THEORY OF COMPUTING, Volume 8 (2012), pp. 375–400
www.theoryofcomputing.org

On the Power of a Unique Quantum Witness

Rahul Jain   Iordanis Kerenidis   Greg Kuperberg   Miklos Santha
Or Sattath   Shengyu Zhang

Received: January 31, 2012; published: August 14, 2012.

Abstract: In a celebrated paper, Valiant and Vazirani (1985) raised the question of whether the difficulty of NP-complete problems was due to the wide variation of the number of witnesses of their instances. They gave a strong negative answer by showing that distinguishing between instances having zero or one witnesses is as hard as recognizing NP, under randomized reductions.

We consider the same question in the quantum setting and investigate the possibility of reducing quantum witnesses in the context of the complexity class QMA, the quantum analogue of NP. The natural way to quantify the number of quantum witnesses is the dimension of the witness subspace W in some appropriate Hilbert space $\mathcal{H}$. We present an efficient deterministic procedure that reduces any problem where the dimension d of W is bounded by a polynomial to a problem with a unique quantum witness. The main idea of our reduction is to consider the alternating subspace of the tensor power $\mathcal{H}^{\otimes d}$. Indeed, the intersection of this subspace with $W^{\otimes d}$ is one-dimensional, and therefore can play the role of the unique quantum witness.

ACM Classification: F.1.3

AMS Classification: 81P68

Key words and phrases: Valiant-Vazirani Theorem, unique witness, quantum, QMA

1 Introduction

One of the most fundamental ideas of modern complexity theory is that the study of decision-making procedures involving a single party should be extended to the study of more complex procedures where several parties interact.

 DOI: 10.4086/toc.2012.v008a017

The notions of *verification* and *witness* are at the heart of those complexity classes whose definition inherently involves interaction. The complexity class P is the set of languages decidable by a polynomial-time deterministic algorithm. Similarly, BPP is the set of promise problems decidable by a polynomial-time bounded-error randomized algorithm. We can think of such an algorithm as a verifier acting alone. The simplest interactive extensions of P and BPP are their non-deterministic analogues, respectively NP and MA [11, 6]. These classes also involve an all-powerful prover that sends a single message which is used by the verifier's decision making procedure together with the input. We require that on positive instances there is some message (called in that case a witness) that makes the verifier accept, whereas on negative instances, all the possible messages are rejected by the verifier. In the case of MA we can fix the permitted error of the verifier, rather arbitrarily to any constant, say $1/3$.

Quantum complexity classes are often defined by analogy to their classical counterparts. Since quantum computation is inherently probabilistic, the quantum analogue of MA is considered to be the right definition of non-deterministic quantum polynomial-time. The quantum extension is twofold: the verifier has the power to decide promise problems in BQP, quantum polynomial-time, and the messages he receives from the prover are also quantum. Thus, QMA is the set of promise problems such that on positive instances there exists a quantum witness accepted with probability at least $2/3$ by the polynomial-time quantum verifier and on negative instances the verifier accepts every quantum state with probability at most $1/3$. While the idea that a quantum state might play the role of a witness goes back to Knill [20], the class was formally defined by Kitaev [19] under the name of BQNP. The currently used name QMA was given to the class by Watrous [32]. Kitaev established several error probability reduction properties of QMA, and proved that the Local Hamiltonian, the quantum analogue of SAT, is complete for it. Watrous showed that Group non-Membership was a problem in QMA and based on this result he constructed an oracle under which MA is strictly included in QMA. Since then, various problems have been proven to be complete for QMA [16, 18, 23, 17, 24]. A potentially weaker quantum extension of MA, namely QCMA, was defined by Aharonov and Naveh [3]: in the case of QCMA, the verifier is still a quantum polynomial-time algorithm, but the message of the prover can only be classical.

The number of witnesses for positive instances of problems in NP can be exponentially high. Also, known NP-complete problems have different instances with widely varying numbers of solutions. In a celebrated paper, Valiant and Vazirani [31] have raised the question of whether the difficulty of the class NP was due to this wide variation. They gave a strong negative answer to this question in the following sense. Let UP be the set of problems in NP where in addition on positive instances there exists a unique witness. We denote by PromiseUP the extension of UP from languages to promise problems. The theorem of Valiant and Vazirani states that any problem in NP can be reduced in randomized polynomial-time to a promise problem in PromiseUP, or in set-theoretical terms, $\mathsf{NP} \subseteq \mathsf{RP}^{\mathsf{PromiseUP}}$.[1] The importance of the class UP also stems from its connection to one-way functions: worst-case one-way functions exist if and only if $\mathsf{UP} \neq \mathsf{P}$ [21, 12].

In a recent paper Aharonov, Ben-Or, Brandão and Sattah [1] have asked a similar question for MA, QCMA and QMA. The restriction of the classical-witness classes MA and QCMA to their unique variants UMA and UQCMA is rather natural: no change for negative instances, but on positive instances

[1] RP is the subclass of problems in BPP where the computation does not err on negative instances, and $\mathsf{RP}^{\mathsf{PromiseUP}}$ is the same, except that the RP machine has oracle access to any PromiseUP problem (the oracle can answer arbitrarily when asked about instances that do not obey the promise).

there has to be exactly one witness that makes the verifier accept with probability at least $2/3$, while all other messages make him accept with probability at most $1/3$. The definition of UQMA, the unique variant of QMA is the following: there is no change for negative instances with respect to QMA, but on positive instances there has to be a quantum witness state $|\psi\rangle$ which is accepted by the verifier with probability at least $2/3$, whereas all states orthogonal to $|\psi\rangle$ are accepted with probability at most $1/3$. Aharonov et al. extended the Valiant-Vazirani proof for the classical witness classes by showing that $\mathsf{MA} \subseteq \mathsf{RP}^{\mathsf{UMA}}$ and $\mathsf{QCMA} \subseteq \mathsf{RP}^{\mathsf{UQCMA}}$. On the other hand, they left the existence of a similar result for QMA as an open problem.

Why is it so difficult to reduce the witnesses to a single witness in the quantum case? The basic idea of Valiant and Vazirani is to use pairwise-independent universal hash functions, having polynomial-size descriptions, that eliminate independently each witness with some constant probability. The size of the original witness set can be guessed approximately by a polynomial-time probabilistic procedure, and in case of a correct guess the hashing keeps alive exactly one witness with again some constant probability. The same idea basically works for MA and QCMA as long as one additional difficulty is overcome: on positive instances there can be exponentially more "pseudo-witnesses," accepted with probability between $1/3$ and $2/3$, than witnesses which are accepted with probability at least $2/3$. In this case, the Valiant-Vazirani proof technique will with high probability eliminate all witnesses before the elimination of the pseudo-witnesses. The solution of Aharonov et al. for this problem is to divide the interval $(1/3, 2/3)$ into polynomially many smaller intervals, and to show that there exists at least one interval such that there are approximately as many witnesses accepted with probability within this interval as above it.

In the quantum case, the set of quantum witnesses can be infinite. For a promise problem in QMA, we can suppose without loss of generality that on positive instances there exists a subspace W such that all unit vectors in W are accepted with high probability. The dimension of W could be large and we wish to reduce it to one. Aharonov et al. [1] considered the special case where the dimension of W is two. Although classically two witnesses are trivially reducible to the unique witness case, they have shown that the natural generalization of the Valiant-Vazirani construction cannot solve even the two-dimensional quantum witness case.

Indeed, the natural generalization of the Valiant-Vazirani construction to this situation is to use random projections and hope that some one-dimensional subspace of W will be accepted with substantially higher probability than its orthogonal. A first difficulty is to implement such projections efficiently. But more importantly, a random projection would not create a polynomial gap in the acceptance probabilities for the pure states of W: fact all states in W which were accepted with exponentially close probabilities, will still be accepted after the random projection with exponentially close probabilities. In fact, if two states in W were accepted with probabilities p_1, p_2 where $|p_1 - p_2| = O(\exp(-n))$, they will be accepted after the random projection with probabilities p'_1, p'_2 such that $|p'_1 - p'_2| = O(\exp(-n))$, with high probability over the choice of the random projection.

Here we describe a fundamentally different proof technique to tackle this problem, which is sufficiently powerful to solve the case when the dimension of the witness subspace W is polynomially bounded in the length of the input. This leads us naturally to the quantum analogue of the promise problem class FewP. This complexity class was defined by Allender [4] as the set of problems in NP with the additional constraint that there is a polynomial q such that on every positive instance of length n, the number of witnesses is at most $q(n)$. The class FewP was extensively studied in the context of counting complexity classes [5, 30, 22, 14, 28]. We define FewQMA, the quantum analogue of FewP, as the set of promise problems in QMA for which there exists a polynomial q with the following properties: on negative instances every message of the prover is accepted by the verifier with probability at most $1/3$; on a positive instance x there exists a subspace W_x of dimension between 1 and $q(|x|)$, such that all pure states in W_x are accepted with probability at least $2/3$, while all pure states orthogonal to W_x are accepted with probability at most $1/3$. Our main theorem extends the result of Valiant and Vazirani to this complexity class. More precisely, we show that FewQMA is deterministic polynomial-time Turing-reducible to UQMA.

Main Theorem. $\mathsf{FewQMA} \subseteq \mathsf{P}^{\mathsf{UQMA}}$.

The main result can alternatively be formulated by using Hamiltonian Complexity terminology. Informally, estimating the ground state energy of a poly-gapped local Hamiltonian with a polynomial degeneracy, is not harder than estimating the ground state energy of a poly-gapped local Hamiltonian with a unique ground state. See Theorem 4.9 for the formal statement.

To explain the intuition behind the construction, let us first examine a simple proof for $\mathsf{FewP} \subseteq \mathsf{P}^{\mathsf{PromiseUP}}$. Suppose the FewP problem had d witnesses. Instead of asking for one witness, we ask for the concatenation of t witnesses and require that each one of the witnesses pass the original FewP test. At first glance, this does not seem to be going in the right direction because the number of witnesses at this point is d^t. Then, we check that the witnesses are in a specific increasing order, such as the increasing lexicographic order, and therefore the total number of witnesses is reduced to $\binom{d}{t}$. If $t = d$, there is exactly one witness. Of course, we do not know d in advance, but since we have a polynomial bound $q(|x|)$ on it, we try every possible value t between 1 and $q(|x|)$.

We use a similar technique in the quantum case: instead of manipulating the states within the original space $\mathcal{H}$ of dimension K, we consider its t-fold tensor powers $\mathcal{H}^{\otimes t}$. At this point, like in the classical case, the dimension of $W^{\otimes t}$ grows as d^t, where d is the dimension of the witness space W. In the quantum case, instead of forcing the witnesses to be in one specific order, we force the witnesses to be a signed superposition of all possible permutations, where the sign depends on the parity (also known as the signature or sign) of the permutation. The subspace spanned by such states is called the alternating subspace Alt of $\mathcal{H}^{\otimes t}$ whose dimension is $\binom{K}{t}$. The important thing to notice is that the dimension of the intersection $\mathsf{Alt} \cap W^{\otimes t}$ is equal to 1 when $t = d$. The reason is that this intersection is in fact equal to the alternating subspace of $W^{\otimes t}$ whose dimension is $\binom{d}{t}$. Therefore, we will choose this one-dimensional subspace as our unique quantum witness. Again, as in the classical case, we do not know the exact dimension of W, but since we have a polynomial upper bound $q(|x|)$ on it, we just try every possible value t between 1 and $q(|x|)$.

