information

We now review some quantum information concepts that we will need. For more details see Nielsen and Chuang [19] for example.

Given two mixed states $\rho$ and $\sigma$, their *trace distance* is $\|\rho - \sigma\|_{\mathrm{tr}} := \frac{1}{2}\sum_{i=1}^{n} |\lambda_i|$, where $(\lambda_1, \ldots, \lambda_n)$ are the eigenvalues of $\rho - \sigma$. We will sometimes say $\sigma$ is $\varepsilon$*-close* to $\rho$ if $\|\rho - \sigma\|_{\mathrm{tr}} \leq \varepsilon$, 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* $p + \varepsilon$*.*

Trace distance also satisfies the triangle inequality:

**Proposition 8.** $\|\rho - \sigma\|_{\mathrm{tr}} + \|\sigma - \xi\|_{\mathrm{tr}} \geq \|\rho - \xi\|_{\mathrm{tr}}$ *for all* $\rho, \sigma, \xi$*.*

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 9.** *Given a* $k$*-partite state* $\rho^{A_1 A_2 \cdots A_k}$*, suppose there are pure states* $|\psi_1\rangle, \ldots, |\psi_k\rangle$ *such that* $\langle\psi_i|\rho^{A_i}|\psi_i\rangle \geq 1 - \varepsilon_i$ *for all* $i$*. Let* $|\Psi\rangle := |\psi_1\rangle \otimes \cdots \otimes |\psi_k\rangle$ *and* $\varepsilon := \varepsilon_1 + \cdots + \varepsilon_k$*. Then* $\langle\Psi|\rho^{A_1 A_2 \cdots A_k}|\Psi\rangle \geq 1 - \varepsilon$*.*

*Proof.* We can assume without loss of generality that $|\psi_i\rangle = |0\rangle$ for all $i$. Then each $\rho^{A_i}$, when measured in the standard basis, yields the outcome $|0\rangle$ with probability at least $1 - \varepsilon_i$. By the union bound, it follows that $\rho^{A_1 A_2 \cdots A_k}$, when measured in the standard basis, yields the outcome $|\Psi\rangle = |0\rangle^{\otimes k}$ with probability at least $1 - \varepsilon$. Hence $\left\langle \Psi \middle| \rho^{A_1 A_2 \cdots A_k} \middle| \Psi \right\rangle \geq 1 - \varepsilon$. □

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

**Proposition 10.** *$\langle \psi | \rho | \psi \rangle + \| \rho - |\psi\rangle \langle \psi | \|_{\text{tr}}^2 \leq 1$ for all $\rho$ and $|\psi\rangle$.*

The most general kind of operation on quantum states is called a *superoperator*. Any superoperator $\Phi$ acting on $n$ qubits can be expressed in the following *operator-sum representation*: $\Phi(\rho) = \sum_{i=1}^{2^{2n}} E_i \rho E_i^\dagger$, where $\sum_{i=1}^{2^{2n}} E_i E_i^\dagger = I$.

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 11.** *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 $\varepsilon/2$.*

*Proof.* Choose a basis that diagonalizes $\rho$, so that $\rho = \text{diag}(\lambda_1, \ldots, \lambda_N)$ where $\lambda_1, \ldots, \lambda_N$ are $\rho$'s eigenvalues. By assumption, $\lambda_i < 1 - \varepsilon$ for every $i$. So given any mixed state $\sigma$, a swap test between $\rho$ and $\sigma$ accepts with probability

$$\frac{1 + \text{tr}(\rho\sigma)}{2} = \frac{1}{2} + \frac{1}{2} \sum_{i=1}^{N} \lambda_i \sigma_{ii} < \frac{1}{2} + \frac{1-\varepsilon}{2} \sum_{i=1}^{N} \sigma_{ii} \leq 1 - \frac{\varepsilon}{2}.$$

□

## 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 a sequence of four steps:

(1) In Section 3.1, we give a sequence of classical reductions, from the original 3Sat problem to a different NP-complete problem that we will actually use.

(2) In Section 3.2, we describe a protocol for the special case where Merlin's message to Arthur is "proper": that is, of the form $\frac{1}{\sqrt{N}} \sum_{i=1}^{N} (-1)^{x_i} |i\rangle$ for some Boolean $x_1, \ldots, x_N$.

(3) In Section 3.3, we generalize our protocol to the case where the Merlins send Arthur $\widetilde{O}(\sqrt{n})$ witnesses, which are not necessarily proper but which are guaranteed to be identical to each other.

(4) In Section 3.4, we remove the restriction that the states be identical.

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 12.** *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 $m \geq n$ 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 [11]. Next we use a reduction due to Papadimitriou and Yannakakis [21], which makes every variable occur in exactly 29 clauses, without destroying the soundness gap. Finally we use a gadget due to Khanna et al. [12], which converts from 3SAT to 2-OUT-OF-4-SAT, without destroying either the soundness gap or the property of being balanced. Note that the reduction of Dinur [11] 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 12, 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 13.** *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 $B_r$:

$$\left|\psi_{ijkl}\right\rangle := \frac{1}{2}\left[(-1)^{x_i}\left|i\right\rangle + (-1)^{x_j}\left|j\right\rangle + (-1)^{x_k}\left|k\right\rangle + (-1)^{x_\ell}\left|\ell\right\rangle\right].$$

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 $\left|\psi_{ijk\ell}\right\rangle$ 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 $\left|\psi_{ijk\ell}\right\rangle$, 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 $\left|\psi\right\rangle$ 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 $\left|\psi\right\rangle$ 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 $\left|\varphi_1\right\rangle, \ldots, \left|\varphi_K\right\rangle$ 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, $\left|\varphi_i\right\rangle = \left|\varphi\right\rangle$ 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 $\left|\varphi\right\rangle$, 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 $\left|\varphi\right\rangle$ in an orthonormal basis, which contains the vectors
>
> $$\frac{\left|i\right\rangle + \left|j\right\rangle}{\sqrt{2}}, \frac{\left|i\right\rangle - \left|j\right\rangle}{\sqrt{2}}$$
>
> for every edge $(i, j) \in \mathcal{M}$. If for some $(i, j) \in \mathcal{M}$, the two outcomes $\frac{\left|i\right\rangle + \left|j\right\rangle}{\sqrt{2}}$ and $\frac{\left|i\right\rangle - \left|j\right\rangle}{\sqrt{2}}$ 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 $\left|\psi\right\rangle$ 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 14.** *The protocol has constant soundness (again, assuming the $\left|\varphi_i\right\rangle$'s are all identical).*

To prove Theorem 14, 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 $\left|\varphi\right\rangle$ is $\varepsilon$-close in trace distance to some proper state $\left|\psi\right\rangle$. Then provided we choose $\varepsilon > 0$ sufficiently small, Lemma 13, combined with Proposition 7, already implies that the Satisfiability Test rejects with constant probability. So our task reduces to proving the following:

**Claim 15.** *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 [9] shows that the Birthday Paradox occurs in the nonuniform case as well:

**Lemma 16** (Generalized Birthday Paradox [9])**.** *Suppose there are $N$ days in the year, and each person's birthday is drawn independently from the same distribution (not necessarily uniform). Then if there are $\Theta(\sqrt{N})$ people, at least two of them share a birthday with $\Omega(1)$ probability. (Indeed, the probability of a collision is minimized precisely when the distribution over birthdays is uniform.)*

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 will 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"

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

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

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

Then the key fact is the following:

**Theorem 17.** *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 17 is deferred to Appendix 8.

Assuming Theorem 17, we can complete the proof of Claim 15, and hence of Theorem 14. 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 17—with overwhelming probability that still leaves $\Theta(\sqrt{N})$

"good" measurement outcomes. Then by Lemma 16, 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, which removes the symmetry restriction.

**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.

It is clear that the above 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 first step in the analysis is simply to discard all states $|\varphi_k\rangle$ such that $|\langle\varphi_1|\varphi_k\rangle| < 1 - \delta$. By Proposition 10, the remaining $K' \geq (1 - \gamma) K$ states are all $\sqrt{2\delta}$-close to $|\varphi_1\rangle$ in trace distance.

Now given a matching $\mathcal{M}$ on $[N]$, let $\mathcal{S}_c$ be the set of edges in $\mathcal{M}$ that are $c$-unbalanced with respect to $|\varphi_1\rangle$, in the sense defined in Section 3.3. Then Theorem 17 implies that $\mathcal{S}_c$ is $d$-large with respect to $|\varphi_1\rangle$ with probability at least 1/3 over the choice of $\mathcal{M}$. Suppose that it is.

Call a measurement outcome $|i\rangle \pm |j\rangle$ *good* if $(i,j) \in \mathcal{S}_c$. Then when Arthur performs the Uniformity Test, we simply discard all states for which the outcome is not good. Since all of the states are $\sqrt{2\delta}$-close to $|\varphi_1\rangle$, and since $\mathcal{S}_c$ is $d$-large with respect to $|\varphi_1\rangle$, with overwhelming probability this still leaves us with $K'' \approx (d - \sqrt{2\delta})K'$ states. Call those states $|\xi_1\rangle, \ldots, |\xi_{K''}\rangle$.

