Quantum Merlin-Arthur with an Internally Separable Proof

Roozbeh Bassirian
University of Chicago, IL, USA

Bill Fefferman
University of Chicago, IL, USA

Itai Leigh
Tel Aviv University, Israel

Kunal Marwaha
University of Chicago, IL, USA

Pei Wu
Pennsylvania State University, University Park, PA, USA

Abstract

While the role of entanglement in quantum proof systems has been extensively studied, the computational power of unentanglement remains poorly understood. Since entanglement admits many inequivalent multipartite structures, it is natural to ask how more fine-grained structural promises affect computational power. In this work we investigate a mild promise: each proof is internally separable, meaning that after tracing out one register, a designated constant-size subsystem is separable from the rest – even though the overall proof may still be entangled across every bipartition. We prove a qualitative jump from one proof to two: with one internally separable proof, the resulting class is contained in EXP (even allowing an inverse-exponential completeness–soundness gap), whereas with two unentangled internally separable proofs, the class equals NEXP at constant gap. Notably, in the NEXP construction, the second proof is used solely to implement a SWAP-based purity test.

2012 ACM Subject Classification Theory of computation → Quantum complexity theory; Theory of computation → Interactive proof systems; Theory of computation → Complexity classes; Theory of computation → Problems, reductions and completeness; Theory of computation → Quantum complexity theory

Keywords and phrases entanglement structures, unentanglement, quantum complexity, QMA(2), NEXP

Digital Object Identifier 10.4230/LIPIcs.TQC.2026.5

Related Version *Full Version*: https://arxiv.org/abs/2410.19152

Funding This work was done in part while a subset of the authors were visiting the Simons Institute for the Theory of Computing, supported by NSF QLCI Grant No. 2016245. BF, RB, and KM acknowledge support from AFOSR (FA9550-21-1-0008). This material is based upon work partially supported by the National Science Foundation under Grant CCF-2044923 (CAREER), by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers (Q-NEXT) and by the DOE QuantISED grant DE-SC0020360.
Any views, opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, European Union, or the European Research Council Executive Agency. Neither the European Union nor any granting authority can be held responsible for them.
Itai Leigh: co-funded by the European Union (ERC, EACTP, 101142020), the Israel Science Foundation (grant number 514/20) and the Len Blavatnik and the Blavatnik Family Foundation.
Kunal Marwaha: acknowledges support from the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1746045.

Acknowledgements KM thanks Tyler Chen, Christopher Musco, and Apoorv Vikram Singh for helpful discussions on convex optimization. RB, KM, and IL thank the Institute for Advanced Study for organizing the summer school where they met and began collaborating. IL and PW thank Thomas Vidick for suggesting them to talk.

1 Overview

Entanglement is a fundamental property of quantum mechanics. One approach to understanding this phenomenon is through the lens of computational complexity, probing the power of quantum proof systems with and without entanglement. A canonical example is the relationship between QMA and QMA(2). Here QMA is the quantum analogue of NP (see [30] for a survey): a quantum polynomial-time verifier (Arthur) receives a quantum state (the proof) from an all-powerful prover (Merlin) and decides whether to accept. The class QMA(2) restricts Merlin's proof to be a product across a fixed bipartition – equivalently, the two halves are sent by two unentangled Merlins. Understanding whether this restriction gives strictly more power is one of the central open problems in quantum complexity theory, unresolved for two decades (e.g. [49, 2, 38]). Despite substantial evidence that QMA(2) can outperform QMA in natural settings (e.g. [15, 2, 22, 58]), our unconditional knowledge in the standard constant-gap setting remains limited to the trivial inclusions

$$\mathsf{QMA} \subseteq \mathsf{QMA}(2) \subseteq \mathsf{NEXP}$$

As posed in [2]: what is the power of unentanglement?

One of the unusual aspects of this landscape is that QMA(2) appears very sensitive to the structure of the proof system. Small changes to the model – extra interaction, very small gaps, or additional promises on the witness – can raise the power all the way to NEXP [58, 45, 10, 32]. This sensitivity complicates attempts to prove meaningful upper bounds for QMA(2). On the other hand, multipartite entanglement admits far richer structure than the binary distinction "entangled vs. unentangled," motivating a finer-grained study of the witness's entanglement structure. In this work, we take a first step exploring whether fine-grained entanglement promises can lead to a robust and analyzable refinement of QMA(2), or whether even very mild structural constraints on entanglement already trigger the same dramatic complexity jumps.

1.1 Internal separability: a fine-grained entanglement promise

We study the power of unentanglement through a natural multipartite promise that constrains only a *marginal* of the witness. We restrict Merlin to quantum states we call internally separable: a proof is a tripartite state ρ_{ABC} where C has only constant qubits and ρ_{BC} is separable. This is a mild constraint – ρ may remain entangled across every bipartition – but it fixes a separable marginal. Equivalently, these witnesses can be viewed as (partial) purifications of separable states. Importantly, internal separability is a constraint within a single proof. Our results show that even this very mild structural promise reveals unentanglement as powerful.

1.2 A qualitative one-versus-two jump, driven only by purity

Our main theorem establishes a sharp one-versus-two prover separation for internally-separable proofs. With one such proof, the resulting proof system is simulable in EXP. With two unentangled such proofs, the system becomes as powerful as NEXP. Thus, the "power of unentanglement" in this setting is as large as the gap between EXP and NEXP.

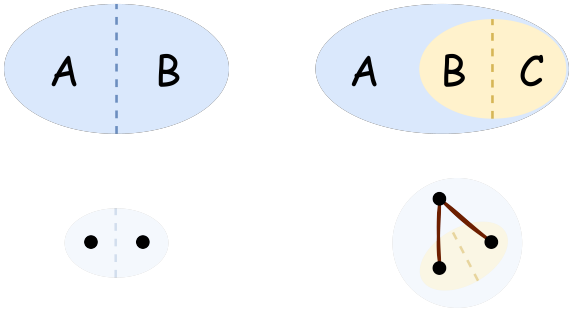

Figure 1 Cartoons of a separable state and a state with a separable subsystem. The left side depicts a separable state, where parts A and B share no entanglement. The right side depicts a state where *after* tracing out part A, parts B and C share no entanglement; this state is in general entangled across every bipartition. The lower cartoons each display a graph where every vertex is a part (A, B, or C), and an edge between two parts X, Y allows bipartite entanglement in the reduced state ρ_{XY}. On the right side, part A is entangled with both part B and part C.