To give a concrete example, suppose the orthogonal states $|1\rangle$, $|2\rangle$ and $|3\rangle$ span W. Then, the state

$$|\psi\rangle = \frac{1}{\sqrt{3!}}(|123\rangle - |213\rangle - |132\rangle - |321\rangle + |231\rangle + |312\rangle) = \frac{1}{\sqrt{3!}} \sum_{\sigma \in S_3} \operatorname{sgn}(\sigma) |\sigma(1)\sigma(2)\sigma(3)\rangle ,$$

is the unique state, up to a global phase, in $\mathsf{Alt} \cap W^{\otimes 3}$, where $|ijk\rangle$ is a shorthand for $|i\rangle \otimes |j\rangle \otimes |k\rangle$.

More details are presented as follows. For illustration, let us assume that there are d orthogonal witnesses that make the verifier accept with probability 1, and all other states in W orthogonal to these

d witnesses make the verifier accept with probability 0. For a fixed t, we would ideally implement $\Pi_{W^{\otimes t}} \cdot \Pi_{\mathsf{Alt}}$, the projection to Alt followed by the projection to $W^{\otimes t}$. This procedure would accept only one state for the following reason. The unique pure state in $\mathsf{Alt} \cap W^{\otimes t}$ (up to a global phase) is clearly accepted with probability 1. On the other hand, we claim that any state $|\phi\rangle$ orthogonal to that is rejected with probability 1. Indeed, $|\phi\rangle$ can be decomposed as $|\phi_1\rangle + |\phi_2\rangle$, where $|\phi_1\rangle \in \mathsf{Alt}^{\perp}$ and $|\phi_2\rangle \in W^{\otimes t\perp}$. Therefore $|\phi_1\rangle$ is rejected by Π_{Alt} and $|\phi_2\rangle$ is rejected by $\Pi_{W^{\otimes t}}$. This implies the claim since we can show that the two projectors actually commute.

We can efficiently implement Π_{Alt} by a procedure we call the Alternating Test. A similar procedure to ours, implementing efficiently the projection to the symmetric subspace Sym of $\mathcal{H}^{\otimes t}$, was proposed by Barenco et al. [7] as the basis of a method for the stabilization of quantum computations. In fact, in the two-fold tensor product case, the two procedures coincide and become the well-known Swap Test which was used by Buhrman et al. [9] for deciding if two given pure states are close or far apart.

We cannot implement $\Pi_{W^{\otimes t}}$ exactly, but we can approximate it efficiently by a procedure called the Witness Test. This test just applies independently to all the t components of the state the procedure at our disposal which decides in $\mathcal{H}$ whether a state is a witness or not, and accepts if all applications accept. There is only one difficulty left: since Π_{Alt} and the Witness Test do not necessarily commute, our previous argument which showed that states in $W^{\otimes t\perp}$ were rejected with probability 1 does not work anymore. We overcome this difficulty by showing that the commutativity of the two projections implies that the projections to Alt of such states are also in $W^{\otimes t\perp}$, and therefore get rejected with high probability by the Witness Test.

Note that our reduction (from FewQMA to UQMA) is a deterministic Turing reduction, while the reduction (from FewP to PromiseUP) used by Valiant-Vazirani is a randomized many-one reduction. As mentioned after the main theorem, a deterministic Turing reduction can be shown, implying that $\mathsf{FewP} \subseteq \mathsf{P}^{\mathsf{PromiseUP}}$. Furthermore, it is easy to convert our reduction to a randomized many-one reduction by choosing t in our algorithm uniformly between 1 and $q(|x|)$, instead of trying all possible t between 1 and $q(|x|)$; by doing that, we lose a multiplicative factor of $1/q(|x|)$ in the completeness parameter. Our reduction is also non-adaptive, similar to the reduction used by Valiant-Vazirani. We believe that reducing QMA to a unique witness, which this paper leaves as an open question, will require a probabilistic or a quantum procedure.

The rest of the paper is structured as follows. In Section 2 we state some facts about the interaction of the tensor products of subspaces with the alternating subspace. In Section 3 we define the complexity classes we are concerned with. We give two definitions for FewQMA and show that they are equivalent. Section 4 is devoted to the proof of our main result, and to the proof of the alternative Hamiltonian complexity formulation of this theorem. Finally, in Section 5 we consider a third definition and show a weak equivalence with the previous ones. The results in the paper appeared initially as Arxiv preprint quant-ph/0906.4425v1 by Jain, Kerenidis, Santha and Zhang. Some of this work was done independently by Sattath and Kuperberg and an initial result in that direction appeared in p. 30 of [29]. The joint version of the paper appears as Arxiv preprint quant-ph/0906.4425v2.

2 Preliminaries

In this section we present definitions and lemmas that we will need in the proof of our main result.

We represent by $[t]$ the set $\{1,2,\ldots,t\}$. For a Hilbert space $\mathcal{H}$, we denote by $\dim(\mathcal{H})$ the dimension of $\mathcal{H}$. For a subspace S of $\mathcal{H}$, let $S^{\perp}$ represent the subspace of $\mathcal{H}$ orthogonal to S, and let Π_S denote the projector onto S. For subspaces S_1, S_2 of $\mathcal{H}$, their *direct sum* $S_1 + S_2$ is defined as $\mathsf{span}(S_1 \cup S_2)$, and when S_1, S_2 are orthogonal subspaces, we denote their (orthogonal) direct sum by $S_1 \oplus S_2$. The following relations are standard.

Fact 2.1.

1. *Let S_1, S_2 be subspaces of a Hilbert space $\mathcal{H}$. Then $(S_1 \cap S_2)^{\perp} = S_1^{\perp} + S_2^{\perp}$.*
2. *Let S_1, S_2 be subspaces of Hilbert spaces $\mathcal{H}_1$, $\mathcal{H}_2$ respectively. Then*

$$(S_1 \otimes S_2)^{\perp} = (S_1^{\perp} \otimes \mathcal{H}_2) + (\mathcal{H}_1 \otimes S_2^{\perp}) = (S_1^{\perp} \otimes \mathcal{H}_2) \oplus (S_1 \otimes S_2^{\perp}).$$

Let $\mathcal{B}$ represent the two-dimensional complex Hilbert space and let $\{|0\rangle, |1\rangle\}$ be the computational basis for $\mathcal{B}$. For a natural number k, the computational basis of $\mathcal{B}^{\otimes k}$ (the k-fold tensor of $\mathcal{B}$) consists of $\{|r\rangle : r \in \{0,1\}^k\}$, where $|r\rangle$ denotes the tensor product $|r_1\rangle \otimes \cdots \otimes |r_k\rangle$ for the k-bit string $r = r_1 \ldots r_k$. Fix k and let $\mathcal{H}$ denote $\mathcal{B}^{\otimes k}$ and let $K = 2^k$. By a *pure state* in $\mathcal{H}$, we mean a unit vector in $\mathcal{H}$. A *mixed state* or just *state* is a positive semi-definite operator in $\mathcal{H}$ with trace 1. We refer the reader to the text [26] for concepts related to quantum information theory. For a natural number $t \in [K]$, we will think of states of $\mathcal{H}^{\otimes t}$ as consisting of t registers, where the content of each register is a state with support in $\mathcal{H}$.

We will consider the intersection of $W^{\otimes t}$, where W is a d-dimensional subspace of $\mathcal{H}$ for some d satisfying $2 \le t \le d \le K$, with the alternating and symmetric subspaces of $\mathcal{H}^{\otimes t}$. Let S_t denote the set of all permutations $\pi : [t] \to [t]$.

For a permutation $\pi \in S_t$, let the unitary operator U_π, acting on $\mathcal{H}^{\otimes t}$, given by

$$U_\pi |\psi_1\rangle \otimes \cdots \otimes |\psi_t\rangle = |\psi_{\pi(1)}\rangle \otimes \cdots \otimes |\psi_{s(t)}\rangle,$$

and extended by linearity to the entire space $\mathcal{H}^{\otimes t}$.

For permutations π_1, π_2, let $\pi_1 \circ \pi_2$ represent their composition. It is easily seen that $U_{\pi_1 \circ \pi_2} = U_{\pi_1} U_{\pi_2}$. For distinct $i, j \in [t]$, let π_{ij} be the transposition of i and j. For all distinct $i, j \in [t]$, the symmetric subspace of $W^{\otimes t}$ with respect to i and j is given by

$$\mathsf{Sym}_{ij}^{W^{\otimes t}} = \{|\phi\rangle \in W^{\otimes t} : U_{\pi_{ij}} |\phi\rangle = |\phi\rangle\},$$

and the *symmetric subspace* of $W^{\otimes t}$ is defined as

$$\mathsf{Sym}^{W^{\otimes t}} = \bigcap_{i \neq j} \mathsf{Sym}_{ij}^{W^{\otimes t}}.$$

Similarly, for all distinct $i, j \in [t]$, the alternating subspace of $W^{\otimes t}$ with respect to i and j is defined as

$$\mathsf{Alt}_{ij}^{W^{\otimes t}} = \{|\phi\rangle \in W^{\otimes t} : U_{\pi_{ij}} |\phi\rangle = -|\phi\rangle\},$$

and the *alternating subspace* of $W^{\otimes t}$ is defined as

$$\mathsf{Alt}^{W^{\otimes t}} = \bigcap_{i \neq j} \mathsf{Alt}_{ij}^{W^{\otimes t}}.$$

The subspaces $\mathsf{Sym}^{W^{\otimes t}}$ and $\mathsf{Alt}^{W^{\otimes t}}$ are of dimension $\binom{d+t-1}{t}$ and $\binom{d}{t}$ respectively [8, Ch. I.5]. In particular, $\mathsf{Alt}^{\mathcal{H}^{\otimes 2}}$ and $\mathsf{Sym}^{\mathcal{H}^{\otimes 2}}$ have respective dimensions $\binom{K+1}{2}$ and $\binom{K}{2}$ and, since they are orthogonal, we have $\mathcal{H}^{\otimes 2} = \mathsf{Alt}^{\mathcal{H}^{\otimes 2}} \oplus \mathsf{Sym}^{\mathcal{H}^{\otimes 2}}$. This implies that for every distinct $i, j \in [t]$, we have

$$\mathsf{Alt}_{ij}^{W^{\otimes t}} \oplus \mathsf{Sym}_{ij}^{W^{\otimes t}} = W^{\otimes t}.$$

Claim 2.2. $\left(\mathsf{Alt}^{W^{\otimes t}}\right)^{\perp} \cap W^{\otimes t} = \sum_{i \neq j} \mathsf{Sym}_{ij}^{W^{\otimes t}}$.