Let $\widetilde{\mathcal{M}} = \{|i\rangle \pm |j\rangle : (i,j) \in \mathcal{M}\}$. Given a state $|\varphi\rangle$, let $\mathcal{D}_{|\varphi\rangle}$ be the probability distribution over $\widetilde{\mathcal{M}}$ induced by measuring $|\varphi\rangle$ according to $\mathcal{M}$. Then we know that $\left\|\mathcal{D}_{|\varphi_1\rangle} - \mathcal{D}_{|\xi_k\rangle}\right\| \leq \sqrt{2\delta}$ for all $k \in [K'']$. Next let $\mathcal{D}'_{|\varphi\rangle}$ be the distribution over $\widetilde{\mathcal{M}}$ induced by measuring $|\varphi\rangle$, and then conditioning on the outcome being good. Then we claim that

$$\left\|\mathcal{D}'_{|\xi_k\rangle} - \mathcal{D}'_{|\varphi_1\rangle}\right\| \leq \frac{\sqrt{2\delta}}{d - \sqrt{2\delta}}$$

for all $k \in [K'']$. This is so because of the following simple fact:

**Proposition 18.** *Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be probability distributions, let $E$ be an event, and let $\mathcal{D}'_1$ and $\mathcal{D}'_2$ denote $\mathcal{D}_1$ and $\mathcal{D}_2$ respectively conditioned on $E$. Suppose $\|\mathcal{D}_1 - \mathcal{D}_2\| \leq \epsilon$ and $\Pr_{x \in \mathcal{D}_1}[E(x)] \geq a$. Then $\|\mathcal{D}'_1 - \mathcal{D}'_2\| \leq \frac{\epsilon}{a-\epsilon}$.*

The proof of Proposition 18 is deferred to Appendix 9.

By construction, every measurement outcome $|i\rangle \pm |j\rangle$ in the support of every $\mathcal{D}'_{|\xi_k\rangle}$ corresponds to an edge $(i,j)$ that is $c$-unbalanced with respect to $|\varphi_1\rangle$. But this still leaves a key question unanswered: is $(i,j)$ reasonably unbalanced with respect to $|\xi_k\rangle$ itself? The following lemma will imply that it is, with high probability over $\mathcal{D}'_{|\xi_k\rangle}$: in particular that

$$\Pr_{|i\rangle \pm |j\rangle \in \mathcal{D}'_{|\xi_k\rangle}}\left[(i,j) \text{ is } \frac{c}{4}\text{-unbalanced w.r.t. } |\xi_k\rangle\right] \geq 1 - \frac{16\sqrt{2\delta}}{c\left(d - \sqrt{2\delta}\right)}$$

for all $k \in [K'']$.

**Lemma 19.** *Let $\mathcal{D} = (p_x, q_x)_{x \in [N]}$ and $\mathcal{D}' = (p'_x, q'_x)_{x \in [N]}$ be any two probability distributions over the set $[N] \times \{0,1\}$. Suppose that $\|\mathcal{D} - \mathcal{D}'\| \leq \mu$, and that $2p_x q_x \geq c(p_x + q_x)^2$ for every $x \in [N]$. Let $\mathcal{S}$ be the set of all $x \in [N]$ such that $2p'_x q'_x \geq c'(p'_x + q'_x)^2$. Then*

$$\sum_{x \in \mathcal{S}} \left(p'_x + q'_x\right) \geq 1 - \frac{8\mu}{c - 2c'},$$

*for all constants $c \in (0, 1/2)$ and $c' \in (0, c/2)$.*

The proof of Lemma 19 is deferred to Appendix 9.

We now need one last conditioning step: discard all states $|\xi_k\rangle$ for which the measurement outcome is not $c/4$-unbalanced with respect to $|\xi_k\rangle$. By Lemma 19, with overwhelming probability this still leaves us with $K''' \approx K''$ states (for suitable choices of $c$, $d$, and $\delta$). Call those states $|\varsigma_1\rangle, \ldots, |\varsigma_{K'''}\rangle$.

Given any state $|\varphi\rangle$, let $\mathcal{D}''_{|\varphi\rangle}$ be the probability distribution over $|i\rangle \pm |j\rangle \in \widetilde{\mathcal{M}}$ obtained by starting from $\mathcal{D}'_{|\varphi\rangle}$, and then conditioning on the edge $(i,j)$ being $c/4$-unbalanced with respect to $|\varphi\rangle$. Then

$$\begin{aligned}
\left\|\mathcal{D}''_{|\varsigma_k\rangle} - \mathcal{D}'_{|\varphi_1\rangle}\right\| &\leq \left\|\mathcal{D}''_{|\varsigma_k\rangle} - \mathcal{D}'_{|\xi_k\rangle}\right\| + \left\|\mathcal{D}'_{|\varsigma_k\rangle} - \mathcal{D}'_{|\varphi_1\rangle}\right\| \\
&\leq \frac{16\sqrt{2\delta}}{c\left(d - \sqrt{2\delta}\right)} + \frac{\sqrt{2\delta}}{d - \sqrt{2\delta}} \\
&\leq \frac{17\sqrt{2\delta}}{c\left(d - \sqrt{2\delta}\right)}
\end{aligned}$$

for all $k \in [K''']$. So by the triangle inequality,

$$\left\| \mathcal{D}''_{|\varsigma_k\rangle} - \mathcal{D}''_{|\varsigma_\ell\rangle} \right\| \leq \frac{34\sqrt{2\delta}}{c\left(d - \sqrt{2\delta}\right)}$$

for all $k, \ell \in [K''']$.

So to sum up: we have $K''' = \Theta(\sqrt{N})$ samples from $\widetilde{\mathcal{M}}$, drawn independently from probability distributions $\mathcal{D}''_{|\varsigma_1\rangle}, \ldots, \mathcal{D}''_{|\varsigma_{K'''}\rangle}$ respectively. The distributions $\mathcal{D}''_{|\varsigma_k\rangle}$ have bounded variation distance from one another. We also know, because of the way the $\mathcal{D}''_{|\varsigma_k\rangle}$'s were constructed, that if Arthur finds a collision among the $K'''$ samples (i.e., two samples of the form $|i\rangle \pm |j\rangle$ for some $(i, j)$), then that collision will also be a disagreement with constant probability. Thus, the one remaining task is to show that Arthur finds a collision with constant probability.

Showing this amounts to generalizing the Birthday Paradox still 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 20.** *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*

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

In Appendix 10, we present a proof of Theorem 20 based on the second moment method. (Indeed, our proof works even if the $X_i$'s are only 4-wise independent.)

By Theorem 20, Arthur will find a collision among the $K''' = \Theta(\sqrt{N})$ remaining samples with constant probability. Then by the definition of the $\mathcal{D}''_{|\varsigma_k\rangle}$'s, this collision will be a disagreement with constant probability, thereby causing Arthur to reject.

So in summary, 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 polylog $m$ factor, and repeat the whole protocol 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 [11]. 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 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 21.** *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.*

*Proof Sketch.* Assume without loss of generality that either $f$ is identically zero, or else there exists a unique "marked item" $x^*$ such that $f(x^*) = 1$. Suppose it were possible to convince Arthur that $x^*$ exists by giving him unentangled witnesses $|\varphi_1\rangle, \ldots, |\varphi_K\rangle$ with $Q$ qubits in total. Then given these witnesses, Arthur's verification algorithm must query $f(x^*)$ with non-negligible amplitude $\beta = \Omega(1/\operatorname{poly}(n))$. For otherwise, by the hybrid argument of Bennett, Bernstein, Brassard, and Vazirani [5], Arthur's verification algorithm would not have $\Omega(1/\operatorname{poly}(n))$ soundness (i.e., Arthur would fail to detect a change in $f(x^*)$ from 1 to 0). But this means that Arthur's algorithm can be modified to one that uses no witnesses, and that *finds* $x^*$ with probability at least $2^{-Q}\beta$ (for Arthur can simply replace $|\varphi_1\rangle, \ldots, |\varphi_K\rangle$ by the $Q$-qubit maximally mixed state).

On the other hand, we know from the result of Bennett et al. [5] mentioned previously that if $x^*$ is uniformly random, then after $T$ queries, Arthur can have found $x^*$ with probability at most $4T^2/2^n$. Solving $2^{-Q}\beta \leq 4T^2/2^n$ for $Q$, we find that

$$Q \geq \log \frac{\beta 2^n}{4T^2} \geq n - O(\log n).$$

□

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 quantum 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 all of quantum information theory. One particular entanglement measure—the *entanglement of formation* $E_F$ defined by Bennett et al. [7]—will be particularly useful for us.