Beyond the separation itself, the proof reveals an unexpected mechanism. In our construction, the role of the two unentangled proofs is minimal: they are used solely to implement a purity test. This stands in contrast with the usual intuition around QMA(2), where one typically relies on separability in a more robust way – most notably via the product test [38]. Conceptually, our results isolate a striking source of power: certifying that a state is pure (as opposed to mixed) can already suffice to unlock NEXP in the presence of even limited internal separability.

More broadly, our work highlights a qualitative distinction between pure and mixed witnesses in multi-prover quantum proof systems, and it points to a concrete route by which structural promises about internal entanglement might drive complexity-theoretic jumps. In particular, it resonates with recent efforts aimed at understanding whether QMA(2) itself could reach NEXP [45, 10]. We hope this perspective – treating entanglement structure as a first-class computational resource – encourages further investigation into how subtle promises about witnesses can have outsized consequences for computational power.

2 Results

Consider the complexity class $\mathsf{QMA}_{\mathsf{IS}}$, a variant of QMA where the quantum proof must have a fixed, separable *subsystem*: that is, after tracing out all qubits beyond the subsystem, a predetermined set of $O(1)$ qubits are *not* entangled with the rest of the subsystem (see formal definition in Appendix A). Unlike in QMA(2), this guarantee refers to *internal separability*: the two parts of the subsystem can be entangled, but only if the entanglement disappears when the subsystem is isolated. See Figure 1 for a visual comparison. We restrict one part of the subsystem to $O(1)$ qubits to zero in on the effect of *internal* entanglement structure; without the restriction, this class trivially contains QMA(2).

We first show that $\mathsf{QMA}_{\mathsf{IS}}$, even with inverse exponentially small gap, is contained in EXP:

▶ **Theorem 1** (Upper bound). *For any polynomial p and $1 > c > s > 0$ such that $c - s > \frac{1}{2^{p(n)}}$, we have $\mathsf{QMA}_{\mathsf{IS}}^{c,s} \subseteq \mathsf{EXP}$, where c, s are the completeness and soundness parameters, respectively.*[1]

[1] We formally define $\mathsf{QMA}_{\mathsf{IS}}$ in Appendix A. In a $\mathsf{QMA}_{\mathsf{IS}}^{c,s}$ protocol, YES instances are accepted with probability at least c, and NO instances are accepted with probability at most s. $\mathsf{QMA}_{\mathsf{IS}}$ includes every $\mathsf{QMA}_{\mathsf{IS}}^{c,s}$ with a constant gap $c - s = \Omega(1)$.

We also define $\mathsf{QMA}_{\mathsf{IS}}(2)$, where the verifier receives *two* unentangled proofs, each with the above guarantee. In contrast to Theorem 1, we find a constant-sized gap where $\mathsf{QMA}_{\mathsf{IS}}(2)$ equals NEXP.

▶ **Theorem 2** (Lower bound). *There exist constants* $1 > c > s > 0$ *where* $\mathsf{QMA}_{\mathsf{IS}}^{c,s}(2) = \mathsf{NEXP}$.

We note two immediate consequences. First, there exists a quantum verifier with a specific guarantee on entanglement structure that is *unconditionally* (and *exponentially*) more powerful than a deterministic classical verifier. This follows as a consequence of the non-deterministic time hierarchy theorem:

▶ **Corollary 3.** $\mathsf{NP} \subsetneq \mathsf{QMA}_{\mathsf{IS}}(2)$.

Second, $\mathsf{QMA}_{\mathsf{IS}}$ and $\mathsf{QMA}_{\mathsf{IS}}(2)$ cannot be equal unless $\mathsf{EXP} = \mathsf{NEXP}$. In other words, the power of unentanglement in this setting is as strong as the difference between EXP and NEXP:

▶ **Corollary 4.** *If* $\mathsf{EXP} \neq \mathsf{NEXP}$, *then* $\mathsf{QMA}_{\mathsf{IS}} \neq \mathsf{QMA}_{\mathsf{IS}}(2)$.

We summarize these results in Table 1.

Table 1 Categorizing quantum verification protocols by the guaranteed entanglement structures of the proof.

internal structure \ *external structure*	GENERAL	SEPARABLE BIPARTITION
GENERAL	QMA	QMA(2)
SEPARABLE SUBSYSTEM	$\mathsf{QMA}_{\mathsf{IS}} \subseteq \mathsf{EXP}$	$\mathsf{QMA}_{\mathsf{IS}}(2) = \mathsf{NEXP}$

3 Themes and implications

We highlight three conceptual takeaways: (i) internal entanglement promises are a natural complexity-theoretic probe of fine-grained separability structures; (ii) in our setting, unentanglement is used only to certify purity, exposing a sharp pure-vs-mixed distinction; (iii) this sheds some light on a gap-amplification route towards showing lower bounds of QMA(2) that avoids known barriers for nearby models.

3.1 A complexity-theoretic probe of multipartite entanglement structure

Despite much attention given to QMA(2)–the study of unentanglement, (un)entanglement can be much richer than simply across a bipartition. For example, the W and GHZ states, which are $\frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle)$ and $\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$ respectively, are entangled in *incomparable* ways: both states are entangled across every bipartition, yet the states are not equivalent under LOCC [27]. The same paper observes that tracing out a qubit from the W state gives an entangled state, but the same action on the GHZ state destroys all entanglement. These examples suggest that "entanglement vs. no entanglement across a cut" is only a coarse proxy for the structure that may matter in proof systems.