Proof. Since $\mathsf{Alt}_{ij}^{W^{\otimes t}} \oplus \mathsf{Sym}_{ij}^{W^{\otimes t}} = W^{\otimes t}$, we have $\left(\mathsf{Alt}_{ij}^{W^{\otimes t}}\right)^{\perp} = \mathsf{Sym}_{ij}^{W^{\otimes t}} \oplus \left(W^{\otimes t}\right)^{\perp}$. Therefore

$$\begin{aligned}
\left(\mathsf{Alt}^{W^{\otimes t}}\right)^{\perp} \cap W^{\otimes t} &= \left(\left(\bigcap_{i \neq j} \mathsf{Alt}_{ij}^{W^{\otimes t}}\right)^{\perp}\right) \cap W^{\otimes t} && \text{(by definition of } \mathsf{Alt}^{W^{\otimes t}}\text{)} \\
&= \left(\sum_{i \neq j} \left(\mathsf{Alt}_{ij}^{W^{\otimes t}}\right)^{\perp}\right) \cap W^{\otimes t} && \text{(from Fact 2.1)} \\
&= \left(\sum_{i \neq j} \left(\mathsf{Sym}_{ij}^{W^{\otimes t}} \oplus \left(W^{\otimes t}\right)^{\perp}\right)\right) \cap W^{\otimes t} \\
&= \left(\left(\sum_{i \neq j} \mathsf{Sym}_{ij}^{W^{\otimes t}}\right) \oplus \left(W^{\otimes t}\right)^{\perp}\right) \cap W^{\otimes t} \\
&= \sum_{i \neq j} \mathsf{Sym}_{ij}^{W^{\otimes t}}.
\end{aligned}$$

The last equality holds since $\sum_{i \neq j} \mathsf{Sym}_{ij}^{W^{\otimes t}} \subseteq W^{\otimes t}$. □

Note that for $W = \mathcal{H}$ the claim states that

$$\left(\mathsf{Alt}^{\mathcal{H}^{\otimes t}}\right)^{\perp} = \sum_{i \neq j} \mathsf{Sym}_{ij}^{\mathcal{H}^{\otimes t}}.$$

For us, a particularly important case is when the number of registers t is equal to d, the dimension of the subspace W. Then the alternating subspace $\mathsf{Alt}^{W^{\otimes d}}$ is one-dimensional. Let $\{|\psi_1\rangle, \ldots, |\psi_d\rangle\}$ be any orthonormal basis of W, and let the vector $|W_{\mathrm{alt}}\rangle \in W^{\otimes d}$ be defined as

$$|W_{\mathrm{alt}}\rangle = \frac{1}{\sqrt{d!}} \sum_{\pi \in S_d} \operatorname{sgn}(\pi)\, U_\pi |\psi_1\rangle \otimes \cdots \otimes |\psi_d\rangle, \tag{2.1}$$

where $\operatorname{sgn}(\pi)$ denotes the sign of the permutation π. The following claim states that $|W_{\mathrm{alt}}\rangle$ spans the one-dimensional subspace $\mathsf{Alt}^{W^{\otimes d}}$. This immediately implies that $|W_{\mathrm{alt}}\rangle$ is independent of the choice of the basis (up to a global phase).

Claim 2.3. $\mathsf{Alt}^{W^{\otimes d}} = \mathsf{span}\{|W_{\mathrm{alt}}\rangle\}$.

Proof. We show that $|W_{\mathrm{alt}}\rangle \in \mathsf{Alt}^{W^{\otimes d}}$. This implies the statement since $\dim\left(\mathsf{Alt}^{W^{\otimes d}}\right) = \binom{d}{d} = 1$. For any distinct $i, j \in [d]$ we show $|W_{\mathrm{alt}}\rangle \in \mathsf{Alt}_{ij}^{W^{\otimes d}}$. For a permutation $\pi \in S_d$, we set $\pi' = \pi_{ij} \circ \pi$. We then have

$$\begin{aligned} U_{\pi_{ij}}|W_{\mathrm{alt}}\rangle &= U_{\pi_{ij}} \frac{1}{\sqrt{d!}} \sum_{\pi \in S_d} \mathrm{sgn}(\pi)\, U_\pi |\psi_1\rangle \otimes \cdots \otimes |\psi_d\rangle \\ &= \frac{1}{\sqrt{d!}} \sum_{\pi \in S_d} \mathrm{sgn}(\pi)\, U_{\pi_{ij}\circ\pi} |\psi_1\rangle \otimes \cdots \otimes |\psi_d\rangle \\ &= \frac{1}{\sqrt{d!}} \sum_{\pi' \in S_d} \mathrm{sgn}(\pi_{ij}^{-1} \circ \pi')\, U_{\pi'} |\psi_1\rangle \otimes \cdots \otimes |\psi_d\rangle \\ &= (-1) \cdot \frac{1}{\sqrt{d!}} \sum_{\pi' \in S_d} \mathrm{sgn}(\pi')\, U_{\pi'} |\psi_1\rangle \otimes \cdots \otimes |\psi_d\rangle \\ &= -|W_{\mathrm{alt}}\rangle\,, \end{aligned}$$

where we used that $\mathrm{sgn}(\pi_{ij}^{-1} \circ \pi') = -\mathrm{sgn}(\pi')$. $\square$

Next, we show that the projections on the spaces $\mathsf{Alt}^{\mathcal{H}^{\otimes t}}$ and $W^{\otimes t}$ commute for any $2 \le t \le d$.

Claim 2.4. *Let $2 \le t \le d$. Then,* $\Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes t}}} \cdot \Pi_{W^{\otimes t}} = \Pi_{W^{\otimes t}} \cdot \Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes t}}}$.

Proof. Set $T = \left(\mathsf{Alt}^{W^{\otimes t}}\right)^{\perp} \cap W^{\otimes t}$. Then $W^{\otimes t} = \mathsf{Alt}^{W^{\otimes t}} \oplus T$ and hence, $\Pi_{W^{\otimes t}} = \Pi_{\mathsf{Alt}^{W^{\otimes t}}} + \Pi_T$. We have

$$\begin{aligned} T &= \left(\mathsf{Alt}^{W^{\otimes t}}\right)^{\perp} \cap W^{\otimes t} \\ &= \sum_{i \ne j} \mathsf{Sym}_{ij}^{W^{\otimes t}} && \text{(from Claim 2.2)} \\ &\subseteq \sum_{i \ne j} \mathsf{Sym}_{ij}^{\mathcal{H}^{\otimes t}} && \text{(by definition)} \\ &= \left(\mathsf{Alt}^{\mathcal{H}^{\otimes t}}\right)^{\perp} && \text{(from Claim 2.2)}. \end{aligned}$$

This implies that $\Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes t}}} \cdot \Pi_T = \Pi_T \cdot \Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes t}}} = \mathbf{0}$. Also, $\mathsf{Alt}^{W^{\otimes t}} \subseteq \mathsf{Alt}^{\mathcal{H}^{\otimes t}}$ since $\mathsf{Alt}^{W^{\otimes t}} = \mathsf{Alt}^{\mathcal{H}^{\otimes t}} \cap W^{\otimes t}$. Therefore, we have

$$\Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes t}}} \cdot \Pi_{W^{\otimes t}} = \Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes t}}} \cdot (\Pi_{\mathsf{Alt}^{W^{\otimes t}}} + \Pi_T) = \Pi_{\mathsf{Alt}^{W^{\otimes t}}}\,.$$

Similarly

$$\Pi_{W^{\otimes t}} \cdot \Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes t}}} = \Pi_{\mathsf{Alt}^{W^{\otimes t}}}\,.$$

Hence $\Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes t}}} \cdot \Pi_{W^{\otimes t}} = \Pi_{W^{\otimes t}} \cdot \Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes t}}}$. $\square$

This commutativity relation enables us to derive the following property

Claim 2.5. *For any state $|\phi\rangle \in \left(\mathsf{Alt}^{W^{\otimes d}}\right)^{\perp}$, we have $\Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes d}}} |\phi\rangle \in (W^{\otimes d})^{\perp}$.*

Proof. First note that $\mathsf{Alt}^{W^{\otimes d}} = \mathsf{Alt}^{\mathcal{H}^{\otimes d}} \cap W^{\otimes d}$, and therefore, $\left(\mathsf{Alt}^{W^{\otimes d}}\right)^{\perp} = \left(\mathsf{Alt}^{\mathcal{H}^{\otimes d}} \cap W^{\otimes d}\right)^{\perp}$ and by Fact 2.1,

$$\left(\mathsf{Alt}^{W^{\otimes d}}\right)^{\perp} = \left(\mathsf{Alt}^{\mathcal{H}^{\otimes d}}\right)^{\perp} + \left(W^{\otimes d}\right)^{\perp}.$$

Hence we can decompose $|\phi\rangle$ as $|\phi_1\rangle + |\phi_2\rangle$, where $|\phi_1\rangle \in \left(\mathsf{Alt}^{\mathcal{H}^{\otimes d}}\right)^{\perp}$ and $|\phi_2\rangle \in \left(W^{\otimes d}\right)^{\perp}$. As

$$\Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes d}}} |\phi_1\rangle = \mathbf{0},$$

it suffices to show that $\Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes d}}} |\phi_2\rangle \in (W^{\otimes d})^{\perp}$. For this, we prove that

$$\Pi_{(W^{\otimes d})^{\perp}} \cdot \Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes d}}} |\phi_2\rangle = \Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes d}}} |\phi_2\rangle .$$

Claim 2.4 implies that

$$\Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes d}}} \cdot \Pi_{(W^{\otimes d})^{\perp}} = \Pi_{(W^{\otimes d})^{\perp}} \cdot \Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes d}}} .$$

Also, $|\phi_2\rangle = \Pi_{(W^{\otimes d})^{\perp}} |\phi_2\rangle$ since $|\phi_2\rangle \in \left(W^{\otimes d}\right)^{\perp}$. Therefore we can conclude by the following equalities:

$$\Pi_{(W^{\otimes d})^{\perp}} \cdot \Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes d}}} |\phi_2\rangle \;=\; \Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes d}}} \cdot \Pi_{(W^{\otimes d})^{\perp}} |\phi_2\rangle \;=\; \Pi_{\mathsf{Alt}^{\mathcal{H}^{\otimes d}}} |\phi_2\rangle. \qquad \square$$

3 Complexity classes

In this section we define the relevant complexity classes and state the facts needed about them. We assume without loss of generality that there is only a measurement of a single qubit at the end of the quantum computation. For a quantum circuit V, we let V also represent the unitary transformation performed by the circuit. We call a *verification procedure* a family of quantum circuits $\{V_x : x \in \{0,1\}^*\}$ uniformly generated in polynomial-time, together with polynomials k and m such that V_x acts on $k(|x|) + m(|x|)$ qubits. We refer to the first $k(|x|)$ qubits as *message qubits* and to the last $m(|x|)$ qubits as *auxiliary qubits*. To simplify notation, when the input x is implicit in the discussion, we refer to $k(|x|)$ by k, and to $m(|x|)$ by m. We will make repeated use of the following projections in $\mathcal{B}^{\otimes(k+m)}$:

$$\Pi_{\text{acc}} = |1\rangle\langle 1| \otimes I_{k+m-1}, \qquad \Pi_{\text{init}} = I_k \otimes |0^m\rangle\langle 0^m|,$$

where I_n is the identity operator on n qubits. We will also make use of the operator Π_x defined as

$$\Pi_x = \Pi_{\text{init}} V_x^{\dagger} \Pi_{\text{acc}} V_x \Pi_{\text{init}}. \tag{3.1}$$

It is easy to see that Π_x is positive semi-definite.