To define $E_F$ we first need some other concepts. Given a mixed state $\sigma$, the *von Neumann entropy* of $\sigma$ is $S(\sigma) := H(\{\lambda_i\})$, where $H(\{p_i\}) = -\sum_i p_i \log_2 p_i$ is the binary entropy function and $\{\lambda_i\}$ are the eigenvalues of $\sigma$. Given a bipartite pure state $|\psi^{AB}\rangle$, let $\sigma^A$ and $\sigma^B$ be the reduced states of the $A$ and $B$ registers respectively. Then it is not hard to show that $S(\sigma^A) = S(\sigma^B)$. We call this quantity the *entanglement entropy* of $|\psi^{AB}\rangle$, or $E(|\psi^{AB}\rangle)$. We can then define $E_F(\rho^{AB})$ as a weighted average of entanglement entropy, minimized over all purifications of $\rho^{AB}$:

**Definition 22.** *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 $\rho^{AB} = \sum_i p_i |\psi_i\rangle \langle\psi_i|$.*

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 23.** *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.*

*Proof.* Let $S(\rho||\sigma)$ be the *quantum relative entropy* between mixed states $\rho$ and $\sigma$ (see Nielsen and Chuang [19] for a definition). Then Vedral and Plenio [24] showed that

$$E_F(\rho^{AB}) \geq \min S\left(\rho^{AB}||\sigma^{AB}\right),$$

where the minimum is taken over all separable states $\sigma^{AB}$. Also, it is known (see Klauck et al. [14] and Ohya and Petz [20]) that

$$S(\rho^{AB}||\sigma^{AB}) \geq \frac{1}{2}\left\|\rho^{AB} - \sigma^{AB}\right\|_{\text{tr}}^2.$$

Putting these results together, if $E_F(\rho^{AB}) \leq \varepsilon$ then there exists a separable state $\sigma^{AB}$ such that $S\left(\rho^{AB}||\sigma^{AB}\right) \leq \varepsilon$, and hence $\left\|\rho^{AB} - \sigma^{AB}\right\|_{\text{tr}} \leq \sqrt{2\varepsilon}$. □

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

**Lemma 24.** *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 25** (Additivity Conjecture). *$E_F$ is superadditive.*

Shor [23] showed that Conjecture 25 is equivalent to several other additivity conjectures in quantum information theory, including the additivity of the Holevo capacity for quantum channels.

For experts in quantum information theory, let us make a few other remarks about the use of entanglement measures in this paper.

First, call an entanglement measure $E$ *weakly superadditive* if it satisfies

$$E\left(\rho^{A_1 A_2 \cdots A_k, B_1 B_2 \cdots B_k}\right) \geq \frac{c}{k} \sum_{i,j=1}^{k} E(\rho^{A_i B_j}),$$

for some constant $c$ independent of $k$. Weak superadditivity is, in particular, implied by the following inequality:

$$E(\rho^{AA',BB'}) \geq \frac{1}{2}\left[E\left(\rho^{AB}\right) + E(\rho^{AB'}) + E(\rho^{A'B}) + E(\rho^{A'B'})\right]$$

which in turn is implied by ordinary superadditivity. Although we will state our results in terms of the usual Additivity Conjecture, it will be evident from the proofs that all we ever need is the following:

**Conjecture 26** (Weak Additivity Conjecture)**.** *$E_F$ is weakly superadditive.*

Of course, the above conjecture might be easier to prove than full Additivity, and might be true even if the Additivity Conjecture fails.

Second, while $E_F$ is conjectured to be superadditive, it badly violates the stronger *monogamy inequality* $E(\rho^{A,BB'}) \geq E\left(\rho^{AB}\right) + E(\rho^{AB'})$.[11] As an example, consider again the maximally antisymmetric state

$$|\psi\rangle = \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} (-1)^{\operatorname{sgn}(\sigma)} |\sigma(1)\rangle \cdots |\sigma(N)\rangle .$$

The entanglement of formation between the first register of $|\psi\rangle$ and the remaining $N-1$ registers is at most $\log N$. Yet the entanglement of formation between the first register and any *one* other register can be shown to be $\Omega(1)$. This is unfortunate for us, for had $E_F$ satisfied the monogamy inequality, we would have been able to use it to show $\mathsf{QMA}(2) \subseteq \mathsf{PSPACE}$.

Third, a different entanglement measure—the *squashed entanglement* $E_{sq}$ of Christandl and Winter [10]—is known to satisfy both superadditivity and the stronger monogamy inequality. The trouble with $E_{sq}$ is that it badly violates the analogue of Lemma 23: there exist $N \times N$-dimensional bipartite states $\rho^{AB}$ such that $E_{sq}(\rho^{AB}) = O\left(\frac{\log N}{N}\right)$, yet $\rho^{AB}$ has trace distance $\Omega(1)$ to any separable state. The example that shows this is once again the maximally antisymmetric state, which seems like the "universal counterexample" of entanglement theory! This is why we cannot use squashed entanglement in this paper, and must instead use entanglement of formation.

[11] Note that the monogamy inequality implies superadditivity, via

$$\begin{aligned} E(\rho^{AA',BB'}) &\geq E(\rho^{AA',B}) + E(\rho^{AA',B'}) \\ &\geq E(\rho^{AB}) + E(\rho^{A'B}) + E(\rho^{AB'}) + E(\rho^{A'B'}) \\ &\geq E(\rho^{AB}) + E(\rho^{A'B'}). \end{aligned}$$

## 4.2 The Two-Prover Case

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

**Theorem 27.** *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 24, 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

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

$$\leq 2T\left(n\right)r\left(n\right)\left(q\left(n\right)-T\left(n\right)\right).$$

So if we define

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

then the convexity of $E_F$ implies that

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

using the fact that $T\left(n\right)\leq q\left(n\right)/2$. By Lemma 23, this means that $\sigma$ is $\sqrt{8T\left(n\right)r\left(n\right)/q\left(n\right)}$-close to a separable state, as claimed. □

## 4.3 The $k$-Prover Case

Recall that Kobayashi et al. [15] showed that amplification of $\mathsf{QMA}\left(k\right)$ protocols implies $\mathsf{QMA}\left(k\right)=\mathsf{QMA}\left(2\right)$ 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 $\mathsf{QMA}\left(k\right)$ to $\mathsf{QMA}\left(2\right)$. Unfortunately, the result of Kobayashi et al. requires amplification for all $\mathsf{QMA}\left(k\right)$, while we have only shown that additivity implies amplification for $\mathsf{QMA}\left(2\right)$. In this section we solve the problem by strengthening Kobayashi et al.'s result. In particular, we will show that *any* $\mathsf{QMA}\left(k\right)$ *protocol with constant soundness can be simulated by a* $\mathsf{QMA}\left(2\right)$ *protocol with soundness* $\Omega\left(1/k\right)$. Combined with Theorem 27, this will then imply that $\mathsf{QMA}\left(k\right)=\mathsf{QMA}\left(2\right)$ for all $k\geq 2$ assuming the Additivity Conjecture.

**Theorem 28.** $\mathsf{QMA}\left(k,a,b\right)\subseteq\mathsf{QMA}\left(2,1-\frac{\left(b-a\right)^2}{8k},1-2^{-n}\right)$.

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

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

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

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 $i\in\left[k\right]$ 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 $\rho^{A_1 A_2 \cdots A_k}$ 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 $\rho^{A_1 A_2 \cdots A_k}$ and $\rho^{B_1 B_2 \cdots B_k}$ causes the $\mathsf{QMA}(2)$ protocol to reject with probability at least $\frac{\delta^2}{8k}$. We consider two cases.

First suppose $\rho^{A_1 A_2 \cdots A_k}$ 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 $\rho^{A_1 A_2 \cdots A_k}$ is $\varepsilon$-far in trace distance from any separable pure state. Then by Proposition 10, we have $\left\langle \Psi | \rho^{A_1 A_2 \cdots A_k} | \Psi \right\rangle < 1 - \varepsilon^2$ for all separable pure states $|\Psi\rangle$. So taking the contrapositive of Proposition 9, 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 11.

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

Combining Theorem 28 with Theorem 27 now yields the following:

**Corollary 29.** *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 27. 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 30.** *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 27, 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 27 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 $\mathsf{QMA}(2; w(n)) \subseteq \mathsf{QMA}(2)$.

### 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, we will now 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 31.** $\mathsf{QMA}(k,a,b) \subseteq \mathsf{SymQMA}(k,a,b) \subseteq \mathsf{QMA}\left(k, 1-\frac{(b-a)^2}{8k}, 1-2^{-n}\right)$.

*Proof.* For the first containment, have each Merlin in the $\mathsf{SymQMA}$ protocol send $k$ witnesses (for a total of $k^2$ witnesses). Then simulate the $\mathsf{QMA}$ protocol by using the $i^{th}$ witness from the $i^{th}$ Merlin for all $i \in [k]$.

For the second containment, first observe that

$$\mathsf{SymQMA}(k,a,b) \subseteq \mathsf{SymQMA}\left(k, 1-(b-a), 1-2^{-n}\right),$$

completely analogously to Lemma 6. Let $\delta = b-a$. Then to simulate a $\mathsf{SymQMA}(k, 1-\delta, 1-2^{-n})$ protocol without the symmetry promise we do the following. Let $|\varphi_i\rangle$ be the witness sent by the $i^{th}$ Merlin. Then

- With 1/2 probability, Arthur performs a swap test between $|\varphi_1\rangle$ and a random other witness, and accepts if and only if the swap test accepts.
- With 1/2 probability, Arthur runs the $\mathsf{SymQMA}$ protocol as if the witnesses were identical.

In the completeness case, it is clear that Arthur accepts with probability greater than $1-2^{-n}$.