Our class $\mathsf{QMA}_{\mathsf{IS}}$ enforces quantum proofs with a specific multipartite entanglement structure, which we call internal separability. To our knowledge, this notion first appeared in [19], where it was called *partial semi-separability*. It is natural to study the power of QMA with *other* types of multipartite entanglement structure. As a first step, we consider

the set of *multiseparable* states [60]: tripartite states where after tracing out *any* register, the reduced density matrix is separable. We show that a quantum proof system over pure, multiseparable states can also decide NEXP:

▶ **Corollary 5.** *There exist constants* $1 > c > s > 0$ *where* $\mathsf{QMA}^{c,s}$ *restricted to pure, multiseparable quantum proofs can decide* NEXP*.*

A useful perspective is that many known NEXP-power phenomena in restricted-witness models arise because the promise enables the verifier to implement some form of *classicality/sparsity detection* primitive – e.g., distinguishing structured "nearly classical" superpositions from genuinely diffuse ones. In our case, the internal separability promise plus purity plays exactly this role: it enables quasirigidity enforcement in the NEXP protocol. An interesting question is whether there is a robust way to characterize which multipartite entanglement promises yield NEXP-power in terms of the verification primitives they enable.

3.2 QMA(2) and purity testing

In our work, $\mathsf{QMA}_{|\mathsf{S}\rangle}(2)$ is no more powerful than $\mathsf{QMA}_{|\mathsf{S}\rangle}$ with a purity test. This resonates with a broader theme in the previous works [54, 38, 9, 64] where they use the fact that a protocol over separable states can simulate a protocol over states with a purity constraint by the purity test. This raises the question: What is the relative power between purity and separability? For example, one QMA(2)-complete candidate was the *pure state marginal problem* [54]: decide if there exists a global pure state consistent with the input local density matrices. This problem is known to be in QMA(2), while the *mixed* variant of this problem is QMA-complete [52].

After our work was put online, [47] showed that the pure state marginal problem is in PSPACE, even with an inverse exponential promise gap. This suggests that the pure state marginal problem is in fact *not* QMA(2)-hard, as QMA(2) with an inverse-exponential completeness-soundness gap is equal to NEXP [58]. Moreover, they propose one systematic way to understand the power of purity by abstracting out the complexity class definition based on the pure version of the super verifier characterization of QMA [5]. Many of the questions remain unclear for their model.

The quantum proof systems are associated with their classical optimization counterparts. Therefore, we also consider the optimization problem over *purifications* of separable states that is enough to decide NP, and even hard to approximate within a constant error:

▶ **Corollary 6.** *Given* $M \in \mathbb{R}^{n\times n}$*, consider the following optimization problem over* n*-dimensional pure, internally separable states (denoted* $\mathcal{W}_{|\mathsf{S}\rangle}$*):*

$$\max_{\rho\in\mathcal{W}_{|\mathsf{S}\rangle}} \mathrm{Tr}[M\rho] \qquad (*)$$

Then there is a constant $0 < \alpha < 1$ *for which it is* NP*-hard to compute an* α*-approximation of* $(*)$*.*

By contrast, optimizing over separable states (of $\mathsf{poly}(n)$ dimension) is known to be NP-hard, but only with a gap inverse polynomial of the system size [35, 2, 29, 15, 58, 28]. If the gap in these problems could be made constant, then $\mathsf{QMA}(2) = \mathsf{NEXP}$.[2]

More generally, assuming mild conditions on a set of quantum states $\mathcal{W}$, we show how a QMA(2) protocol with both proofs restricted to $\mathcal{W}$ can implement a QMA protocol with a proof restricted to *pure states* in $\mathcal{W}$. We use this as a recipe to generate other proof systems

[2] One must also assume that the problems can be made succinct.

where one quantum proof is provably weaker than two *unentangled* quantum proofs, up to standard complexity-theoretic assumptions. We give two examples: the pure quantum proofs in QMA^+, and the *proper states* of [2, 12]. Each set represents the pure states of some larger set $\mathcal{W}$. Then, $\mathsf{QMA}_{\mathcal{W}}$ can be decided with a semidefinite program, but $\mathsf{QMA}_{\mathcal{W}}(2)$ can restrict to the more capable pure states. This technique is likely to generate other examples that demonstrate the power of unentanglement.

3.3 A revived approach towards QMA(2)

Corollary 4 suggests a potentially cleaner gap-amplification route towards understanding the power of $\mathsf{QMA}(2)$. For a constant r, denote by $\mathsf{QMA}_{\mathsf{IS}}^{c,s}[r](k)$ the class of problems in $\mathsf{QMA}_{\mathsf{IS}}^{c,s}(k)$ with at most r dimensions for the witnesses' bounded part of the separable subsystem. Since r is constant, we have that for any constant k, $\mathsf{QMA}_{\mathsf{IS}}[r](k)$ at large enough completeness-soundness gap is equal to $\mathsf{QMA}(k)$:

▷ Claim 7. For any constants k and r, there exist constants $1 > c > s > 0$ so that $\mathsf{QMA}_{\mathsf{IS}}^{c,s}[r](k) = \mathsf{QMA}(k)$.

By Claim 7 and Theorem 2,[3] gap amplification of $\mathsf{QMA}_{\mathsf{IS}}[r](2)$ would imply $\mathsf{QMA}(2) = \mathsf{NEXP}$:

▶ **Corollary 8.** *If for some r $\mathsf{QMA}_{\mathsf{IS}}[r](2)$ exhibits gap amplification, then $\mathsf{QMA}(2) = \mathsf{NEXP}$.*

Corollary 8 is analogous to an approach suggested in [45]. That work introduced the class $\mathsf{QMA}^+(k)$, and showed that $\mathsf{QMA}^+(2)$ at some completeness-soundness gap is equal to NEXP, while at a larger gap, is equal to $\mathsf{QMA}(2)$. Therefore, if $\mathsf{QMA}^+(2)$ admits gap amplification, then $\mathsf{QMA}(2) = \mathsf{NEXP}$. However, [10] showed that the same is true for QMA^+; as a consequence, QMA^+ *cannot* have gap amplification unless $\mathsf{QMA} = \mathsf{NEXP}$. Conceptually, this limits the techniques available to prove gap amplification of $\mathsf{QMA}^+(2)$ in order to show $\mathsf{QMA}(2) = \mathsf{NEXP}$. By contrast, gap amplification of $\mathsf{QMA}_{\mathsf{IS}}(2)$ can be plausibly found by studying gap amplification of $\mathsf{QMA}_{\mathsf{IS}}$, since the latter implies no such complexity collapse. This raises the possibility of using Corollary 8 to fully characterize $\mathsf{QMA}(2)$. However, any such gap amplification must be sensitive to *purity*, as $\mathsf{QMA}_{|\mathsf{IS}\rangle}$ suffers the same issue as QMA^+. We reinterpret the barrier of [10] by constructing a mixed-state variant of QMA^+ that also avoids the complexity collapse, as broadly explained in the previous subsection.