Given a verification procedure, on input x, a Quantum Merlin-Arthur protocol proceeds in the following way: the prover Merlin sends a pure state $|\psi\rangle \in \mathcal{B}^{\otimes k}$ to the verifier Arthur, who then applies the circuit V_x to $|\psi\rangle \otimes |0^m\rangle$, and accepts if the measurement of the first qubit of the result gives 1. We will denote the probability that Arthur accepts x with state $|\psi\rangle$ by $\Pr[V_x \text{ outputs Accept on } |\psi\rangle]$, which is equal to $\|\Pi_{\text{acc}} V_x(|\psi\rangle \otimes |0^m\rangle)\|^2$.

A promise problem is a tuple $L = (L_{\text{yes}}, L_{\text{no}})$ with $L_{\text{yes}} \cup L_{\text{no}} \subseteq \{0,1\}^*$ and $L_{\text{yes}} \cap L_{\text{no}} = \emptyset$. We now define the following complexity classes.

Definition 3.1. A promise problem $L = (L_{\text{yes}}, L_{\text{no}})$ is in the complexity class *Quantum Merlin-Arthur* (denoted QMA) if there exists a verification procedure $\{V_x : x \in \{0,1\}^*\}$ with polynomials k and m such that

1. for all $x \in L_{\text{yes}}$, there exists a state $|\psi\rangle$, called a witness, such that $\left\|\Pi_{\text{acc}} V_x(|\psi\rangle \otimes |0^m\rangle)\right\|^2 \geq 2/3$,
2. for all $x \in L_{\text{no}}$, and for all state $|\psi\rangle$, $\left\|\Pi_{\text{acc}} V_x(|\psi\rangle \otimes |0^m\rangle)\right\|^2 \leq 1/3$.

If the input x is not in $L_{\text{yes}} \cup L_{\text{no}}$ then there is no requirement put on the verification procedure. Same thing holds for all the relevant definitions below.

For a promise problem in QMA, it can be shown that on positive instances there exists a subspace W such that all unit vectors in W are accepted with probability higher than the completeness parameter; see equation (3.3) for a precise definition. Hence, by putting a polynomial upper bound on the dimension of this subspace, we derive the following definition:

Definition 3.2. Let $c, w, s : \mathbb{N} \to [0,1]$ be polynomial-time computable functions such that $c(n) > \max\{w(n), s(n)\}$ for all $n \in \mathbb{N}$. A promise problem $L = (L_{\text{yes}}, L_{\text{no}})$ is in the complexity class *Few Quantum Merlin-Arthur* (denoted $\mathsf{FewQMA}(c,w,s)$) if there exists a verification procedure $\{V_x : x \in \{0,1\}^*\}$ with polynomials k and m, and a polynomial q such that

1. for all $x \in L_{\text{yes}}$ there exists a subspace W_x of $\mathcal{B}^{\otimes k}$ with $\dim(W_x) \in [q(|x|)]$ such that
   (a) for all states $|\psi\rangle \in W_x$, $\left\|\Pi_{\text{acc}} V_x(|\psi\rangle \otimes |0^m\rangle)\right\|^2 \geq c(|x|)$, and
   (b) for all states $|\phi\rangle \in W_x{}^{\perp}$, $\left\|\Pi_{\text{acc}} V_x(|\phi\rangle \otimes |0^m\rangle)\right\|^2 \leq w(|x|)$,
2. for all $x \in L_{\text{no}}$ and for all pure states $|\psi\rangle \in \mathcal{B}^{\otimes k}$, $\left\|\Pi_{\text{acc}} V_x(|\psi\rangle \otimes |0^m\rangle)\right\|^2 \leq s(|x|)$.

Definition 3.3. The class *Few Quantum Merlin-Arthur* (denoted FewQMA) is $\mathsf{FewQMA}(2/3, 1/3, 1/3)$.

We next provide an alternative definition of $\mathsf{FewQMA}(c,w,s)$ and show that the two definitions are equivalent.

Definition 3.4. Let $c, w, s : \mathbb{N} \to [0,1]$ be polynomial-time computable functions such that, for all $n \in \mathbb{N}$, $c(n) > \max\{w(n), s(n)\}$. A promise problem $L = (L_{\text{yes}}, L_{\text{no}})$ is in the complexity class *Alternative Few Quantum Merlin-Arthur* (denoted $\mathsf{Alt\text{-}FewQMA}(c,w,s)$) if there exists a verification procedure $\{V_x : x \in \{0,1\}^*\}$ with polynomials k and m, and a polynomial q such that

1. for all $x \in L_{\text{yes}}$, the number of eigenvalues of Π_x which are at least $c(|x|)$ is in $[q(|x|)]$, and no eigenvalue of Π_x is in the open interval $(w(|x|), c(|x|))$,
2. for all $x \in L_{\text{no}}$, all eigenvalues of Π_x are at most $s(|x|)$.

We prove the following equivalence between the two definitions.

Theorem 3.5. *Let $c, w, s : \mathbb{N} \to [0,1]$ be polynomial-time computable functions such that, for all $n \in \mathbb{N}$, $c(n) > \max\{w(n), s(n)\}$. Then* $\mathsf{FewQMA}(c,w,s) = \mathsf{Alt\text{-}FewQMA}(c,w,s)$.

Proof. **Part 1** (Definition 3.4 $\Rightarrow$ Definition 3.2): Let L be a promise problem in $\mathsf{Alt\text{-}FewQMA}(c,w,s)$ with some verification procedure $\{V_x\}$ and polynomial q according to Definition 3.4. We show that $\{V_x\}$ and q satisfy also Definition 3.2 with the same parameters.

We first consider the case $x \in L$. For every $|u\rangle \in \mathcal{B}^{\otimes(k+m)}$, we have

$$\left\|\Pi_{\mathrm{acc}} V_x \Pi_{\mathrm{init}} |u\rangle\right\|^2 = \langle u|\Pi_x|u\rangle\,. \tag{3.2}$$

It is easy to check that all eigenvectors of Π_x are also eigenvectors of Π_{init}. The 1-eigenvectors of the projector Π_{init} are of the form $|u\rangle \otimes |0^m\rangle$ with $|u\rangle \in \mathcal{B}^{\otimes k}$, and any vector orthogonal to these is an eigenvector with eigenvalue 0. Let r be the number of eigenvalues of Π_x that are at least $c(|x|)$, by hypothesis $r \in [q(|x|)]$. Let $\{|v_i\rangle \otimes |0^m\rangle : i \in [r]\}$ be a set of orthonormal eigenvectors of Π_x with respective eigenvalues $\{\lambda_i \geq c(|x|) : i \in [r]\}$, we set $W_x = \mathsf{span}(\{|v_i\rangle : i \in [r]\})$. Then all remaining $2^{k+m} - r$ eigenvalues of Π_x are less than or equal to $w(|x|)$. Let $\{|v_i\rangle \otimes |0^m\rangle : i \in \{r+1,\ldots,2^{k+m}\}\}$ be a set of orthonormal eigenvectors with such eigenvalues. It is clear then that $W_x^{\perp} = \mathsf{span}(\{|v_i\rangle : r < i \leq 2^{k+m}\})$.

We consider a pure state $|\psi\rangle = \sum_{i\in[r]} \alpha_i |v_i\rangle$ in W_x. Then,

$$\begin{aligned}
\left\|\Pi_{\mathrm{acc}} V_x (|\psi\rangle \otimes |0^m\rangle)\right\|^2 &= \left\|\Pi_{\mathrm{acc}} V_x \Pi_{\mathrm{init}} (|\psi\rangle \otimes |0^m\rangle)\right\|^2 = \big(\langle\psi| \otimes \langle 0^m|\big)\big(\Pi_x|\psi\rangle \otimes |0^m\rangle\big) \\
&= \left(\langle\psi| \otimes \langle 0^m|\right)\left(\Big(\sum_{i\in[r]} \lambda_i \alpha_i \cdot |v_i\rangle\Big) \otimes |0^m\rangle\right) \\
&= \sum_{i\in[r]} |\alpha_i|^2 \cdot \lambda_i \;\geq\; c(|x|)\,.
\end{aligned}$$

If $|\phi\rangle \in W_x^{\perp}$ is a pure state then by similar arguments we get $\|\Pi_{\mathrm{acc}} V_x (|\phi\rangle \otimes |0^m\rangle)\|^2 \leq w(|x|)$.

When $x \notin L$, condition 2 of Definition 3.2 gets satisfied analogously from condition 2 of Definition 3.4.

Part 2 (Definition 3.2 $\Rightarrow$ Definition 3.4): Let $L \in \mathsf{FewQMA}(c,w,s)$ with some verification procedure $\{V_x\}$ and polynomial q according to Definition 3.2. We claim that $\{V_x\}$ and q satisfy also Definition 3.4 with the same parameters.

First consider the case $x \in L$. The cardinality of the dimension of the subspace of witnesses W_x in $\mathcal{B}^{\otimes k}$ is in $[q(|x|)]$ by hypothesis. We set

$$W_c = \mathsf{span}\{|v\rangle \in \mathcal{B}^{\otimes k} : |v\rangle \otimes |0^m\rangle \text{ is an eigenvector of } \Pi_x \text{ with eigenvalue} \geq\; c(|x|)\} \tag{3.3}$$

and

$$W_w = \mathsf{span}\{|v\rangle \in \mathcal{B}^{\otimes k} : |v\rangle \otimes |0^m\rangle \text{ is an eigenvector of } \Pi_x \text{ with eigenvalue} >\; w(|x|)\}$$

We will show that $\dim(W_x) = \dim(W_c)$ and that $W_c = W_w$, from which the claim follows. For this, it is sufficient to prove that $\dim(W_c) = \dim(W_x) = \dim(W_w)$, since clearly W_c is a subspace of W_w.

First observe that the definitions of W_c and W_w imply that

$$W_c^{\perp} = \mathsf{span}\{|v\rangle \in \mathcal{B}^{\otimes k} : |v\rangle \otimes |0^m\rangle \text{ is an eigenvector of } \Pi_x \text{ with eigenvalue} <\; c(|x|)\}$$

and

$$W_w^{\perp} = \mathsf{span}\{|v\rangle \in \mathcal{B}^{\otimes k} : |v\rangle \otimes |0^m\rangle \text{ is an eigenvector of } \Pi_x \text{ with eigenvalue} \leq\; w(|x|)\}\,.$$

Let us suppose that $\dim(W_x) < \dim(W_c)$. Then there exists a vector $|u\rangle$ in $W_c \cap W_x^{\perp}$. Since $|u\rangle \in W_c$, using arguments as in Part 1 above, we have $\|\Pi_{\text{acc}} V_x(|u\rangle \otimes |0^m\rangle)\|^2 \geq c(|x|)$. However, since $|u\rangle \in W_x^{\perp}$, from condition 1b of Definition 3.2 we have $\|\Pi_{\text{acc}} V_x(|u\rangle \otimes |0^m\rangle)\|^2 \leq w(|x|) < c(|x|)$ which is a contradiction. We similarly reach a contradiction assuming $\dim(W_x) > \dim(W_c)$ and hence $\dim(W_x) = \dim(W_c)$.