To show soundness we consider two cases, just like in Theorem 28. Let $|\Phi\rangle = |\varphi_1\rangle \otimes \cdots \otimes |\varphi_k\rangle$. First suppose $|\Phi\rangle$ is $\varepsilon$-close in trace distance to $|\varphi_1\rangle^{\otimes k}$. Then by Proposition 7, when he runs the $\mathsf{SymQMA}$ protocol Arthur will reject with probability at least $\delta - \varepsilon$.

Next suppose $|\Phi\rangle$ is $\varepsilon$-far from $|\varphi_1\rangle^{\otimes k}$. Then $|\langle\varphi_1|^{\otimes k}|\Phi\rangle|^2 < 1-\varepsilon^2$ by Proposition 10. So by Proposition 9, we have

$$\sum_{i=1}^{k}\left(1-|\langle\varphi_1|\varphi_i\rangle|^2\right) > \varepsilon^2.$$

Hence when Arthur performs a swap test, he rejects with probability greater than $\frac{\varepsilon^2}{2k}$.

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

Combining Lemma 31 with Theorem 28, we immediately get the following.

**Theorem 32.** *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 33** (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 33 might be true. In studying probabilistic complexity classes, one typically assumes amplification theorems will hold unless there is some obvious 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 $\mathsf{NP} \subseteq \mathsf{DTIME}(n^{\operatorname{polylog} n})$. 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 33 implies $\mathsf{QMA}(k) \subseteq \mathsf{PSPACE}$ for all $k$. We know from Kobayashi et al. [15] 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 33 implies $\mathsf{QMA}(2) \subseteq \mathsf{PSPACE}$.

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

**Lemma 34** (**[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 35.** *Conjecture 33 implies* $\mathsf{QMA}(2) \subseteq \mathsf{QMA}_{\mathsf{PSPACE}}$*.*

*Proof.* Let $L$ be a language in $\mathsf{QMA}(2)$. By Conjecture 33, 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 34, 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 34. 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 $T2^{-2s(n)} < 1/3$. □

**Lemma 36.** $\mathsf{QMA}_{\mathsf{PSPACE}} = \mathsf{PSPACE}$*.*

*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 $\operatorname{Tr}(A^{2p(n)})$ is just an exponential-size linear algebra problem, which can be solved in $\mathsf{PSPACE}$. On the other hand $\operatorname{Tr}(A^{2p(n)})$ 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. □

Combining Lemma 35 with Lemma 36 now yields the main result.

**Theorem 37.** *Conjecture 33 implies* $\mathsf{QMA}(2) \subseteq \mathsf{PSPACE}$.

Or if we "scale down by an exponential," Conjecture 33 implies that

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

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 $\operatorname{polylog} n$. Assuming Conjecture 33, this means in particular that the 3-Coloring protocol of Blier and Tapp [8] cannot be amplified to constant soundness, unless $\mathsf{NP} \subseteq \mathsf{DTIME}(n^{\operatorname{polylog} n})$.

Theorem 37 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.[12] (Of course, such an algorithm would require a strong amplification procedure as a subroutine.) We now state the connection more precisely.

**Theorem 38.** *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 33, there exists a deterministic algorithm that takes $M$ as input, approximates $p(M)$ to within additive error $\varepsilon$, and runs in time $N^{\operatorname{polylog} N/\operatorname{poly}(\varepsilon)}$.*

# 6 Nonexistence of Perfect Disentanglers

**Definition 39.** *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 $\log \dim \mathcal{H} = O(\operatorname{poly}(\log \dim \mathcal{K}))$—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 40** (Watrous)**.** *For all constants $\varepsilon, \delta < 1$, any $(\varepsilon, \delta)$-disentangler requires* $\dim \mathcal{H} = 2^{\Omega(\dim \mathcal{K})}$.

A proof of Conjecture 40 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.

Here we show that, at least in the case $\varepsilon = \delta = 0$, no disentangler exists in *any* finite dimension. The counterexamples in Section 1.6 imply that this result would be false if we let either $\varepsilon$ or $\delta$ be nonzero.

**Theorem 41.** *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* $\dim \mathcal{H} = \infty$.

[12] 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 [8].

*Proof.* For any pure state $|\alpha\rangle \in \mathcal{K}$, by assumption there exists a state $\rho_\alpha$ such that $\Phi(\rho_\alpha) = |\alpha\rangle\langle\alpha| \otimes |\alpha\rangle\langle\alpha|$. We can assume without loss of generality that $\rho_\alpha = |\phi_\alpha\rangle\langle\phi_\alpha|$ is pure. Also, suppose $\dim \mathcal{H}$ is finite. Then $\Phi$ admits an operator-sum representation $\Phi(\rho) = \sum_{i=1}^k E_i \rho E_i^\dagger$ where $\sum_{i=1}^k E_i^\dagger E_i = I$. We then have

$$\Phi(|\phi_\alpha\rangle\langle\phi_\alpha|) = \sum_{i=1}^k E_i |\phi_\alpha\rangle\langle\phi_\alpha| E_i^\dagger = |\alpha\rangle\langle\alpha| \otimes |\alpha\rangle\langle\alpha| .$$

Hence $E_i|\phi_\alpha\rangle$ must be a multiple of $|\alpha\rangle|\alpha\rangle$ for all $i$ and $\alpha$; that is, there exist constants $c_{\alpha,i}$ such that $E_i|\phi_\alpha\rangle = c_{\alpha,i}|\alpha\rangle|\alpha\rangle$.

Now let $|\alpha\rangle, |\beta\rangle$ be any two pure states in $\mathcal{K}$ with $|\alpha\rangle \neq |\beta\rangle$. Also let $|\psi\rangle = a|\phi_\alpha\rangle + b|\phi_\beta\rangle$ for some nonzero real numbers $a, b$. Then

$$\begin{aligned}\Phi(|\psi\rangle\langle\psi|) &= a^2\Phi(|\phi_\alpha\rangle\langle\phi_\alpha|) + b^2\Phi(|\phi_\beta\rangle\langle\phi_\beta|) + ab\Phi(|\phi_\alpha\rangle\langle\phi_\beta|) + ab\Phi(|\phi_\beta\rangle\langle\phi_\alpha|) \\ &= a^2|\alpha\rangle\langle\alpha| \otimes |\alpha\rangle\langle\alpha| + b^2|\beta\rangle\langle\beta| \otimes |\beta\rangle\langle\beta| + abc|\alpha\rangle\langle\beta| \otimes |\alpha\rangle\langle\beta| + ab\overline{c}|\beta\rangle\langle\alpha| \otimes |\beta\rangle\langle\alpha| ,\end{aligned}$$

where

$$c = \sum_{i=1}^k c_{\alpha,i}\overline{c}_{\beta,i}.$$

We claim that $c = 0$. To see this, recall that $\Phi(|\psi\rangle\langle\psi|)$ is a separable mixed state, and consider any decomposition of $\Phi(|\psi\rangle\langle\psi|)$ into separable pure states. Every pure state in the support of $\Phi(|\psi\rangle\langle\psi|)$ must have the form $x|\alpha\rangle|\alpha\rangle + y|\beta\rangle|\beta\rangle$. But by the assumption $|\alpha\rangle \neq |\beta\rangle$, such a state cannot be separable unless $x = 0$ or $y = 0$. Hence the only separable pure states in the support of $\Phi(|\psi\rangle\langle\psi|)$ are $|\alpha\rangle|\alpha\rangle$ and $|\beta\rangle|\beta\rangle$. Therefore $abc = 0$. But $a$ and $b$ were nonzero, so $c = 0$ as claimed.

This means in particular that $\Phi(|\phi_\alpha\rangle\langle\phi_\beta|) = 0$ for all $|\alpha\rangle \neq |\beta\rangle$. Hence

$$\begin{aligned}\langle\phi_\beta|\phi_\alpha\rangle &= \sum_{i=1}^k \langle\phi_\beta| E_i^\dagger E_i |\phi_\alpha\rangle \\ &= \operatorname{Tr}\left(\sum_{i=1}^k E_i|\phi_\alpha\rangle\langle\phi_\beta| E_i^\dagger\right) \\ &= \operatorname{Tr}\left(\Phi(|\phi_\alpha\rangle\langle\phi_\beta|)\right) \\ &= 0.\end{aligned}$$

So for different $|\alpha\rangle$'s, the states $|\phi_\alpha\rangle$ are all orthogonal, and since the number of $|\alpha\rangle$'s is infinite, $\dim\mathcal{H}$ must be infinite as well. □

# 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 [26]) 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 $\widetilde{O}\left(\sqrt{n}\right)$, or alternatively, give evidence against this possibility? For example, can we show that $\Omega\left(\sqrt{n}\right)$ witnesses are information-theoretically required for the Uniformity Test? Finally, can we show unconditionally that $\mathsf{QMA}\left(2\right) \subseteq \mathsf{EXP}$?