4 Techniques

4.1 Upper bound

To show Theorem 1, we rely on a well-known duality of membership testing of a convex set and optimizing a convex function over this set. We use this duality to reduce any $\mathsf{QMA}_{\mathsf{IS}}$ problem to membership testing of the set of internally separable quantum states $\mathcal{W}_{\mathsf{IS}}$. In order to use tools from convex optimization, we represent quantum states as vectors in a generalized Bloch sphere. We use a non-ellipsoidal reduction that permits an error analysis matching our setup [53, 29].

We then show how to test membership of $\mathcal{W}_{\mathsf{IS}}$ in EXP. It is essential that only $O(1)$ qubits are unentangled from the rest of the subsystem. Our algorithm tests membership of two convex sets whose intersection forms $\mathcal{W}_{\mathsf{IS}}$; note that this requires testers with inverse

[3] The theorem also holds for $\mathsf{QMA}_{\mathsf{IS}}^{c,s}[r](2)$ for some fixed constant r.

exponential precision. The first set represents all Bloch vectors of quantum states, which can be tested using exponentially many explicit conditions (e.g. [48]). The second set represents all vectors whose reduced density matrix on the subsystem is separable. This can be tested by solving the Doherty-Parrilo-Spedalieri SDP [25, 26, 56] which decides if a state has a *symmetric extension*. A careful error analysis shows that separability of a bipartite state ρ with local dimensions M, N can be tested to error δ in time $\tilde{O}((2/\delta)^{9N}\mathsf{poly}(M,N))$ [44]. When N is a fixed constant, this SDP runs in exponential time.

4.2 Lower bound

To prove Theorem 2, we consider $\mathsf{QMA}_{|\mathsf{IS}\rangle}$, a variant of $\mathsf{QMA}_{\mathsf{IS}}$ where the proof must also be a *pure state*. Note that unlike QMA and $\mathsf{QMA}(2)$, restricting $\mathsf{QMA}_{\mathsf{IS}}$ to pure states can change the power of the complexity class, as an internally separable ensemble of pure states may not decompose into an *ensemble* of internally separable pure states.[4] We then develop a $\mathsf{QMA}_{|\mathsf{IS}\rangle}$ protocol for an NEXP-complete problem. This protocol builds on the recent proof of $\mathsf{QMA}^+ = \mathsf{NEXP}$ [15, 45, 10], where QMA^+ is restricted to pure quantum proofs with real nonnegative amplitudes.

As in [45, 10], we consider a succinct constraint satisfaction problem (CSP) with constant promise gap, which is NEXP-hard by the PCP theorem [7, 6, 40]. In the QMA^+ protocol, one enforces a *rigidity* property of the quantum proof: the state must be close to $\frac{1}{\sqrt{R}}\sum_{j\in[R]}|j\rangle|a_j\rangle$. It is easy to ensure that all constraints $j \in [R]$ are included in the state ("density"), but the positive-amplitude guarantee is used to ensure there is at most one assignment a_j per constraint ("quasirigidity"). We find a way to instead enforce quasirigidity using both internal separability *and* purity.

▶ **Lemma 9** (informal). *There is a test A and a fixed quantum circuit C such that for any internally separable $|\psi\rangle$, if $A(|\psi\rangle) \geq 1-\epsilon$, then $C|\psi\rangle$ has $1-O(\epsilon)$ squared overlap with a* quasirigid *state.*

In our $\mathsf{QMA}_{|\mathsf{IS}\rangle}$ protocol, the verifier enforces tripartite proofs of the specific form $\sum_{j,a} z_{ja}|j\rangle|a,j\rangle|a\rangle$. This can be done by measuring all registers and accepting iff we see two copies of the same constraint j, and two copies of the same assignment a. We then observe that having multiple assignments per constraint adds entanglement to the second and third registers, even after tracing out the first register. Thus, if the state is *internally separable*, it has high overlap with a state with one assignment a per constraint j. Notably, this test is inspired by recasting the QMA^+ protocol in the language of "superposition detection" and non-collapsing measurement [11, 4], an unphysical operation related to a problem of Aaronson [1].

We then prove that $\mathsf{QMA}_{|\mathsf{IS}\rangle} \subseteq \mathsf{QMA}_{\mathsf{IS}}(2)$. Without loss of generality, the best proof Merlin can send in a $\mathsf{QMA}_{\mathsf{IS}}(2)$ protocol is a product state $\rho_1 \otimes \rho_2$, since separable states form the convex hull of the set of product states. The SWAP test accepts a product state with probability $\frac{1}{2}(1+\mathrm{Tr}[\rho_1\rho_2])$: this is 1 only if ρ_1, ρ_2 are *pure*. The $\mathsf{QMA}_{\mathsf{IS}}(2)$ simulator thus applies a SWAP test with some probability, and otherwise applies the $\mathsf{QMA}_{|\mathsf{IS}\rangle}$ circuit.

To make this formal, we first generalize the above statement. Given a set of quantum states $\mathcal{W}$, let $\mathsf{QMA}_{\mathcal{W}}$ be the variant of QMA where the proof must be in $\mathcal{W}$. Let $\mathrm{rank1}(\mathcal{W}) \subseteq \mathcal{W}$ be the pure states contained in $\mathcal{W}$. Then provided $\mathcal{W}$ satisfies a mild *continuity condition*, we have that $\mathsf{QMA}_{\mathrm{rank1}(\mathcal{W})} \subseteq \mathsf{QMA}_{\mathcal{W}}(2)$.

[4] For example, the same is true in QMA^+: this is the difference between doubly non-negative matrices and completely positive matrices.

▶ **Lemma 10** (informal). *Consider a protocol V in $\mathsf{QMA}_{\mathrm{rank1}(\mathcal{W})}$. If $\mathcal{W}$ has the continuity condition, then $\mathsf{QMA}_{\mathcal{W}}(2)$ can simulate V.*

Intuitively, the continuity condition requires that every density operator in $\mathcal{W}$ that is "almost pure" is close to a pure state in $\mathcal{W}$; i.e. in $\mathrm{rank1}(\mathcal{W})$. The condition ensures that when the SWAP test accepts with high probability, both ρ_1 and ρ_2 are close to a state in $\mathrm{rank1}(\mathcal{W})$.