The equality $\dim(W_w) = \dim(W_x)$ can be proven by an argument analogous to the proof of $\dim(W_x) = \dim(W_c)$.

In the case $x \notin L$, assume for contradiction that there is an eigenvalue $\lambda > s(|x|)$ of Π_x with eigenvector $|v\rangle \otimes |0^m\rangle$. Then as before,

$$\begin{aligned} \left\|\Pi_{\text{acc}} V_x(|\psi\rangle \otimes |0^m\rangle)\right\|^2 &= \left\|\Pi_{\text{acc}} V_x \Pi_{\text{init}}(|v\rangle \otimes |0^m\rangle)\right\|^2 \\ &= (\langle v| \otimes \langle 0^m|)\Pi_x(|v\rangle \otimes |0^m\rangle) \;=\; \lambda \;>\; s(|x|), \end{aligned}$$

which contradicts condition 2 of Definition 3.2. □

The alternative definition of $\mathsf{FewQMA}(c,w,s)$ is useful in arriving at the following strong error-probability reduction theorem whose proof is the same as the QMA strong error-probability reduction proof in Marriott and Watrous [25] and hence is skipped.

Theorem 3.6 (Error reduction)**.** *Let $c,w,s : \mathbb{N} \to [0,1]$ be polynomial-time computable functions such that for some polynomial p, for all n, they satisfy $c(n) > \max\{w(n),s(n)\} + 1/p(n)$. Let r be any polynomial. Then for any $L \in \mathsf{FewQMA}(c,w,s)$ having a verification procedure with polynomials k,m,q, there exists a verification procedure for L with parameters $(c',w',s') = (1-2^{-r},2^{-r},2^{-r})$ and polynomials $k' = k$, $m' = \mathrm{poly}(m,r)$ and $q' = q$.*

Definition 3.7. A promise problem $L = (L_{\text{yes}}, L_{\text{no}})$ is in the complexity class *Unique Quantum Merlin-Arthur* (denoted UQMA) if L is in FewQMA with the additional constraint that for all $x \in L_{\text{yes}}$, the subspace W_x of witnesses in the definition of FewQMA is one-dimensional.

Definition 3.8. UQMA-CPP is the promise problem $(\text{UQMA-CPP}_{\text{yes}}, \text{UQMA-CPP}_{\text{no}})$ where the elements of $\text{UQMA-CPP}_{\text{yes}} \cup \text{UQMA-CPP}_{\text{no}}$ are descriptions of quantum circuits V with k message qubits and m auxiliary qubits, such that

1. $V \in \text{UQMA-CPP}_{\text{yes}}$ if
   (a) there exists a witness $|\psi\rangle$ such that $\Pr[V$ outputs Accept on $|\psi\rangle] \geq 2/3$,
   (b) for all states $|\phi\rangle$ orthogonal to $|\psi\rangle$, $\Pr[V$ outputs Accept on $|\phi\rangle] \leq 1/3$,
2. $V \in \text{UQMA-CPP}_{\text{no}}$ if for all pure states $|\psi\rangle$, $\Pr[V$ outputs Accept on $|\psi\rangle] \leq 1/3$.

UQMA-CPP can be regarded as the canonical UQMA-complete promise problem.

4 Reducing the dimension of the witness subspace

This section will be entirely devoted to the proof of our main result.

Theorem 4.1. (Main Theorem) $\mathsf{FewQMA} \subseteq \mathsf{P}^{\mathsf{UQMA}}$.

Proof. Let $L \in \mathsf{FewQMA}$ have a verification procedure $\{V'_x : x \in \{0,1\}^*\}$ with polynomials k, m', q. Let r be a polynomial such that $q 2^{-r} \le 1/3$. Then, we know from Theorem 3.6 that

$$L \in \mathsf{FewQMA}(1-2^{-r}, 2^{-r}, 2^{-r})$$

with verification procedure $\{V_x : x \in \{0,1\}^*\}$ and polynomials k, m, q.

Our goal is to describe a deterministic polynomial-time algorithm $\mathcal{A}$, with access to the oracle $\mathcal{O}$ for the promise problem UQMA-CPP, that decides the promise problem L. Broadly, our algorithm works in the following way. On input x and for all $t \in [q(|x|)]$, $\mathcal{A}$ calls $\mathcal{O}$ with a quantum circuit A_x^t that uses $t \cdot k$ message qubits and $t \cdot m$ auxiliary qubits, and outputs Accept if and only if the message has the following two properties: first, it belongs to the alternating subspace of $\mathcal{H}^{\otimes t}$ and second the circuit V_x, when performed on each of the t registers separately, outputs Accept on all of them. $\mathcal{A}$ accepts iff for some t, oracle $\mathcal{O}$ accepts. We will prove that for $x \in L_{\mathrm{yes}}$, we have $A_x^d \in \text{UQMA-CPP}_{\mathrm{yes}}$, where $d = \dim(W_x)$. Hence $\mathcal{O}$ accepts A_x^d and therefore $\mathcal{A}$ accepts. On the other hand, for $x \in L_{\mathrm{no}}$, we show that for all $t \in [q(|x|)], A_x^t \in \text{UQMA-CPP}_{\mathrm{no}}$ and hence $\mathcal{A}$ rejects.

We first describe in detail the Alternating Test and the Witness Test that appear in the algorithm. In our descriptions below k, m, q, r represent the integers $k(|x|), m(|x|), q(|x|)$ and $r(|x|)$ respectively.

4.1 Alternating test

Let $\mathcal{H}$ be the Hilbert space $\mathcal{B}^{\otimes k}$ and let $t \in [2^k]$. Let us fix some polynomial-time computable bijection between the set $[t!]$ and the set of permutations S_t. Let $\mathcal{P}_t$ be the $(t!)$-dimensional Hilbert space spanned by vectors $|i\rangle$, for $i \in [t!]$. We will use the elements of S_t for describing the above basis vectors via the fixed bijection.

The Alternating Test with parameter t receives, as input, a pure state in $\mathcal{H}^{\otimes t}$ and performs a unitary operation in the Hilbert space $\mathcal{P}_t \otimes \mathcal{H}^{\otimes t}$, followed by a measurement. We will refer to the elements of $\mathcal{P}_t \otimes \mathcal{H}^{\otimes t}$ as consisting of two registers R and S, where the content of each register is a mixed state with support over the corresponding Hilbert space.

Let us define

$$|\mathrm{sgn}_t\rangle = \frac{1}{\sqrt{t!}} \sum_{\pi \in S_t} \mathrm{sgn}(\pi) |\pi\rangle \,.$$

Alternating Test(t)
Input: A pure state $|\psi\rangle \in \mathcal{H}^{\otimes t}$ in the $(t \cdot k)$-qubit register S.
Output: The content of S and Accept or Reject.

1. Create the state $\left(\frac{1}{\sqrt{t!}} \sum_{\pi \in S_t} |\pi\rangle \right) \otimes |\psi\rangle$.
2. Apply the unitary $U : |\pi\rangle \otimes |\psi\rangle \to |\pi\rangle \otimes U_\pi |\psi\rangle$.
3. Perform the measurement $(M, I-M)$, where $M = |\mathrm{sgn}_t\rangle\langle \mathrm{sgn}_t| \otimes I_{\mathcal{H}^{\otimes t}}$. Output the content of S. Output Accept if the state has been projected onto M and output Reject otherwise.

It is easily verified that the $\mathsf{Alternating\,Test}(t)$ runs in time polynomial in $t \cdot k$. Since we will only call it with $t \in [q]$, its running time will be polynomial in $|x|$. The following lemma states that the $\mathsf{Alternating\,Test}(t)$ is a projection onto the subspace $\mathsf{Alt}^{\mathcal{H}^{\otimes t}}$.

Lemma 4.2.

1. *For any pure state* $|\psi\rangle \in \mathsf{Alt}^{\mathcal{H}^{\otimes t}}$*, the* $\mathsf{Alternating\,Test}(t)$ *outputs the state* $|\psi\rangle$ *and* Accept *with probability 1.*

2. *For any* $|\phi\rangle \in \left(\mathsf{Alt}^{\mathcal{H}^{\otimes t}}\right)^{\perp}$*, the* $\mathsf{Alternating\,Test}(t)$ *outputs* Accept *with probability 0.*

Proof. **Part 1**: Since $|\psi\rangle \in \mathsf{Alt}^{\mathcal{H}^{\otimes t}}$ we have $U_\pi|\psi\rangle = \operatorname{sgn}(\pi)\cdot|\psi\rangle$, and therefore the state after Step 2 is

$$\frac{1}{\sqrt{t!}}\sum_{\pi\in S_t}|\pi\rangle\otimes U_\pi|\psi\rangle \;=\; \frac{1}{\sqrt{t!}}\sum_{\pi\in S_t}\operatorname{sgn}(\pi)\cdot|\pi\rangle\otimes|\psi\rangle \;=\; |\mathrm{sgn}_t\rangle\otimes|\psi\rangle\,.$$

Part 2: By Claim 2.2, we have

$$\left(\mathsf{Alt}^{\mathcal{H}^{\otimes t}}\right)^{\perp} \;=\; \sum_{i\neq j}\mathsf{Sym}_{ij}^{\mathcal{H}^{\otimes t}}\,.$$

Hence it is enough to show that for all distinct $i,j\in[t]$ and for any vector $|\phi\rangle\in\mathsf{Sym}_{ij}^{\mathcal{H}^{\otimes t}}$, the probability p that the $\mathsf{Alternating\,Test}(t)$ outputs Accept is 0. We have

$$\begin{aligned} p \;&=\; \left(\frac{1}{\sqrt{t!}}\sum_{\sigma\in S_t}\langle\sigma|\otimes\langle\phi|U_\sigma^\dagger\right)\left(|\mathrm{sgn}_t\rangle\langle\mathrm{sgn}_t|\otimes I_{\mathcal{H}^{\otimes t}}\right)\left(\frac{1}{\sqrt{t!}}\sum_{\pi\in S_t}|\pi\rangle\otimes U_\pi|\phi\rangle\right)\\ &=\; \left(\frac{1}{t!}\right)^2\sum_{\sigma,\pi\in S_t}\operatorname{sgn}(\sigma)\cdot\operatorname{sgn}(\pi)\cdot\langle\phi|U_\sigma^\dagger U_\pi|\phi\rangle\,. \end{aligned}$$

We define $\pi'=\pi\circ\pi_{ij}$. Then $\pi=\pi'\circ\pi_{ij}^{-1}=\pi'\circ\pi_{ij}$ and $\operatorname{sgn}(\pi)=-\operatorname{sgn}(\pi')$. Since $|\phi\rangle\in\mathsf{Sym}_{ij}^{\mathcal{H}^{\otimes t}}$, we have $U_{\pi_{ij}}|\phi\rangle=|\phi\rangle$ and hence $U_{\pi'\circ\pi_{ij}}|\phi\rangle=U_{\pi'}\cdot(U_{\pi_{ij}}|\phi\rangle)=U_{\pi'}|\phi\rangle$. Therefore,

$$\begin{aligned} p \;&=\; \left(\frac{1}{t!}\right)^2\sum_{\sigma,\pi\in S_t}\operatorname{sgn}(\sigma)\cdot\operatorname{sgn}(\pi)\cdot\langle\phi|U_\sigma^\dagger U_\pi|\phi\rangle\\ &=\; \frac{-1}{(t!)^2}\sum_{\sigma,\pi'\in S_t}\operatorname{sgn}(\sigma)\cdot\operatorname{sgn}(\pi')\cdot\langle\phi|U_\sigma^\dagger U_{\pi'\circ\pi_{ij}}|\phi\rangle\\ &=\; \frac{-1}{(t!)^2}\sum_{\sigma,\pi'\in S_t}\operatorname{sgn}(\sigma)\cdot\operatorname{sgn}(\pi')\cdot\langle\phi|U_\sigma^\dagger U_{\pi'}|\phi\rangle\\ &=\; -p\,. \end{aligned}$$

Hence $p=0$. □

4.2 Witness test

The $\mathsf{Witness\,Test}$ with parameter $t \in [q]$ receives as input a pure state in $\mathcal{H}^{\otimes t}$ and performs a unitary operation in the Hilbert space $\mathcal{H}^{\otimes t} \otimes \mathcal{B}^{\otimes (tm)}$ followed by a measurement. We will refer to the elements of $\mathcal{H}^{\otimes t} \otimes \mathcal{B}^{\otimes (tm)}$ as consisting of t pairs of registers (T_i, Z_i) respectively on k and m qubits, for $i \in [t]$. All registers Z_i will be initialized to $|0^m\rangle$.