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 $\mathsf{QMA}\left(k\right)$, does it matter if the amplitudes are reals or complex numbers? For $\mathsf{BQP}$ and $\mathsf{QMA}$, it is not hard to show that this distinction is irrelevant. Interestingly, though, the usual equivalence proofs break down for $\mathsf{QMA}\left(k\right)$. As evidence that $\mathsf{QMA}\left(k\right)$ might actually be sensitive to the difference between reals and complex numbers, consider the analysis of our 3Sat protocol: in Appendix 8, the result that we need becomes much simpler prove when we assume all amplitudes are real.

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

Can we prove Conjecture 40: that there is no $(\varepsilon,\delta)$-disentangler with $\operatorname{poly}\left(n\right)$ 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 $\mathsf{QMA}^U \neq \mathsf{QMA}^U\left(2\right)$? Can we at least show that Conjecture 40 would imply the existence of such an oracle?

### 7.3 $\mathsf{QMA}\left(k\right)$ 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 $\mathsf{BellQMA}\left(k\right)$ to be the subclass of $\mathsf{QMA}\left(k\right)$ 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 $\mathsf{BellQMA}\left(k\right) = \mathsf{QMA}\left(k\right)$? Does $\mathsf{BellQMA}\left(k\right) = \mathsf{BellQMA}\left(2\right)$ for all $k \geq 2$? Note that it is trivial to show amplification for $\mathsf{BellQMA}\left(k\right)$. 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 27 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 40; an anonymous reviewer for greatly simplifying the proof of Lemma 24 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] M. Ben-Or, S. Goldwasser, J. Kilian, and A. Wigderson. Multi-prover interactive proofs: how to remove the intractability assumptions. In *Proc. ACM STOC*, pages 113–131, 1988.

[5] C. Bennett, E. Bernstein, G. Brassard, and U. Vazirani. Strengths and weaknesses of quantum computing. *SIAM J. Comput.*, 26(5):1510–1523, 1997. quant-ph/9701001.

[6] 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.

[7] 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.

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

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

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

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

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

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

[14] H. Klauck, A. Nayak, A. Ta-Shma, and D. Zuckerman. Interaction in quantum communication. *IEEE Trans. Information Theory*, 53(6):1970–1982, 2007. Earlier version in STOC'2001. quant-ph/0603135.

[15] 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.

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

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

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

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

[20] M. Ohya and D. Petz. *Quantum Entropy and its Use.* Springer, 1993.

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

[22] R. Raz. A parallel repetition theorem. In *Proc. ACM STOC*, pages 447–456, 1995.

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

[24] V. Vedral and M. B. Plenio. Entanglement measures and purification procedures. *Phys. Rev. A*, 57:1619–1633, 1998. quant-ph/9707035.

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

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

[27] 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.

# 8 Appendix: Unbalanced Edges in Random Matchings

The goal of this appendix is to prove Theorem 17, which we now restate in a more careful way.

**Theorem.** *There exist constants $c, d > 0$ for which the following holds. Let $N$ be even and sufficiently large. Suppose the state $|\varphi\rangle = \alpha_1 |1\rangle + \cdots + \alpha_N |N\rangle$ is $\varepsilon$-far in trace distance from any proper state (that is, any state of the form $\frac{1}{\sqrt{N}} \sum_{i=1}^{N} (-1)^{x_i} |i\rangle$ where $x_1, \ldots, x_N \in \{0,1\}$). Let $\mathcal{M}$ be a matching on $[N]$ chosen uniformly at random, and let $\mathcal{S}$ be the set of edges $(i,j) \in \mathcal{M}$ that are "$c\varepsilon^8$-unbalanced," meaning that*

$$\left|\alpha_i^2 - \alpha_j^2\right|^2 \geq 2c\varepsilon^8 \left(|\alpha_i|^2 + |\alpha_j|^2\right)^2.$$

*Then*

$$\sum_{(i,j)\in\mathcal{S}} \left(|\alpha_i|^2 + |\alpha_j|^2\right) \geq d\varepsilon^4$$

*with probability at least $1/3$ over the choice of $\mathcal{M}$.*

Given a state $|\varphi\rangle = \alpha_1 |1\rangle + \cdots + \alpha_N |N\rangle$, define the *nonuniformity* of $|\varphi\rangle$ to be

$$\operatorname{NU}(|\varphi\rangle) := \frac{1}{2} \sum_{i=1}^{N} \left| |\alpha_i|^2 - \frac{1}{N} \right|.$$

Intuitively, $\operatorname{NU}(|\varphi\rangle)$ measures whether the distribution induced by measuring $|\varphi\rangle$ in the standard basis is close to uniform or not. We will divide the proof of Theorem 17 into two cases: first that $\operatorname{NU}(|\varphi\rangle) > \varepsilon^4/100$ (the "nonuniform case"), and second that $\operatorname{NU}(|\varphi\rangle) \leq \varepsilon^4/100$ (the "uniform case").

### 8.1 The Nonuniform Case

We now prove Theorem 17 in the case $\operatorname{NU}(|\varphi\rangle) > \varepsilon^4/100$. For convenience, define $\epsilon := \varepsilon^4/100$ and $p_i := |\alpha_i|^2$. Then the condition

$$\left|\alpha_i^2 - \alpha_j^2\right|^2 \geq 2c\varepsilon^8 \left(|\alpha_i|^2 + |\alpha_j|^2\right)^2$$

is equivalent to

$$p_i^2 + p_j^2 - 2\operatorname{Re}\alpha_i^2\overline{\alpha_j}^2 \geq 2c\varepsilon^8 \left(p_i + p_j\right)^2,$$

which will certainly be true whenever $(p_i - p_j)^2 \geq 2c\varepsilon^8 (p_i + p_j)^2$, or equivalently

$$\frac{p_i}{p_j} + \frac{p_j}{p_i} \geq \frac{2 + 2c\varepsilon^8}{1 - 2c\varepsilon^8}.$$

(If $p_i = 0$ or $p_j = 0$ then we stipulate that the above inequality holds.) Thus, it suffices to prove the following classical lemma.

**Lemma 42.** *Let $N$ be even and sufficiently large. Let $(p_1, \ldots, p_N)$ be a probability distribution, and suppose*

$$\frac{1}{2}\sum_{i=1}^{N}\left|p_i - \frac{1}{N}\right| \geq \epsilon.$$

*Let $\mathcal{M}$ be a uniform random matching on $[N]$, and let $\mathcal{S}$ be the set of edges $(i,j) \in \mathcal{M}$ such that $p_i/p_j + p_j/p_i \geq 2 + \epsilon^2/16$. Then*

$$\sum_{(i,j)\in\mathcal{S}} (p_i + p_j) \geq \frac{\epsilon}{12}$$

*with probability at least $1/3$ over $\mathcal{M}$.*

Let $H$ (the "heavy elements") be the set of $i \in [N]$ such that $p_i \geq 1/N$, and let $H^* \subseteq H$ (the "very heavy elements") be the set of $i \in [N]$ such that $p_i \geq \frac{1+\epsilon/2}{N}$. Let $L$ (the "light elements") be the set of $i \in [N]$ such that $p_i < 1/N$, and let $L^* \subseteq L$ (the "very light elements") be the set of $i \in [N]$ such that $p_i \leq \frac{1-\epsilon/4}{N}$. Clearly

$$\sum_{i\in H}\left(p_i - \frac{1}{N}\right) = \sum_{i\in L}\left(\frac{1}{N} - p_i\right) = \epsilon.$$

Using this, we can prove two simple facts: that there are $\Omega(N)$ very light elements, and that the very heavy elements have total weight $\Omega(\epsilon)$.

**Proposition 43.** *$|L^*| \geq \epsilon N/2$.*

*Proof.* We have

$$\epsilon = \sum_{i\in L}\left(\frac{1}{N} - p_i\right) \leq \frac{|L^*|}{N} + (|L| - |L^*|)\frac{\epsilon}{4N}.$$

Now use $|L| \leq N$ and rearrange. □

Given any subset $A \subseteq [N]$, define the "weight" of $A$ to be $W_A := \sum_{i\in A} p_i$.

**Proposition 44.** *$W_{H^*} \geq \epsilon/2$.*

*Proof.* We have

$$\begin{aligned}\epsilon &= \sum_{i\in H}\left(p_i - \frac{1}{N}\right)\\ &= \sum_{i\in H\setminus H^*}\left(p_i - \frac{1}{N}\right) + \sum_{i\in H^*}\left(p_i - \frac{1}{N}\right)\\ &\leq N\frac{\epsilon}{2N} + W_{H^*}.\end{aligned}$$

Now subtract $\epsilon/2$ from both sides. □

To prove Lemma 42, we divide into two cases.

The first case is that $|H| \geq N/2$ (in other words, at least half the elements are heavy). In this case, we begin constructing the matching $\mathcal{M}$ by randomly assigning partners to the "very light elements" $i \in L^*$. Recall from Proposition 43 that $|L^*| \geq \epsilon N/2$. So by a standard Chernoff bound, it is easy to see that at least (say) $|L^*|/6$ elements $i \in L^*$ will be matched to partners $j \in H$, with probability $1 - o(1)$ over $\mathcal{M}$. Notice that every edge $(i,j)$ with $i \in L^*$ and $j \in H$ satisfies $p_i \leq \frac{1-\epsilon/4}{N}$ and $p_j \geq 1/N$, and therefore

$$\frac{p_i}{p_j} + \frac{p_j}{p_i} \geq 1 - \epsilon/4 + \frac{1}{1-\epsilon/4} > 2 + \frac{\epsilon^2}{16}.$$

Thus, all of these edges go into the set $\mathcal{S}$. We then have

$$\sum_{(i,j)\in\mathcal{S}}(p_i + p_j) \geq \frac{|L^*|}{6}\cdot\frac{1}{N} \geq \frac{\epsilon}{12}$$

and are done.