Ideally, we would show that $\mathcal{W}_{\mathsf{IS}}$ satisfies the continuity condition, but a proof of this seems difficult. Nonetheless, we find a large subset $\mathcal{W}_{\mathsf{IS}^\star} \subseteq \mathcal{W}_{\mathsf{IS}}$ that has the continuity condition, such that the NEXP protocol is still contained in $\mathsf{QMA}_{\mathrm{rank1}(\mathcal{W}_{\mathsf{IS}^\star})}$. By the above lemma, NEXP is contained in $\mathsf{QMA}_{\mathcal{W}_{\mathsf{IS}^\star}}(2)$. To complete the proof, we show how to simulate any protocol in $\mathsf{QMA}_{\mathcal{W}_{\mathsf{IS}^\star}}(2)$ using a protocol in $\mathsf{QMA}_{\mathcal{W}_{\mathsf{IS}}}(2)$.

5 Related work

5.1 The complexity class QMA(2)

One sign that QMA(2) is more powerful than QMA is that QMA with a $O(\log n)$-qubit proof is equal to BQP [55]. By contrast, QMA(2) with two unentangled $O(\log n)$-qubit proofs can decide NP, although with inverse polynomially small completeness-soundness gap [2, 12, 15, 22, 23, 28, 58]. Improving the completeness-soundness gap to a constant would likely imply $\mathsf{QMA}(2) = \mathsf{NEXP}$; see [45, 10, 46] for a recent attempt using quantum proofs with non-negative amplitudes (QMA^+). In general, one only needs two unentangled proofs since $\mathsf{QMA}(2) = \mathsf{QMA}(k)$ [38]; the proof of this result uses an iterated SWAP test. There are few complete problems known for QMA(2) [21, 41, 36]. One proposal is the *pure state marginal problem*, which is known to be QMA-hard and in QMA(2) [52, 54, 53, 63, 65, 64]. Recently, [47] proved that this problem is in PSPACE, and gave evidence that it is *not* QMA(2)-hard.

5.2 Convex optimization and quantum entanglement

Some quantum papers build a Turing reduction using the duality between convex optimization and membership testing of a convex set, for example NP-hardness of separability testing [35, 44, 29], and QMA-hardness of the mixed state marginal problem [52, 54, 53, 63]. This reduction can be implemented in several ways, including the ellipsoid method [66, 34], random walks [14], and a perceptron-like algorithm [53]. One can use a hierarchy of semidefinite programs to detect entanglement [25, 26, 56, 39] or optimize over separable states [16, 17, 9]. In some cases, carefully chosen ϵ-nets provide similar performance [59, 8, 18].

5.3 Measuring quantum entanglement

There are numerous ways to quantify entanglement; see the surveys [67, 37, 43, 42]. One approach considers the set of reachable states under transformations consisting only of local operations and classical communication (LOCC) [27, 13, 62]. Other works classify multipartite entanglement by whether a *subsystem* is separable [60, 19]. Permutation-symmetric approaches include the Positive Partial Transpose (PPT) criterion [57] and quantum de Finetti theorems [20, 24]. Mysteriously, both approaches identify entangled states that "look" separable under LOCC transformations [61, 16, 17, 50, 51]. There exist ensembles of low-entanglement quantum states that are computationally indistinguishable from maximally entangled quantum states [3].

6 Outlook

There are several ways to further the ideas within this work. In our view, the most promising future direction is to study other quantum proof systems with other multipartite entanglement structures. For example, what is the power of $\mathsf{QMA}_{\mathcal{W}}$ when the states in $\mathcal{W}$ are the Absolutely Maximally Entangled states (see review in [31])?

Another direction is to probe the relationship of $\mathsf{QMA}(2)$ and purity testing. The proof of Theorem 2 uses two quantum proofs only to implement a purity test. Is there a set of proofs $\mathcal{W}$ such that $\mathsf{QMA}_{\mathcal{W}}(2)$ is strictly more powerful than $\mathsf{QMA}_{\mathrm{rank1}(\mathcal{W})}$? If so, there is a setting where *unentanglement* is useful for something beyond purity testing. A related technical question is whether the *continuity condition* is necessary to show $\mathsf{QMA}_{\mathrm{rank1}(\mathcal{W})} \subseteq \mathsf{QMA}_{\mathcal{W}}(2)$. Is there a simple characterization of sets which do *not* satisfy the continuity condition?

This work also reignites the possibility of showing the power of $\mathsf{QMA}(2)$ by proving a gap amplification result (Corollary 8). As a starting point, does the complexity class $\mathsf{QMA}_{\mathsf{IS}}$ exhibit gap amplification? (Even a small amount of gap amplification implies the existence of new complete problems for $\mathsf{QMA}_{\mathsf{IS}}$.) One can ask similar questions about the additional quantum proof systems we discuss, see Section 3.1. Unlike with QMA^{+} [10], each of these complexity classes demonstrates that unentanglement can be as powerful as the difference between EXP vs. NEXP.

References

1 Scott Aaronson. Quantum miscellany — shtetl-optimized, 2023. URL: https://scottaaronson.blog/?p=7516.

2 Scott Aaronson, Salman Beigi, Andrew Drucker, Bill Fefferman, and Peter Shor. The power of unentanglement. *Theory of Computing*, 5(1):1–42, 2009. doi:10.4086/toc.2009.v005a001.

3 Scott Aaronson, Adam Bouland, Bill Fefferman, Soumik Ghosh, Umesh Vazirani, Chenyi Zhang, and Zixin Zhou. Quantum Pseudoentanglement. In *15th Innovations in Theoretical Computer Science Conference (ITCS 2024)*, 2024. doi:10.4230/LIPIcs.ITCS.2024.2.

4 Scott Aaronson, Sabee Grewal, Vishnu Iyer, Simon C. Marshall, and Ronak Ramachandran. Pdqma = dqma = nexp: Qma with hidden variables and non-collapsing measurements, 2024. doi:10.48550/arXiv.2403.02543.