> $\mathsf{Witness\,Test}(t)$
>
> Input: A pure state $|\psi\rangle \in \mathcal{H}^{\otimes t}$ in the k-qubit registers T_i, for $i \in [t]$.
> Output: Accept or Reject.
>
> 1. For all $i \in [t]$, append a register Z_i initialized to $|0^m\rangle$ and apply the circuit V_x on registers (T_i, Z_i).
> 2. Output Accept if for all $i \in [t]$, V_x outputs Accept; otherwise output Reject.

We can describe the $\mathsf{Witness\,Test}(t)$ as the operator $(\Pi_{\text{acc}} V_x)^{\otimes t}$ acting on a state[2] $|\psi\rangle \otimes |0^{tm}\rangle$. Hence, $\Pr[\mathsf{Witness\,Test}(t) \text{ outputs Accept on } |\psi\rangle] = \|(\Pi_{\text{acc}} V_x)^{\otimes t}(|\psi\rangle \otimes |0^{tm}\rangle)\|^2$. Note that the description of the circuit $V_x^{\otimes t}$ can be generated in polynomial-time, since the circuit family $\{V_x, x \in \{0,1\}^*\}$ is uniformly generated in polynomial-time.

In what follows we will have to argue about the probability that the verification procedure V_x outputs Accept when its input is some mixed state. Even though we have only considered pure states as inputs in the definition of the class FewQMA, we will see that it is not hard to extend our arguments to mixed states.

Lemma 4.3.

1. *If $x \in L_{\text{yes}}$, then for every $|\psi\rangle \in W_x^{\otimes t}$, the* $\mathsf{Witness\,Test}(t)$ *outputs* Accept *with probability at least* $2/3$.
2. *If $x \in L_{\text{yes}}$, then for every $|\phi\rangle \in (W_x^{\otimes t})^{\perp}$, the* $\mathsf{Witness\,Test}(t)$ *outputs* Accept *with probability at most* $1/3$.
3. *If $x \in L_{\text{no}}$, then for every $|\psi\rangle \in \mathcal{H}^{\otimes t}$, the* $\mathsf{Witness\,Test}(t)$ *outputs* Accept *with probability at most* $1/3$.

Proof. **Part 1**: By completeness we know that for any pure state $|\psi'\rangle \in W_x$, we have

$$\Pr[V_x \text{ outputs Reject on } |\psi'\rangle] \leq 2^{-r}.$$

Let ρ_i denote the reduced density matrix of $|\psi\rangle$ on register T_i. Since $|\psi\rangle \in W_x{}^{\otimes t}$, then for every $i \in [t]$, the density matrix ρ_i is a distribution of pure states that all belong to W_x and hence

$$\Pr[V_x \text{ outputs Reject on } \rho_i] \leq 2^{-r}.$$

[2] More accurately, the operator $(\Pi_{\text{acc}} V_x)^{\otimes t}$ acts on a reordered version of $|\psi\rangle \otimes |0^{tm}\rangle$, but this will be ignored for the sake of clarity.

It follows from the union bound that

$$\begin{aligned}\mathsf{Pr}[\mathsf{Witness\,Test}(t) \text{ outputs Accept on } |\psi\rangle] &\geq 1 - \sum_{i=1}^{t} \mathsf{Pr}[V_x \text{ outputs Reject on } \rho_i] \\ &\geq 1 - t \cdot 2^{-r} \geq 2/3\,,\end{aligned}$$

where the last inequality follows from the choice of r.

Part 2: For $i \in [t]$, let

$$S_i = W_x \otimes \cdots \otimes W_x \otimes W_x^{\perp} \otimes \mathcal{H} \otimes \cdots \otimes \mathcal{H}\,,$$

where $W_x^{\perp}$ stands in the i^{th} component of the tensor product. By Fact 2.1 we have $(W_x{}^{\otimes t})^{\perp} = \bigoplus_{i\in[t]} S_i$.

Let us therefore consider a pure state $|\phi\rangle \in \bigoplus_{i\in[t]} S_i$ and let $|\phi\rangle = \sum_{i\in[t]} a_i |\phi_i\rangle$, where $|\phi_i\rangle$ is a pure state in S_i. Then $\sum_{i=1}^{t} |a_i|^2 = 1$ because the $|\phi_i\rangle$'s are also orthogonal. Furthermore, let ρ_i be the reduced density matrix of $|\phi_i\rangle$ on register T_i. Then the support of ρ_i is over $W_x^{\perp}$. Since for any pure state $|\phi'\rangle \in W_x^{\perp}$ we have $\mathsf{Pr}[V_x \text{ outputs Accept on } |\phi'\rangle] \leq 2^{-r}$, we conclude that $\mathsf{Pr}[V_x \text{ outputs Accept on } \rho_i] \leq 2^{-r}$.

The probability that the $\mathsf{Witness\,Test}(t)$ outputs Accept on $|\phi_i\rangle$ is equal to the probability that all t applications of V_x output Accept, which is less than the probability that the i^{th} application of V_x outputs Accept, since the projections, performed in different registers, commute. Hence,

$$\mathsf{Pr}[\mathsf{Witness\,Test}(t) \text{ outputs Accept on } |\phi_i\rangle] \leq \mathsf{Pr}[V_x \text{ outputs Accept on } \rho_i] \leq 2^{-r}.$$

Now for the input $|\phi\rangle$ we have

$$\begin{aligned}\mathsf{Pr}[\mathsf{Witness\,Test}(t) \text{ outputs Accept on } |\phi\rangle] &= \left\|(\Pi_{\text{acc}} V_x)^{\otimes t}(|\phi\rangle \otimes |0^{tm}\rangle)\right\|^2 \\ &= \Big\|\sum_{i\in[t]} a_i (\Pi_{\text{acc}} V_x)^{\otimes t}(|\phi_i\rangle \otimes |0^{tm}\rangle)\Big\|^2 \\ &\leq \left(\sum_{i\in[t]} |a_i| \cdot \left\|(\Pi_{\text{acc}} V_x)^{\otimes t}(|\phi_i\rangle \otimes |0^{tm}\rangle)\right\|\right)^2 \\ &\leq \left(\sum_{i\in[t]} |a_i|^2\right) \cdot \left(\sum_{i\in[t]} \left\|(\Pi_{\text{acc}} V_x)^{\otimes t}(|\phi_i\rangle \otimes |0^{tm}\rangle)\right\|^2\right) \\ &\leq t \cdot 2^{-r} \leq 1/3\,.\end{aligned}$$

In the above calculation the first two inequalities follow respectively from the triangle inequality and the Cauchy-Schwarz inequality.

Part 3: By the soundness of the original protocol we know that for any pure state $|\psi'\rangle \in \mathcal{H}$,

$$\mathsf{Pr}[V_x \text{ outputs Accept on } |\psi'\rangle] \leq 2^{-r}.$$

The same holds for any mixed state as well. Since the probability that the $\mathsf{Witness\,Test}(t)$ outputs Accept is at most the probability that the procedure V_x accepts the state on the first register T_1 we conclude that

$$\mathsf{Pr}[\mathsf{Witness\,Test}(t) \text{ outputs Accept on } |\psi\rangle] \leq \mathsf{Pr}[V_x \text{ outputs Accept on } \rho_1] \leq 2^{-r} \leq 1/3\,. \qquad \square$$

4.3 Putting it all together

Finally we describe the algorithm $\mathcal{A}$ in the figure below and proceed to analyze its properties.

Running time We have seen that the description of the circuit that performs the $\mathsf{Alternating\,Test}(t)$ can be generated in time polynomial in $t \cdot k$ which is polynomial in $|x|$. The description of the circuit that performs the $\mathsf{Witness\,Test}(t)$ can also be generated in polynomial-time, since the circuit family $\{V_x : x \in \{0,1\}^*\}$ can be generated uniformly in polynomial-time. Hence the description of the circuit A_x^t can be generated in polynomial-time and the overall algorithm $\mathcal{A}$ runs in polynomial-time.

Algorithm $\mathcal{A}$

Input: $x \in L_{\text{yes}} \cup L_{\text{no}}$.

Output: Accept or Reject.

1. For $t = 1, \ldots, q(|x|)$ do:
   (a) Call the oracle $\mathcal{O}$ with input A_x^t, where A_x^t is the description of the circuit of the following procedure on $t \cdot k$ message qubits and $t \cdot m$ auxiliary qubits

   Input: A pure state $|\psi\rangle \in \mathcal{H}^{\otimes t}$; Output: Accept or Reject.

      i. Run the $\mathsf{Alternating\,Test}(t)$ with input $|\psi\rangle$.

      ii. Run the $\mathsf{Witness\,Test}(t)$ with input being the output state of the $\mathsf{Alternating\,Test}(t)$.

      iii. Output Accept iff both Tests output Accept.

   (b) If $\mathcal{O}$ outputs Accept then output Accept and halt.
2. Output Reject.

Correctness in case $x \in L_{\text{yes}}$ Let us consider the oracle call with input A_x^d where d is the dimension of W_x. We prove that $A_x^d \in \mathsf{UQMA\text{-}CPP}_{\text{yes}}$, hence the oracle $\mathcal{O}$ outputs Accept and therefore $\mathcal{A}$ outputs Accept as well. Our claim is immediate from the following lemma.

Lemma 4.4.

1. $\Pr[A_x^d$ *outputs* Accept *on* $|W_{\text{alt}}\rangle] \geq 2/3$, *where* $|W_{\text{alt}}\rangle$ *is defined in equation* (2.1).
2. *Let* $|\phi\rangle \in \mathcal{H}^{\otimes d}$ *be orthogonal to* $|W_{\text{alt}}\rangle$. *Then* $\Pr[A_x^d$ *outputs* Accept *on* $|\phi\rangle] \leq 1/3$.