The second case is that $|H| < N/2$ (in other words, there are more light elements than heavy ones). In this case, we begin constructing $\mathcal{M}$ by randomly assigning partners to the "very heavy elements" $i \in H^*$. Let $B$ be the set of elements $i \in H^*$ that get matched to partners in $H$. Then since $|H| < N/2$, every element of $H^*$ goes into $B$ with probability less than $1/2$, and hence

$$\operatorname*{E}_{\mathcal{M}}[W_B] < \sum_{i\in H^*}\frac{p_i}{2} = \frac{W_{H^*}}{2}.$$

Therefore

$$\Pr_{\mathcal{M}}\left[W_B > \frac{3}{4}W_{H^*}\right] < \frac{2}{3}$$

by Markov's inequality. In other words, with probability greater than $1/3$, at least $1/4$ of the probability weight in $H^*$ gets matched to partners in $L$. Suppose this happens.

Notice that every edge $(i,j)$ with $i \in H^*$ and $j \in L$ satisfies $p_i \geq \frac{1+\epsilon/2}{N}$ and $p_j < 1/N$, and therefore

$$\frac{p_i}{p_j} + \frac{p_j}{p_i} > 1 + \epsilon/2 + \frac{1}{1-\epsilon/2} > 2 + \frac{\epsilon^2}{4}.$$

Thus, all of these edges go into the set $\mathcal{S}$. Furthermore, by the assumption $W_B \leq \frac{3}{4}W_{H^*}$, we have

$$\sum_{(i,j)\in\mathcal{S}}(p_i + p_j) \geq \sum_{i\in H^*\setminus B} p_i = W_{H^*\setminus B} \geq \frac{W_{H^*}}{4} \geq \frac{\epsilon}{8}$$

and are done.

## 8.2 The Uniform Case

We now prove Theorem 17 for states $|\varphi\rangle = \alpha_1 |1\rangle + \cdots + \alpha_N |N\rangle$ such that $\mathrm{NU}(|\varphi\rangle) \leq \varepsilon^4/100$. The first step is to define a measure of the distance from $|\varphi\rangle$ to the closest proper state, which we call the *impropriety* of $|\varphi\rangle$ or $\mathrm{imp}(|\varphi\rangle)$:

$$\mathrm{imp}(|\varphi\rangle) := \min_{|r|=1/N} \sum_{i=1}^{N} \left|\alpha_i^2 - r\right|.$$

Clearly $0 \leq \mathrm{imp}(|\varphi\rangle) \leq 2$ for all $|\varphi\rangle$, with $\mathrm{imp}(|\varphi\rangle) = 0$ if and only if $|\varphi\rangle$ is equivalent to a proper state up to a phase shift. We also have the following:

**Lemma 45.** *Suppose $|\varphi\rangle$ is $\varepsilon$-far in trace distance from any proper state. Then* $\mathrm{imp}(|\varphi\rangle) > \varepsilon^2$.

*Proof.* By Proposition 10, we have $|\langle\varphi|\psi\rangle| < \sqrt{1-\varepsilon^2} < 1 - \varepsilon^2/2$ for all proper states $|\psi\rangle$. On the other hand, suppose $\mathrm{imp}(|\varphi\rangle) \leq \varepsilon^2$. Then we will construct a proper state $|\psi\rangle$ such that $|\langle\varphi|\psi\rangle| \geq 1 - \varepsilon^2/2$, thereby obtaining the desired contradiction.

Let $r$ be a complex number with $|r| = 1/N$ that minimizes $\sum_{i=1}^{N} \left|\alpha_i^2 - r\right|$, and let $\sqrt{r}$ be a canonical square root of $r$. Also let $\beta_i := \left|\alpha_i^2 - r\right|$. Then

$$\left|\alpha_i + \sqrt{r}\right| \left|\alpha_i - \sqrt{r}\right| = \left|\alpha_i^2 - r\right| = \beta_i,$$

which means that either $|\alpha_i + \sqrt{r}| \leq \sqrt{\beta_i}$ or $|\alpha_i - \sqrt{r}| \leq \sqrt{\beta_i}$. So by setting the $\gamma_i$'s to $\sqrt{r}$ or $-\sqrt{r}$ appropriately, we can construct a state $|\psi\rangle = \gamma_1 |1\rangle + \cdots + \gamma_N |N\rangle$ that is proper up to a trivial phase factor, such that $|\alpha_i - \gamma_i| \leq \sqrt{\beta_i}$ for all $i$. Then

$$\begin{aligned}
2 - 2|\langle\varphi|\psi\rangle| &= 2 - 2\sum_{i=1}^{N} \alpha_i \overline{\gamma_i} \\
&= \sum_{i=1}^{N} \left(|\alpha_i|^2 + |\gamma_i|^2 - 2\alpha_i \overline{\gamma_i}\right) \\
&= \sum_{i=1}^{N} |\alpha_i - \gamma_i|^2 \\
&\leq \sum_{i=1}^{N} \beta_i \\
&= \mathrm{imp}(|\varphi\rangle) \\
&\leq \varepsilon^2,
\end{aligned}$$

and hence $|\langle\varphi|\psi\rangle| \geq 1 - \varepsilon^2/2$ as claimed. □

In what follows, assume $\mathrm{imp}(|\varphi\rangle) > \varepsilon^2$.

Now as in Section 8.1, let $p_i := |\alpha_i|^2$, and for any subset $A \subseteq [N]$, define the "probability weight" of $A$ to be $W_A := \sum_{i \in A} p_i$. Also, let $\delta := \varepsilon^2/5$, and let $U$ (the "$\delta$-uniform subset") be the set of all $i \in [N]$ such that $|p_i - 1/N| \leq \delta/N$. The following proposition shows that $U$ encompasses "most" of $|\varphi\rangle$, whether in terms of cardinality or in terms of probability weight.

**Proposition 46.** $|U| \geq N\left(1 - \varepsilon^2/10\right)$ *and* $W_U \geq 1 - 3\varepsilon^2/10$.

*Proof.* We have

$$\frac{\varepsilon^4}{100} \geq \mathrm{NU}\left(|\varphi\rangle\right) \geq \frac{1}{2}\left(N - |U|\right)\frac{\delta}{N},$$

hence

$$|U| \geq N\left(1 - \frac{\varepsilon^4}{50\delta}\right) = N\left(1 - \frac{\varepsilon^2}{10}\right),$$

hence

$$W_U \geq N\left(1 - \frac{\varepsilon^2}{10}\right)\left(\frac{1}{N} - \frac{\delta}{N}\right) \geq 1 - \frac{3\varepsilon^2}{10}.$$

□

Let

$$\operatorname{imp}_U\left(|\varphi\rangle\right) := \min_{|r|=1/N} \sum_{i \in U} \left|\alpha_i^2 - r\right|$$

be an analogue of impropriety that is restricted to the set $U$. By combining Lemma 45 with Proposition 46, we can now lower-bound $\operatorname{imp}_U\left(|\varphi\rangle\right)$.

**Proposition 47.** $\operatorname{imp}_U\left(|\varphi\rangle\right) \geq 3\varepsilon^2/5$.

*Proof.* For all $r$ with $|r| = 1/N$, we have

$$\sum_{i \notin U} \left|\alpha_i^2 - r\right| \leq \sum_{i \notin U} p_i + \sum_{i \notin U} \frac{1}{N} \leq \left(1 - W_U\right) + \frac{N - |U|}{N} \leq \frac{3\varepsilon^2}{10} + \frac{\varepsilon^2}{10} = \frac{2\varepsilon^2}{5}$$

by Proposition 46. Hence

$$\begin{aligned}\operatorname{imp}_U\left(|\varphi\rangle\right) &= \min_{|r|=1/N} \sum_{i \in U} \left|\alpha_i^2 - r\right| \\ &\geq \min_{|r|=1/N} \sum_{i \in [N]} \left|\alpha_i^2 - r\right| - \max_{|r|=1/N} \sum_{i \notin U} \left|\alpha_i^2 - r\right| \\ &\geq \operatorname{imp}\left(|\varphi\rangle\right) - \frac{2\varepsilon^2}{5} \\ &\geq \frac{3\varepsilon^2}{5}.\end{aligned}$$

□

We are finally ready for the geometric core of our result. Let $V$ be a collection of vectors in $\mathbb{R}^2$ (possibly with multiplicity), which consists of $\left(N \operatorname{Re} \alpha_i^2, N \operatorname{Im} \alpha_i^2\right)$ for every $i \in U$. Let $\|v\|$ be the 2-norm of $v$. Then we know by the definition of $U$ that $1 - \delta \leq \|v\| \leq 1 + \delta$ for all $v \in V$. We also know from Proposition 46 that

$$|V| = |U| \geq N\left(1 - \frac{\varepsilon^2}{10}\right) \geq 0.9N,$$

and from Proposition 47 that for all unit vectors $w \in \mathbb{R}^2$,

$$\sum_{v \in V} \|v - w\| \geq \frac{3\varepsilon^2 N}{5} \geq \frac{3\varepsilon^2 |V|}{5}.$$

Based on this information, we want to find two subsets $X, Y \subseteq V$, both of size $\Omega(|V|)$, such that $\|x - y\| = \Omega(1)$ for all $x \in X$ and $y \in Y$.