5 Dorit Aharonov and Oded Regev. A lattice problem in quantum np, 2003. arXiv:quant-ph/0307220.

6 Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, 1998. doi:10.1145/278298.278306.

7 Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs; A new characterization of NP. In *33rd Annual Symposium on Foundations of Computer Science*, 1992. doi:10.1109/SFCS.1992.267824.

8 Boaz Barak, Jonathan A. Kelner, and David Steurer. Rounding sum-of-squares relaxations. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, 2014. doi:10.1145/2591796.2591886.

9 Boaz Barak, Pravesh K. Kothari, and David Steurer. Quantum entanglement, sum of squares, and the log rank conjecture. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, 2017. doi:10.1145/3055399.3055488.

10 Roozbeh Bassirian, Bill Fefferman, and Kunal Marwaha. Quantum Merlin-Arthur and proofs without relative phase. In *15th Innovations in Theoretical Computer Science Conference (ITCS 2024)*, 2023. doi:10.4230/LIPIcs.ITCS.2024.9.

11 Roozbeh Bassirian and Kunal Marwaha. Superposition detection and qma with non-collapsing measurements, 2024. doi:10.48550/arXiv.2403.02532.

12 Salman Beigi. Np vs qma_log(2), 2008. arXiv:0810.5109.

13 Charles H. Bennett, Sandu Popescu, Daniel Rohrlich, John A. Smolin, and Ashish V. Thapliyal. Exact and asymptotic measures of multipartite pure-state entanglement. *Phys. Rev. A*, 63, December 2000. doi:10.1103/PhysRevA.63.012307.

14 Dimitris Bertsimas and Santosh Vempala. Solving convex programs by random walks. *J. ACM*, 51(4):540–556, June 2004. doi:10.1145/1008731.1008733.

15 Hugue Blier and Alain Tapp. A quantum characterization of np. *Comput. Complex.*, September 2012. doi:10.1007/s00037-011-0016-2.

16 Fernando G.S.L. Brandão, Matthias Christandl, and Jon Yard. A quasipolynomial-time algorithm for the quantum separability problem. In *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing*, June 2011. doi:10.1145/1993636.1993683.

17 Fernando G.S.L. Brandão and Aram W. Harrow. Quantum de finetti theorems under local measurements with applications. In *Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing*, 2013. doi:10.1145/2488608.2488718.

18 Fernando G.S.L. Brandão and Aram W. Harrow. Estimating operator norms using covering nets, 2015. arXiv:1509.05065.

19 Gilles Brassard and Tal Mor. Multi-particle entanglement via two-party entanglement. *Journal of Physics A: Mathematical and General*, 34(35), August 2001. doi:10.1088/0305-4470/34/35/306.

20 Carlton M. Caves, Christopher A. Fuchs, and Rüdiger Schack. Unknown quantum states: The quantum de finetti representation. *Journal of Mathematical Physics*, 43(9), September 2002. doi:10.1063/1.1494475.

21 Andre Chailloux and Or Sattath. The complexity of the separable hamiltonian problem. In *2012 IEEE 27th Conference on Computational Complexity*. IEEE, June 2012. doi:10.1109/ccc.2012.42.

22 Jing Chen and Andrew Drucker. Short multi-prover quantum proofs for sat without entangled measurements, 2010. arXiv:1011.0716.

23 Alessandro Chiesa and Michael A. Forbes. Improved soundness for qma with multiple provers. *Chicago Journal of Theoretical Computer Science*, January 2013. arXiv:1108.2098.

24 Matthias Christandl, Robert König, Graeme Mitchison, and Renato Renner. One-and-a-half quantum de finetti theorems. *Communications in Mathematical Physics*, 273(2):473–498, March 2007. doi:10.1007/s00220-007-0189-3.

25 Andrew C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. Distinguishing separable and entangled states. *Phys. Rev. Lett.*, 88, April 2002. doi:10.1103/PhysRevLett.88.187904.

26 Andrew C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. Complete family of separability criteria. *Physical Review A*, 69(2), February 2004. doi:10.1103/physreva.69.022308.

27 W. Dür, G. Vidal, and J. I. Cirac. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A*, 62, November 2000. doi:10.1103/PhysRevA.62.062314.

28 François Le Gall, Shota Nakagawa, and Harumichi Nishimura. On QMA protocols with two short quantum proofs. *Quantum Inf. Comput.*, 12(7-8):589–600, 2012. doi:10.26421/QIC12.7-8-4.

29 Sevag Gharibian. Strong np-hardness of the quantum separability problem. *Quantum Info. Comput.*, 10(3):343–360, March 2010. doi:10.5555/2011350.2011361.

30 Sevag Gharibian. Guest column: The 7 faces of quantum NP. *SIGACT News*, 54(4):54–91, 2023. doi:10.1145/3639528.3639535.

31 Dardo Goyeneche, Daniel Alsina, José I Latorre, Arnau Riera, and Karol Życzkowski. Absolutely maximally entangled states, combinatorial designs, and multiunitary matrices. *Physical Review A*, 92(3), 2015. arXiv:1506.08857.

32 Sabee Grewal and William Kretschmer. Unentanglement and post-measurement branching in quantum interactive proofs. *CoRR*, abs/2509.15319, 2025. doi:10.48550/arXiv.2509.15319.

33 Alex B. Grilo, Iordanis Kerenidis, and Jamie Sikora. Qma with subset state witnesses. In *International Symposium on Mathematical Foundations of Computer Science*, 2015. arXiv:1410.2882.

34 Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2. Springer Science & Business Media, 1988. doi:10.1007/978-3-642-97881-4.

35 Leonid Gurvits. Classical deterministic complexity of edmonds' problem and quantum entanglement. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, 2003. doi:10.1145/780542.780545.

36 Gus Gutoski, Patrick Hayden, Kevin Milner, and Mark M. Wilde. Quantum interactive proofs and the complexity of separability testing. *Theory of Computing*, 11(1):59–103, 2015. doi:10.4086/toc.2015.v011a003.

37 Otfried Gühne and Géza Tóth. Entanglement detection. *Physics Reports*, 474(1–6):1–75, April 2009. doi:10.1016/j.physrep.2009.02.004.