Proof. **Part (1)**: By Claim 2.3, $|W_{\text{alt}}\rangle \in \mathsf{Alt}^{\mathcal{H}^{\otimes d}}$ and Lemma 4.2 tells us that the $\mathsf{Alternating\,Test}(d)$ outputs the state $|W_{\text{alt}}\rangle$ and Accept with probability 1. Then, since for every $i \in [d]$ the support of the reduced density matrix of $|W_{\text{alt}}\rangle$ on register T_i is on W_x, Lemma 4.3 tells us that the $\mathsf{Witness\,Test}$ outputs Accept with probability at least $2/3$.

Part (2): By Claim 2.5 and the fact that the state $|\phi\rangle$ is orthogonal to $|W_{\text{alt}}\rangle$, we can conclude that if the Alternating Test(d) outputs Accept then the output state is a pure state $|\phi'\rangle \in (W^{\otimes d})^{\perp}$. Now, by Lemma 4.3, the probability that the Witness Test(d) outputs Accept on input $|\phi'\rangle$ is at most $1/3$. □

Correctness in case $x \in L_{\text{no}}$ By Lemma 4.3 it follows easily that for all $t \in [q]$: $A_x^t \in$ UQMA-CPP$_{\text{no}}$. In this case, $\mathcal{O}$ outputs Reject in every iteration, and hence $\mathcal{A}$ outputs Reject.

This concludes the proof of Theorem 4.1. □

4.4 Implications to gapped Hamiltonians

The Local Hamiltonian is a QMA-complete problem defined as follows.

Definition 4.5 (k-local Hamiltonian)**.** A Hermitian H acting on n qubits is a k-local Hamiltonian if $H = \sum_{i=1}^{m} H_i$, where $m \le \text{poly}(n)$, and for all i, H_i acts non-trivially on at most k qubits, and $\|H_i\| \le \text{poly}(n)$ where $\|\cdot\|$ is the operator norm.

Definition 4.6 (k-Local Hamiltonian promise problem)**.** The input contains a k-local Hamiltonian H, and two real numbers a and b such that $b - a > 1/\text{poly}(n)$. In a YES instance, the smallest eigenvalue of H is at most a, and in a NO instance, all eigenvalues of H are at least b.

The gap between the two lowest eigenvalues of the Hamiltonian plays an important role. For example, Hastings [13] has shown that under certain conditions on the Hamiltonian and the gap, the ground state has an efficient classical description. Therefore, a certain variant of the Local Hamiltonian problem is in NP (which implies that this variant of the problem is not QMA-complete, under the assumption that NP $\neq$ QMA).

Also, in the adiabatic computation paradigm [10], the gap governs the running time. Therefore, it is natural to define a variant of the Local Hamiltonian problem with an additional promise on the gap.

Definition 4.7 (Unique k-Local Hamiltonian promise problem)**.** The input is the same as that of the k-Local Hamiltonian problem. In a YES instance, there is a unique eigenvector of H with eigenvalue at most a, and no eigenvalues in the interval (a,b), and in a NO instance, all the eigenvalues of H are at least b.

Since we are interested in the relationship between a unique quantum witness, and polynomially many quantum witnesses, it is natural to define the following problem:

Definition 4.8 (Few k-Local Hamiltonian promise problem)**.** The input is the same as that of the k-Local Hamiltonian problem, together with a function $q(n) < \text{poly}(n)$. In a YES instance, there are at most q orthogonal eigenvectors of H with eigenvalues at most a, and no eigenvalues in the interval (a,b), where in a NO instance, all eigenvalues of H are at least b.

Note that Unique k-Local Hamiltonian is Karp-reducible to Few k-Local Hamiltonian: every Unique k-Local Hamiltonian instance is also a Few k-Local Hamiltonian instance with $q = 1$. Now we show that these two problems are actually equivalent.

Theorem 4.9. Few k-Local Hamiltonian *is Cook-reducible to* Unique k-Local Hamiltonian *in polynomial time.*

Proof. Aharonov et al. [1] observed that known proofs [19, 27, 18, 2] of the fact that k-Local Hamiltonian is QMA-hard, preserve the dimension of the satisfying subspace. Using this observation, they have shown that Unique k-Local Hamiltonian is UQMA-complete. The same argument also shows that Few k-Local Hamiltonian is FewQMA-complete.

Since Few k-Local Hamiltonian $\in$ FewQMA, Theorem 4.1 shows that Few k-Local Hamiltonian is Cook-reducible to UQMA-CPP in polynomial time. Combining this with the above facts shows that Few k-Local Hamiltonian is Cook-reducible to Unique k-Local Hamiltonian in polynomial time. $\square$

5 Yet another definition of FewQMA

We have seen two definitions for FewQMA and have proven their equivalence. In high level, one says that in the yes instances there is a polynomial number of eigenvalues of the projection operator Π_x that are larger than $2/3$, while no eigenvalue is in the interval $(1/3, 2/3)$. The other definition says that in a yes instance there exists a subspace of polynomial dimension such that every state in the subspace is accepted with probability at least $2/3$ and every state orthogonal to this subspace is accepted with probability at most $1/3$.

A natural question is whether the use of a subspace is necessary in the second definition or we could have just talked about a set of orthonormal vectors of polynomial size, where each vector is accepted with probability $2/3$ and every vector orthogonal to these ones is accepted with probability at most $1/3$. While we are unable to show the equivalence of the complexity class defined this way and FewQMA, a weak equivalence can indeed be shown. For this we include the parameters of the verification procedure and the bound on the number of witnesses in the definition of the class, and we also require strong amplification. More precisely, consider the following definition.

Definition 5.1. Let $c, w, s : \mathbb{N} \to [0,1]$ be polynomial-time computable functions such that $c(n) > \max\{w(n), s(n)\}$ for all $n \in \mathbb{N}$. Let q, k, m be polynomials such that $q(n) \leq 2^{k(n)}$ for all $n \in \mathbb{N}$. A promise problem $L = (L_{\text{yes}}, L_{\text{no}})$ is in the complexity class *Vector Few Quantum Merlin-Arthur* (denoted Vector-FewQMA(c, w, s, q, k, m)) if there exists a verification procedure $\{V_x : x \in \{0,1\}^*\}$ with k witness qubits and m ancilla qubits, such that

1. for all $x \in L_{\text{yes}}$ there exists an orthonormal basis $\{|\psi_1\rangle, \ldots, |\psi_{2^k}\rangle\}$ of the witness space and $d \in [q(|x|)]$ such that
   (a) for all pure states $|\psi_i\rangle$ with $i \in [d]$, $\left\|\Pi_{\text{acc}} V_x(|\psi_i\rangle \otimes |0^m\rangle)\right\|^2 \geq c(|x|)$,
   (b) for all pure states $|\psi_i\rangle$ with $d+1 \leq i \leq 2^k$, $\left\|\Pi_{\text{acc}} V_x(|\psi_i\rangle \otimes |0^m\rangle)\right\|^2 \leq w(|x|)$,
2. for all $x \in L_{\text{no}}$ and for all pure states $|\psi\rangle \in \mathcal{B}^{\otimes k}$, $\left\|\Pi_{\text{acc}} V_x(|\psi\rangle \otimes |0^m\rangle)\right\|^2 \leq s(|x|)$.

Finally we define

$$\text{Vector-FewQMA} = \bigcup_{\substack{q,k,m \\ q \leq 2^k}} \text{Vector-FewQMA}\left(1 - \frac{1}{3q}, \frac{1}{3 \cdot 2^k}, \frac{1}{3}, q, k, m\right).$$

We show Vector-FewQMA $=$ FewQMA by using Horn's Theorem that states that for a Hermitian matrix, the vector of the eigenvalues majorizes the diagonal.

Theorem 5.2 ([15]). *Let R be a natural number. Let $\Lambda = \{\lambda_i\}_{i=1}^R$ and $A = \{\mu_i\}_{i=1}^R$ where*

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_R \quad \text{and} \quad \mu_1 \geq \mu_2 \geq \ldots \geq \mu_R .$$

Then there exists an $R \times R$ Hermitian matrix with set of eigenvalues Λ and set of diagonal elements A if and only if $\sum_{i=1}^t (\lambda_i - \mu_i) \geq 0$ for all $t \in [R]$ and with equality for $t = R$.

Theorem 5.3. FewQMA $=$ Vector-FewQMA.

Proof. We first show FewQMA $\subseteq$ Vector-FewQMA. Let $L \in$ FewQMA have a verification procedure $\{V'_x : x \in \{0,1\}^*\}$ with polynomials k, m', q. Let r be a polynomial such that $2^{-r} \leq 1/(3 \cdot 2^k)$ and $2^{-r} \leq 1/(3q)$. We know from Theorem 3.6 that $L \in \mathsf{FewQMA}(1-2^{-r}, 2^{-r}, 2^{-r})$ with verification procedure $\{V_x : x \in \{0,1\}^*\}$ and polynomials k, m, q, where $m = \mathrm{poly}(m', r)$. Then the eigenbasis of Π_x (see equation (3.1)) satisfies the requirements of the definition of

$$\mathsf{Vector\text{-}FewQMA}\left(1 - \frac{1}{3q}, \frac{1}{3 \cdot 2^k}, \frac{1}{3}, q, k, m\right).$$

This shows that $L \in$ Vector-FewQMA.

We now show Vector-FewQMA $\subseteq$ FewQMA. Let

$$L \in \mathsf{Vector\text{-}FewQMA}\left(1 - \frac{1}{3q}, \frac{1}{3 \cdot 2^k}, \frac{1}{3}, q, k, m\right),$$

for some polynomials q, k, m such that $q \leq 2^k$. Let $N = 2^k$ and $\mathcal{H} = \mathcal{B}^{\otimes k}$. If $x \in L_{\text{yes}}$, then there exist an orthonormal basis $\{|\psi_1\rangle, \ldots, |\psi_N\rangle\}$ for $\mathcal{H}$ and $d \in [q]$, such that for $i \in [d]$, $\mu_i \geq 1 - 1/(3q)$ and for $d+1 \leq i \leq N$, $\mu_i \leq 1/(3N)$, where by definition

$$\mu_i \stackrel{\text{def}}{=} \left\| \Pi_{\text{acc}} V_x(|\psi_i\rangle \otimes |0^m\rangle) \right\|^2 \;=\; \langle \psi_i| \otimes \langle 0^m | \Pi_x | \psi_i \rangle \otimes |0^m\rangle .$$

Consider now the Hermitian matrix M that describes the projection operator Π_x in a basis that is an extension of $\{|\psi_1\rangle \otimes |0^m\rangle, \ldots, |\psi_N\rangle \otimes |0^m\rangle\}$. Note that $\mu_i, i \in [N]$ are the first N diagonal elements of M. Observe that an eigenvector of Π_x with non-zero eigenvalue is also an eigenvector of Π_{init} with non-zero eigenvalue. Since there are N non-zero eigenvalues of Π_{init} (counting with multiplicity), there are at most N non-zero eigenvalues of Π_x. This also implies that $\mu_i = 0$ for $N < i \leq 2^{k+m}$. Let the first N eigenvalues of Π_x in decreasing order be $\lambda_i, i \in [N]$. Then, using Horn's theorem,

$$\sum_{i=1}^d \lambda_i \;\geq\; \sum_{i=1}^d \mu_i \;\geq\; d \cdot \left(1 - \frac{1}{3q}\right)$$

which implies (since $\lambda_i \leq 1$, for all $i \in [N]$) that

$$\lambda_d \;\geq\; -(d-1) + d \cdot (1 - \frac{1}{3q}) \;\geq\; 1 - \frac{d}{3q} \;\geq\; \frac{2}{3}.$$

Also, we have that

$$\lambda_{d+1} \leq \sum_{i=d+1}^{2^{k+m}} \lambda_i \leq \sum_{i=d+1}^{2^{k+m}} \mu_i = \sum_{i=d+1}^{2^k} \mu_i \leq (N-d)\frac{1}{3N} \leq \frac{1}{3}.$$

If $x \in L_{\text{no}}$ then by the soundness condition $\lambda_1 \leq 1/3$. This shows that $L \in \mathsf{FewQMA}$. □

Acknowledgments

Most of this work was conducted when R.J., I.K., M.S. and S.Z. were at the Centre for Quantum Technologies (CQT) in Singapore, and partially funded by the Singapore Ministry of Education and the National Research Foundation. Research partially supported by European Commission IST project Quantum Computer Science (QCS) 25596, by the CHIST-ERA project DIQIP, by the French ANR programs under contract ANR-08-EMER-012 (QRAC project) and ANR-09-JCJC-0067-01 (CRYQ project), by the French MAEE STIC-Asie program FQIC, by NSF grant DMS-0606795 and by Research Grants Council of the Hong Kong S.A.R. (Project no. CUHK419309, CUHK418710, CUHK419011). R.J., I.K., M.S. and S.Z. would like to thank Hartmut Klauck for several insightful discussions during the process of this work. G.K. and O.S. would like to thank Scott Aaronson and Dorit Aharonov for discussions.