For suppose we can do this. Then just as in Section 8.1, when a matching $\mathcal{M}$ on $[N]$ is chosen uniformly at random, by a Chernoff bound it will have $\Omega(N)$ edges between the subsets of $[N]$ corresponding to $X$ and $Y$ with overwhelming probability. Assuming that happens, we will have $\left|\alpha_i^2 - \alpha_j^2\right| = \Omega(1/N)$ for every such edge $(i,j) \in \mathcal{M}$, and hence all of these edges will get added to the set $\mathcal{S}$. We will therefore have

$$\sum_{(i,j)\in\mathcal{S}} (p_i + p_j) = \Omega(1)$$

as desired. (For simplicity, we have suppressed the dependence on $\varepsilon$ here.)

What we need, then, is the following geometric lemma.

**Lemma 48.** *Let $V$ be a collection of vectors in the plane. Suppose that $1 - \delta \leq \|v\| \leq 1 + \delta$ for every $v \in V$, and that $\sum_{v\in V} \|v - w\| \geq \epsilon |V|$ for every unit vector $w \in \mathbb{R}^2$. Then provided $\delta \leq \epsilon/2$, there exist subsets $X, Y \subseteq V$, both of size at least $\epsilon |V| /40$, such that $\|x - y\| \geq \epsilon/20$ for all $x \in X$ and $y \in Y$.*[13]

*Proof.* Divide the plane into $K \in [30/\epsilon, 40/\epsilon]$ equal-sized, half-open angular sectors, centered about the origin. By the pigeonhole principle, one of these sectors (call it $S$) must contain at least $|V|/K \geq \epsilon |V|/40$ of the vectors.

Let $S'$ be the union of $S$ and its two adjacent sectors. Then we claim that at least $\epsilon |V|/40$ of the vectors must lie outside of $S'$. For suppose not. Then let $z$ be the unit vector that bisects $S$, and let

$$\theta = \frac{3}{2}\left(\frac{2\pi}{K}\right) \leq \frac{3}{2}\left(\frac{2\pi}{30/\epsilon}\right) = \frac{\pi\epsilon}{10}$$

be the angle between $z$ and the border of $S'$. Notice that by the triangle inequality, we have

$$\|v - z\| \leq \sqrt{2 - 2\cos\theta} + \delta \leq \theta + \delta \leq \frac{\pi\epsilon}{10} + \delta$$

for every $v$ in $S'$ (where we have used the bound $\cos\theta \geq 1 - \theta^2/2$). We also have $\|v - z\| \leq 2 + \delta$ for every $v \in V$. Hence

$$\begin{aligned} \sum_{v\in V} \|v - z\| &\leq \sum_{v\in S'} \left(\frac{\pi\epsilon}{10} + \delta\right) + \sum_{v\notin S'} (2 + \delta) \\ &\leq \frac{\pi\epsilon}{10}|V| + 2\frac{\epsilon|V|}{40} + \delta|V| \\ &< \epsilon |V| \end{aligned}$$

which is a contradiction.

Now let $X$ be the set of all $v$'s in $S$, and let $Y$ be the set of all $v$'s outside $S'$. Then $|X| \geq \epsilon |V|/40$ and $|Y| \geq \epsilon n/40$. Also, let

$$\tau = \frac{2\pi}{K} \geq \frac{\pi\epsilon}{20}$$

be the angle of a single sector. Then it is not hard to see that for all $x \in X$ and $y \in Y$,

$$\|x - y\| \geq (1 - \delta)\sqrt{2 - 2\cos\tau} \geq (1 - \delta)\frac{\tau}{\sqrt{2}} \geq \left(1 - \frac{\epsilon}{2}\right)\frac{\pi\epsilon}{20\sqrt{2}} \geq \frac{\epsilon}{20}$$

where we have used the bound $\cos\tau \leq 1 - \tau^2/4$ for all $\tau \in [0, \pi/2]$. □

[13] We did not try to optimize the constants.

Now set $\epsilon := 3\varepsilon^2/5$. Then $\delta = \varepsilon^2/5 < \epsilon/2$ and the condition of Lemma 48 is satisfied. So considering the sets $X, Y$ from the lemma, we have

$$|X|, |Y| \geq \frac{\epsilon |V|}{40} > \frac{\varepsilon^2 N}{2},$$

and also

$$\|x - y\| \geq \frac{\epsilon}{20} \geq \frac{3\varepsilon^2}{100}$$

for all $x \in X$ and $y \in Y$. This means that we can find subsets $X', Y' \subseteq [N]$ such that

(i) $|X'|, |Y'| \geq \varepsilon^2 N/2$ and

(ii) $\left|\alpha_i^2 - \alpha_j^2\right| \geq \frac{3\varepsilon^2}{100N}$ for all $i \in X'$ and $j \in Y'$.

Property (ii) implies that

$$\left|\alpha_i^2 - \alpha_j^2\right|^2 \geq \frac{9\varepsilon^4}{10000N^2} \geq 2c\varepsilon^8 \left(|\alpha_i|^2 + |\alpha_j|^2\right)^2$$

for some suitable constant $c$. Hence every edge $(i, j) \in \mathcal{M}$ with $i \in X'$ and $j \in Y'$ will get added to the set $\mathcal{S}$ (again assuming a suitable $c$).

Property (i), together with a Chernoff bound, implies that with probability $1 - o(1)$ over the choice of matching $\mathcal{M}$, there are at least (say) $\varepsilon^4 N/8$ edges $(i, j) \in \mathcal{M}$ such that $i \in X'$ and $j \in Y'$. Suppose this happens. Then

$$\sum_{(i,j)\in\mathcal{S}} (p_i + p_j) \geq \frac{\varepsilon^4 N}{8} \cdot 2\left(\frac{1-\delta}{N}\right) = \Omega\left(\varepsilon^4\right)$$

as desired. This completes the proof of Theorem 17.

## 9 Appendix: Missing Proofs from Section 3.4

*Proof of Proposition 18.* Let $b = \Pr_{x\in\mathcal{D}_2}[E(x)]$, and note that $|a - b| \leq \epsilon$. Also let $p_x = \Pr_{\mathcal{D}_1}[x]$ and $q_x = \Pr_{\mathcal{D}_2}[x]$. Then

$$\begin{aligned}
\left\|\mathcal{D}_1' - \mathcal{D}_2'\right\| &= \frac{1}{2}\sum_{x:E(x)} \left|\Pr_{\mathcal{D}_1'}[x] - \Pr_{\mathcal{D}_2'}[x]\right| \\
&= \frac{1}{2}\sum_{x:E(x)} \left|\frac{p_x}{a} - \frac{q_x}{b}\right| \\
&\leq \frac{1}{2b}\sum_{x:E(x)} \left(|p_x - q_x| + \left|p_x - \frac{b}{a}p_x\right|\right) \\
&\leq \frac{\epsilon}{2b} + \frac{1}{2}\left|1 - \frac{b}{a}\right| \\
&\leq \frac{\epsilon}{2b} + \frac{\epsilon}{2a} \\
&\leq \frac{\epsilon}{a - \epsilon}.
\end{aligned}$$

□

*Proof of Lemma 19.* Assume for simplicity that $p_x, q_x > 0$ for all $x$ (it is not hard to remove this restriction). Let $\varepsilon_x = p_x' - p_x$ and $\delta_x = q_x' - q_x$. Then by assumption,

$$\sum_{x=1}^{N} \left(|\varepsilon_x| + |\delta_x|\right) \leq 2\mu.$$

Let $\overline{\mathcal{S}}$ be the complement of $\mathcal{S}$. Consider an adversary with a "budget" of $2\mu$, who is trying to perturb $\mathcal{D}$ so as to maximize $\sum_{x \in \overline{\mathcal{S}}} \left(p_x' + q_x'\right)$. Define the "price per pound" of $x$ to be

$$\$_x := \frac{|\varepsilon_x| + |\delta_x|}{p_x' + q_x'}.$$

Intuitively, $\$_x$ is the amount the adversary has to "spend" on perturbing $p_x$ and $q_x$, divided by the amount of probability mass that gets added to $\overline{\mathcal{S}}$ as a result. We will show that $\$_x \geq \left(c - 2c'\right)/4$ for all $x \in \overline{\mathcal{S}}$. This will suffice to prove the lemma, since we then have

$$\sum_{x \in \overline{\mathcal{S}}} \left(p_x' + q_x'\right) = \sum_{x \in \overline{\mathcal{S}}} \frac{|\varepsilon_x| + |\delta_x|}{\$_x} \leq \frac{8\mu}{c - 2c'}.$$