38 Aram W. Harrow and Ashley Montanaro. Testing product states, quantum merlin-arthur games and tensor optimization. *Journal of the ACM (JACM)*, 60(1):1–43, 2013. doi:10.1145/2432622.2432625.

39 Aram W. Harrow, Anand Natarajan, and Xiaodi Wu. An improved semidefinite programming hierarchy for testing entanglement. *Communications in Mathematical Physics*, 352(3):881–904, March 2017. doi:10.1007/s00220-017-2859-0.

40 Prahladh Harsha. *Robust PCPs of proximity and shorter PCPs*. PhD thesis, Massachusetts Institute of Technology, 2004. URL: https://dspace.mit.edu/bitstream/handle/1721.1/26720/59552830-MIT.pdf.

41 Patrick Hayden, Kevin Milner, and Mark M. Wilde. Two-message quantum interactive proofs and the quantum separability problem. In *2013 IEEE Conference on Computational Complexity*, June 2013. doi:10.1109/ccc.2013.24.

42 Pawel Horodecki, Lukasz Rudnicki, and Karol Życzkowski. Multipartite entanglement, 2024. arXiv:2409.04566.

43 Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, June 2009. doi:10.1103/RevModPhys.81.865.

44 Lawrence M. Ioannou. Computational complexity of the quantum separability problem. *Quantum Info. Comput.*, 7(4):335–370, May 2007. doi:10.26421/QIC7.4-5.

45 Fernando Granha Jeronimo and Pei Wu. The power of unentangled quantum proofs with non-negative amplitudes. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, 2023. doi:10.1145/3564246.3585248.

46 Fernando Granha Jeronimo and Pei Wu. Dimension Independent Disentanglers from Unentanglement and Applications. In *39th Computational Complexity Conference (CCC 2024)*, 2024. doi:10.4230/LIPIcs.CCC.2024.26.

47 Jonas Kamminga and Dorian Rudolph. The Pure-State Consistency of Local Density Matrices Problem: In PSPACE and Complete for a Class Between QMA and QMA(2). In *17th Innovations in Theoretical Computer Science Conference (ITCS 2026)*, 2026. doi:10.4230/LIPIcs.ITCS.2026.83.

48 Gen Kimura. The bloch vector for n-level systems. *Physics Letters A*, 314(5–6):339–349, August 2003. doi:10.1016/s0375-9601(03)00941-1.

49 Hirotada Kobayashi, Keiji Matsumoto, and Tomoyuki Yamakami. Quantum merlin-arthur proof systems: Are multiple merlins more helpful to arthur? In *Algorithms and Computation*, 2003. doi:10.1007/978-3-540-24587-2_21.

50 Cécilia Lancien and Andreas Winter. Distinguishing multi-partite states by local measurements. *Communications in Mathematical Physics*, 323(2):555–573, August 2013. doi:10.1007/s00220-013-1779-x.

51 Ke Li and Graeme Smith. Quantum de finetti theorem under fully-one-way adaptive measurements. *Physical Review Letters*, 114(16), April 2015. doi:10.1103/physrevlett.114.160503.

52 Yi-Kai Liu. Consistency of local density matrices is qma-complete. In *APPROX and RANDOM*, 2006. arXiv:quant-ph/0604166.

53 Yi-Kai Liu. *The Complexity of the Consistency and N-representability Problems for Quantum States*. PhD thesis, University of California, San Diego, 2007. arXiv:0712.3041.

54 Yi-Kai Liu, Matthias Christandl, and F. Verstraete. N-representability is qma-complete. *Physical Review Letters*, 98(11), March 2007. doi:10.1103/physrevlett.98.110503.

55 Chris Marriott and John Watrous. Quantum arthur-merlin games. *Comput. Complex.*, 2005. doi:10.1007/s00037-005-0194-x.

56 Miguel Navascués, Masaki Owari, and Martin B. Plenio. Power of symmetric extensions for entanglement detection. *Physical Review A*, 80(5), November 2009. doi:10.1103/physreva.80.052306.

57 Asher Peres. Separability criterion for density matrices. *Physical Review Letters*, 77(8):1413–1415, August 1996. doi:10.1103/physrevlett.77.1413.

58 Attila Pereszlényi. Multi-prover quantum merlin-arthur proof systems with small gap, 2012. arXiv:1205.2761.

59 Yaoyun Shi and Xiaodi Wu. Epsilon-net method for optimizations over separable states. *Theoretical Computer Science*, 598:51–63, 2015. doi:10.1016/j.tcs.2015.03.031.

60 Ashish V. Thapliyal. Multipartite pure-state entanglement. *Physical Review A*, 59(5):3336–3342, May 1999. doi:10.1103/physreva.59.3336.

61 Géza Tóth and Otfried Gühne. Entanglement and permutational symmetry. *Phys. Rev. Lett.*, 102, May 2009. doi:10.1103/PhysRevLett.102.170503.

62 Péter Vrana and Matthias Christandl. Asymptotic entanglement transformation between W and GHZ states. *Journal of Mathematical Physics*, 56(2), February 2015. arXiv:1310.3244.

63 Tzu-Chieh Wei, Michele Mosca, and Ashwin Nayak. Interacting boson problems can be qma hard. *Phys. Rev. Lett.*, 104, January 2010. doi:10.1103/PhysRevLett.104.040501.

64 Xiao-Dong Yu, Timo Simnacher, H. Chau Nguyen, and Otfried Gühne. Quantum-inspired hierarchy for rank-constrained optimization. *PRX Quantum*, 3(1), March 2022. doi:10.1103/prxquantum.3.010340.

65 Xiao-Dong Yu, Timo Simnacher, Nikolai Wyderka, H. Chau Nguyen, and Otfried Gühne. A complete hierarchy for the pure state marginal problem in quantum mechanics. *Nature Communications*, 12(1), February 2021. doi:10.1038/s41467-020-20799-5.

66 D.B. Yudin and A.S. Nemirovskii. Informational complexity and efficient methods for the solution of convex extremal problems. *Ekonomika i Matematicheskie Metody*, 12:357–369, 1976. English translation: Matekon 13 (3) pp.25-45 (1977). URL: https://books.google.com/books?id=0kEzAAAAIAAJ.

67 Karol Życzkowski and Ingemar Bengtsson. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge University Press, 2 edition, 2017. The arXiv link is based on a book chapter. doi:10.1017/9781139207010.