We now lower-bound $\$_x$. If we simply divide through by $p_x q_x$, the condition $2 p_x q_x \geq c \left(p_x + q_x\right)^2$ is equivalent to $p_x/q_x + q_x/p_x \leq \left(2 - 2c\right)/c$. Let $A = \left(2 - 2c\right)/c$; then in particular, we have $p_x \leq A q_x$ and $q_x \leq A p_x$ for all $x$. On the other hand, to get $x \in \overline{\mathcal{S}}$ we need $p_x'/q_x' + q_x'/p_x' > \left(2 - 2c'\right)/c'$, and hence either $p_x'/q_x' > B$ or $q_x'/p_x' > B$ where $B = \left(1 - c'\right)/c'$. Suppose $p_x'/q_x' > B$ without loss of generality. Then

$$p_x + \varepsilon_x > B\left(q_x + \delta_x\right) > B\left(\frac{p_x}{A} + \delta_x\right),$$

which rearranging means

$$\varepsilon_x - B\delta_x > \left(\frac{B}{A} - 1\right) p_x.$$

Likewise

$$B\left(q_x + \delta_x\right) < p_x + \varepsilon_x < A q_x + \varepsilon_x,$$

which rearranging means

$$\varepsilon_x - B\delta_x > \left(B - A\right) q_x > \left(\frac{B}{A} - 1\right) q_x.$$

Combining,

$$\varepsilon_x - B\delta_x > \left(\frac{B}{A} - 1\right) \frac{p_x + q_x}{2}$$

and hence

$$|\varepsilon_x| + |\delta_x| > \frac{1}{B}\left(\varepsilon_x - B\delta_x\right) > \left(\frac{1}{A} - \frac{1}{B}\right) \frac{p_x + q_x}{2}.$$

Therefore

$$\begin{aligned}
\$_x &= \frac{|\varepsilon_x| + |\delta_x|}{p'_x + q'_x} \\
&\geq \frac{|\varepsilon_x| + |\delta_x|}{p_x + q_x + |\varepsilon_x| + |\delta_x|} \\
&> \frac{\frac{1}{2}\left(1/A - 1/B\right)}{1 + \frac{1}{2}\left(1/A - 1/B\right)} \\
&= \frac{c/\left(2-2c\right) - c'/\left(1-c'\right)}{2 + c/\left(2-2c\right) - c'/\left(1-c'\right)} \\
&= \frac{c - 2c' + cc'}{4 - 3c - 6c' + 5cc'} \\
&\geq \frac{c - 2c'}{4}
\end{aligned}$$

as claimed. □

## 10 Appendix: The Generalized Birthday Paradox

The purpose of this appendix is to prove the Birthday Paradox, even in the very general situation where

(1) the distributions over birthdays need not be uniform, and

(2) the distributions need not be the same for every person, but only $\varepsilon$-close in variation distance, and

(3) the distributions need not be independent, but only 4-wise independent.

First we need two lemmas.

**Lemma 49.** *Let $\mathcal{D}_1, \mathcal{D}_2$ be probability distributions over $[n]$ such that $\|\mathcal{D}_1 - \mathcal{D}_2\| \leq \varepsilon$. Then $\Pr_{x \in \mathcal{D}_1, y \in \mathcal{D}_2}[x = y] \geq (1-\varepsilon)^2/n$.*

*Proof.* Let $p_x = \Pr_{\mathcal{D}_1}[x]$ and let $q_x = \Pr_{\mathcal{D}_2}[x]$. Then

$$\begin{aligned}
\Pr_{x \in \mathcal{D}_1, y \in \mathcal{D}_2}[x = y] &= \sum_{x \in [n]} p_x q_x \\
&\geq \sum_{x \in [n]} \min\left(p_x, q_x\right)^2 \\
&\geq \frac{1}{n}\left(\sum_{x \in [n]} \min\left(p_x, q_x\right)\right)^2 \\
&= \frac{1}{n}\left(1 - \varepsilon\right)^2
\end{aligned}$$

where the third line follows from Cauchy-Schwarz. □

**Lemma 50.** *Let $p_1, \ldots, p_K$ be nonnegative reals, and let $r = \sum_{i<j<k} p_i p_j p_k$ and $s = \sum_{i<j} p_i p_j$. Then $r^2 \leq 2s^3$.*

*Proof.* Let $\mathcal{S}$ be the set of 6-tuples $(i, j, k, \ell, m, n)$ such that $i < j$, $k < \ell$, and $m < n$, and let $\mathcal{R}$ be the set of 6-tuples such that $i < j < k$ and $\ell < m < n$. Then

$$s^3 = \sum_{\mathcal{S}} p_i p_j p_k p_\ell p_m p_n$$

while

$$r^2 = \sum_{\mathcal{R}} p_i p_j p_k p_\ell p_m p_n.$$

Now define a mapping from $\mathcal{R}$ to $\mathcal{S}$, by simply swapping $k$ and $\ell$ if $k > \ell$, or swapping $\ell$ and $m$ if $k = \ell$. It is easily checked that this mapping is two-to-one. Hence $r^2 \leq 2s^3$ as claimed. □

We now prove Theorem 20, which we restate for convenience.

**Theorem.** *Let $X_1, \ldots, X_K$ be 4-wise independent random variables over $[n]$, and let $\mathcal{D}_i$ be the marginal 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*

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

*Proof.* Let $Y_{ij}$ be 1 if $X_i = X_j$ and 0 otherwise, and let $Y := \sum_{i<j} Y_{ij}$. By Lemma 49, we have

$$\mathrm{E}\left[Y\right] \geq \binom{K}{2} \frac{(1 - 1/10)^2}{n} \geq 12.$$

The remainder of the proof will involve upper-bounding the second moment $\mathrm{E}\left[Y^2\right]$. Let us write

$$\mathrm{E}\left[Y^2\right] = \sum_{i<j, k<\ell} \mathrm{E}\left[Y_{ij} Y_{k\ell}\right] = \tau_2 + \tau_3 + \tau_4,$$

where $\tau_N$ contains the terms in which $N$ distinct indices appear among $\{i, j, k, l\}$. It is easy to see that

$$\tau_2 = \sum_{i<j} \mathrm{E}\left[Y_{ij}\right] = \mathrm{E}\left[Y\right]$$

and (by 4-wise independence) that

$$\tau_4 = \sum_{i<j, k<l \text{ all distinct}} \mathrm{E}\left[Y_{ij}\right] \mathrm{E}\left[Y_{k\ell}\right] \leq \mathrm{E}\left[Y\right]^2.$$

So the nontrivial part is to upper-bound $\tau_3$. Let $p_{i,x} := \Pr\left[X_i = x\right]$. Also, let

$$r_x := \sum_{i<j<k} p_{i,x} p_{j,x} p_{k,x},$$

$$s_x := \sum_{i<j} p_{i,x} p_{j,x},$$

and notice that $\sum_{x\in[n]} s_x = \mathrm{E}[Y]$. Then

$$\begin{aligned}
\tau_3 &= 6 \sum_{i<j<k} \sum_{x\in[n]} p_{i,x} p_{j,x} p_{k,x} \\
&= 6 \sum_{x\in[n]} r_x \\
&\leq 6 \sum_{x\in[n]} \sqrt{2 s_x^3} \\
&\leq 6\sqrt{2} \sum_{x\in[n]} \left( \frac{1}{10} s_x^2 + 3 s_x \right) \\
&\leq 6\sqrt{2} \left( \frac{1}{10} \left( \sum_{x\in[n]} s_x \right)^2 + 3 \sum_{x\in[n]} s_x \right) \\
&= 6\sqrt{2} \left( \frac{\mathrm{E}[Y]^2}{10} + 3\,\mathrm{E}[Y] \right).
\end{aligned}$$

Here the third line follows from Lemma 50, and the fourth line follows from the basic calculus fact that $s^{3/2} \leq \frac{1}{10} s^2 + 3s$ for all nonnegative $s$.

Hence

$$\begin{aligned}
\Pr[Y=0] &\leq \Pr[|Y - \mathrm{E}[Y]| \geq \mathrm{E}[Y]] \\
&= \Pr\left[ (Y - \mathrm{E}[Y])^2 \geq \mathrm{E}[Y]^2 \right] \\
&\leq \frac{\mathrm{Var}[Y]}{\mathrm{E}[Y]^2} \\
&= \frac{\mathrm{E}\left[Y^2\right] - \mathrm{E}[Y]^2}{\mathrm{E}[Y]^2} \\
&= \frac{\tau_2 + \tau_3 + \tau_4 - \mathrm{E}[Y]^2}{\mathrm{E}[Y]^2} \\
&\leq \frac{\mathrm{E}[Y] + \left( \mathrm{E}[Y]^2/10 + 3\,\mathrm{E}[Y] \right) + \mathrm{E}[Y]^2 - \mathrm{E}[Y]^2}{\mathrm{E}[Y]^2} \\
&= \frac{4}{\mathrm{E}[Y]} + \frac{1}{10} \\
&\leq \frac{1}{2}.
\end{aligned}$$

Finally,

$$\Pr[\exists i, j : X_i = X_j] = 1 - \Pr[Y=0] \geq \frac{1}{2}$$

and we are done. □