A Formal Definitions

Fix any set of quantum states $\mathcal{W}$. Let $\mathsf{QMA}_{\mathcal{W}}(m)$ as $\mathsf{QMA}(m)$ where Merlin *must* send m unentangled proofs from $\mathcal{W}$. Here we modify both completeness and soundness. Only modifying completeness sometimes has no effect, as seen in [33].

▶ **Definition 11** ($\mathsf{QMA}_{\mathsf{IS}}(m), \mathsf{QMA}_{|\mathsf{IS}\rangle}$)**.** *Let $c, s : \mathbb{N} \to [0,1]$ and $\hat{a}, \hat{b}, \hat{c} : \mathbb{N} \to \mathbb{N}$ be polynomial-time computable functions, and $m : \mathbb{N} \to \mathbb{N}$ be a function.*

A promise problem $A = (A_{yes}, A_{no})$ is in $\mathsf{QMA}_{\mathsf{IS}}^{c,s}[\hat{a}, \hat{b}, \hat{c}](m)$ if and only if there exist a polynomial p and a polynomial-time uniform family of quantum circuits $\{Q_n\}$ that satisfies the following:

1. Q_n *takes as input:*
   - *a string* $x \in \{0,1\}^n$*;*
   - m *quantum states* $\rho_1, \rho_2, \ldots, \rho_m \in A \otimes B \otimes C$*, each over three registers* A, B, C *of dimensions* $\hat{a}, \hat{b}, \hat{c}$ *respectively, and such that each is* internally separable*, meaning that* $\mathrm{Tr}_A(\rho_i)$ *is separable;*
   - *and at most* $p(n)$ *scratch qubits in the state* $|0\rangle$*.*
2. Q_n *is complete and sound for* A*:*
   - *(completeness) If* $x \in A_{yes}$*, there is a set of* m *such quantum states* $\rho_1, \ldots, \rho_m$ *such that* Q_n *accepts* $|x\rangle\langle x| \otimes (\rho_1 \otimes \cdots \otimes \rho_m) \otimes (|0\rangle\langle 0| \otimes \cdots \otimes |0\rangle\langle 0|)$ *with probability at least* $c(n)$*.*
   - *(soundness) If* $x \in A_{no}$*, then for any set of* m *such quantum states* $\rho_1, \ldots, \rho_m$*,* Q_n *accepts* $|x\rangle\langle x| \otimes (\rho_1 \otimes \cdots \otimes \rho_m) \otimes (|0\rangle\langle 0| \cdots \otimes |0\rangle\langle 0|)$ *with probability at most* $s(n)$*.*

Let

$$\mathsf{QMA}_{\mathsf{IS}}^{c,s}(m) := \bigcup_{\substack{\hat{a},\hat{b}=O(poly(n)),\\ \hat{c}=O(1)}} \mathsf{QMA}_{\mathsf{IS}}^{c,s}[\hat{a},\hat{b},\hat{c}](m).$$

Define $\mathsf{QMA}_{|\mathsf{IS}\rangle}^{c,s}$ *similarly, where each quantum state* $\rho_i = |\psi_i\rangle\langle\psi_i|$ *is always pure.*

We also generalize the above definition for other sets of quantum states.

▶ **Definition 12** ($\mathsf{QMA}_{\mathcal{W}}(m)$)**.** *Let* $c, s : \mathbb{N} \to \mathbb{R}^+$ *be polynomial-time computable functions, and* m *a polynomial. Fix an arbitrary set of quantum states* $\mathcal{W}$*.*

A promise problem $A = (A_{yes}, A_{no})$ *is in* $\mathsf{QMA}_{\mathcal{W}}^{c,s}(m)$ *if and only if there exist polynomials* p, q *and a polynomial-time uniform family of quantum circuits* $\{Q_n\}$*, where* Q_n *takes as input a string* $x \in \{0,1\}^n$*,* m $p(n)$*-qubit quantum states in* $\mathcal{W}$*, and* $q(n)$ *scratch qubits in state* $|0\rangle^{\otimes q(n)}$*, such that:*

- *(completeness) If* $x \in A_{yes}$*, there exists a set of* m $p(n)$*-qubit quantum states* $\rho_1, \ldots, \rho_m \in \mathcal{W}$ *such that* Q_n *accepts* $|x\rangle\langle x| \otimes (\rho_1 \otimes \cdots \otimes \rho_m) \otimes |0\rangle\langle 0|^{\otimes q(n)}$ *with probability at least* $c(n)$*.*
- *(soundness) If* $x \in A_{no}$*, then for any set of* m $p(n)$*-qubit quantum states* $\rho_1, \ldots, \rho_m \in \mathcal{W}$*,* Q_n *accepts* $|x\rangle\langle x| \otimes (\rho_1 \otimes \cdots \otimes \rho_m) \otimes |0\rangle\langle 0|^{\otimes q(n)}$ *with probability at most* $s(n)$*.*

▶ Remark 13. Following the convention in [45], we drop the c, s when the completeness-soundness gap is allowed to be any absolute constant: $\mathsf{QMA}_{\mathcal{W}}(m) \stackrel{\text{def}}{=} \bigcup_{c-s=\Omega(1)} \mathsf{QMA}_{\mathcal{W}}^{c,s}(m)$. However, note that $\mathsf{QMA}_{\mathcal{W}}^{c,s}(m)$ is not in general equal to $\mathsf{QMA}_{\mathcal{W}}^{c',s'}(m)$ [10].

We will only consider $\mathsf{QMA}_{|\mathsf{IS}\rangle}$ with one proof, as we prove $\mathsf{QMA}_{|\mathsf{IS}\rangle} = \mathsf{NEXP}$. This implies $\mathsf{QMA}_{|\mathsf{IS}\rangle}(m) = \mathsf{NEXP}$ for all $m \geq 1$.

We occasionally select another property P to choose a set of quantum states which we denote $\mathcal{W}_P$. When this occurs, we use the shorthand $\mathsf{QMA}_P(m) \stackrel{\text{def}}{=} \mathsf{QMA}_{\mathcal{W}_P}(m)$. We use the label $|P\rangle$ for the pure states in $\mathcal{W}_P$.