References

[1] DORIT AHARONOV, MICHAEL BEN-OR, FERNANDO G. S. L. BRANDÃO, AND OR SATTATH: The pursuit for uniqueness: extending Valiant-Vazirani theorem to the probabilistic and quantum settings, 2008. arXiv:0810.4840. 376, 377, 393

[2] DORIT AHARONOV, DANIEL GOTTESMAN, SANDY IRANI, AND JULIA KEMPE: The power of quantum systems on a line. *Comm. Math. Phys.*, 287(1):41–65, 2009. Preliminary version in FOCS'07. [doi:10.1007/s00220-008-0710-3] 393

[3] DORIT AHARONOV AND TOMER NAVEH: Quantum NP - A survey, 2002. arXiv:quant-ph/0210077. 376

[4] ERIC W. ALLENDER: The complexity of sparse sets in P (preliminary report). In *Proc. 1st Structure in Complexity Theory Conf. (1986)*, volume 223 of *Lecture Notes in Comput. Sci.*, pp. 1–11. Springer, 1986. [doi:10.1007/3-540-16486-3_85] 377

[5] ERIC W. ALLENDER AND ROY S. RUBINSTEIN: P-printable sets. *SIAM J. Comput.*, 17(6):1193–1202, 1988. [doi:10.1137/0217075] 378

[6] LÁSZLÓ BABAI: Trading group theory for randomness. In *Proc. 17th STOC*, pp. 421–429. ACM Press, 1985. [doi:10.1145/22145.22192] 376

[7] ADRIANO BARENCO, ANDRÉ BERTHIAUME, DAVID DEUTSCH, ARTUR EKERT, RICHARD JOZSA, AND CHIARA MACCHIAVELLO: Stabilization of quantum computations by symmetrization. *SIAM J. Comput.*, 26(5):1541–1557, 1997. [doi:10.1137/S0097539796302452] 379

[8] RAJENDRA BHATIA: *Matrix Analysis*. Volume 169 of *Graduate Texts in Mathematics*. Springer, 1997. Springer. 381

[9] HARRY BUHRMAN, RICHARD CLEVE, JOHN WATROUS, AND RONALD DE WOLF: Quantum fingerprinting. *Phys. Rev. Lett.*, 87(16), 2001. arXiv:quant-ph/0102001. [doi:10.1103/PhysRevLett.87.167902] 379

[10] EDWARD FARHI, JEFFREY GOLDSTONE, SAM GUTMANN, AND MICHAEL SIPSER: Quantum computation by adiabatic evolution. 2000. arXiv:quant-ph/0001106. 392

[11] SHAFI GOLDWASSER, SILVIO MICALI, AND CHARLES RACKOFF: The knowledge complexity of interactive proof systems. *SIAM J. Comput.*, 18(1):186–208, 1989. Preliminary version in STOC'85. [doi:10.1137/0218012] 376

[12] JOACHIM GROLLMANN AND ALAN L. SELMAN: Complexity measures for public-key cryptosystems. *SIAM J. Comput.*, 17(2):309–335, 1988. Preliminary version in FOCS'84. [doi:10.1137/0217018] 376

[13] MATTHEW B. HASTINGS: An area law for one-dimensional quantum systems. *J. Statistical Mechanics: Theory and Experiment*, 2007(8):P08024, 2007. arXiv:0705.2024. [doi:10.1088/1742-5468/2007/08/P08024] 392

[14] LANE A. HEMASPAANDRA, SANJAY JAIN, AND NIKOLAI K. VERESHCHAGIN: Banishing robust Turing completeness. *Internat. J. Found. Comput. Sci.*, 4(3):245–265, 1993. Preliminary version in LFCS'92. [doi:10.1142/S012905419300016X] 378

[15] ALFRED HORN: Doubly stochastic matrices and the diagonal of a rotation matrix. *Amer. J. Math.*, 76(3):620–630, 1954. [doi:10.2307/2372705] 394

[16] DOMINIK JANZING, PAVEL WOCJAN, AND THOMAS BETH: Identity check is QMA-complete, 2003. arXiv:quant-ph/0305050. 376

[17] ALASTAIR KAY: Quantum-Merlin-Arthur-complete translationally invariant Hamiltonian problem and the complexity of finding ground-state energies in physical systems. *Phys. Rev. A*, 76(3):030307, 2007. arXiv:0704.3142. [doi:10.1103/PhysRevA.76.030307] 376

[18] JULIA KEMPE, ALEXEI KITAEV, AND ODED REGEV: The complexity of the Local Hamiltonian problem. *SIAM J. Comput.*, 35(5):1070–1097, 2006. Preliminary version in FSTTCS'04. [doi:10.1137/S0097539704445226] 376, 393

[19] ALEXEI KITAEV, ALEXANDER SHEN, AND MIKHAIL VYALYI: *Classical and Quantum Computation*. Volume 47 of *Graduate Studies in Mathematics*. Amer. Math. Soc., 2002. Translated from the 1999 Russian original by Lester J. Senechal, AMS. 376, 393

[20] EMANUEL KNILL: Quantum randomness and nondeterminism, 1996. arXiv:quant-ph/9610012. 376

[21] KER-I KO: On some natural complete operators. *Theoret. Comput. Sci.*, 37(1):1–30, 1985. [doi:10.1016/0304-3975(85)90085-4] 376

[22] JOHANNES KÖBLER, UWE SCHÖNING, SEINOSUKE TODA, AND JACOBO TORÁN: Turing machines with few accepting computations and low sets for PP. *J. Comput. System Sci.*, 44(2):272–286, 1992. Preliminary version in Structure In Complexity Theory, 1989. [doi:10.1016/0022-0000(92)90022-B] 378

[23] YI-KAI LIU: Consistency of local density matrices is QMA-complete. In *Proc. 10th Internat. Workshop on Randomization and Computation (RANDOM'06)*, volume 4110 of *Lecture Notes in Comput. Sci.*, pp. 438–449. Springer, 2006. arXiv:quant-ph/0604166. [doi:10.1007/11830924_40] 376

[24] YI-KAI LIU, MATTHIAS CHRISTANDL, AND FRANK VERSTRAETE: Quantum computational complexity of the N-representability problem: QMA-complete. *Phys. Rev. Lett.*, 98(11), 2007. arXiv:quant-ph/0609125. [doi:10.1103/PhysRevLett.98.110503] 376

[25] CHRIS MARRIOTT AND JOHN WATROUS: Quantum Arthur-Merlin games. *Comput. Complexity*, 14(2):122–152, 2005. Preliminary version in CCC'04. [doi:10.1007/s00037-005-0194-x] 386

[26] MICHAEL A. NIELSEN AND ISAAC L. CHUANG: *Quantum Computation and Quantum Information*. Cambridge University Press, 2000. [doi:10.1017/CBO9780511976667] 380

[27] ROBERTO OLIVEIRA AND BARBARA TERHAL: The complexity of quantum spin systems on a two-dimensional square lattice. *Quantum Inf. Comput.*, 8(10):900–924, 2008. arXiv:quant-ph/0504050. [ACM:2016987] 393

[28] RAJESH P. N. RAO, JÖRG ROTHE, AND OSAMU WATANABE: Upward separation for FewP and related classes. *Inform. Process. Lett.*, 52(4):175 – 180, 1994. [doi:10.1016/0020-0190(94)90123-6] 378

[29] OR SATTATH: The pursuit of uniqueness: Extending Valiant-Vazirani theorem to the probabilistic and quantum settings. Master's thesis, Hebrew University in Jerusalem, 2009. 379

[30] JACOBO TORÁN: Counting the number of solutions. A survey of recent inclusion results in the area of counting classes. In *Mathematical Foundations of Computer Science (MFCS'90)*, volume 452 of *Lecture Notes in Comput. Sci.*, pp. 121–134. Springer, 1990. [doi:10.1007/BFb0029600] 378

[31] LESLIE G. VALIANT AND VIJAY V. VAZIRANI: NP is as easy as detecting unique solutions. *Theoret. Comput. Sci.*, 47(1):85–93, 1986. Preliminary version in STOC'85. [doi:10.1016/0304-3975(86)90135-0] 376

[32] JOHN WATROUS: Succinct quantum proofs for properties of finite groups. In *Proc. 41st FOCS*, pp. 537–546. IEEE Comp. Soc. Press, 2000. arXiv:cs/0009002. [doi:10.1109/SFCS.2000.892141] 376

AUTHORS

Rahul Jain [About the author]
Assistant professor
National University of Singapore
rahul@comp.nus.edu.sg
http://comp.nus.edu.sg/~rahul

Iordanis Kerenidis [About the author]
Senior Researcher
Université Paris Diderot 7, Paris, France
jkeren@liafa.univ-paris-diderot.fr
http://www.liafa.jussieu.fr/~jkeren

Greg Kuperberg [About the author]
Professor
UC Davis, Davis, CA
greg@math.ucdavis.edu
http://www.math.ucdavis.edu/~greg

Miklos Santha [About the author]
Directeur de Recherche
Université Paris Diderot
miklos.santha@lri.fr
http://www.liafa.univ-paris-diderot.fr/~santha

Or Sattath [About the author]
Ph. D. candidate
The Hebrew University, Jerusalem, Israel
sattath@cs.huji.ac.il
http://www.cs.huji.ac.il/~sattath

Shengyu Zhang [About the author]
Assistant professor
The Chinese University of Hong Kong
syzhang@cse.cuhk.edu.hk
http://www.cse.cuhk.edu.hk/~syzhang

ABOUT THE AUTHORS

RAHUL JAIN obtained his Ph. D. in computer science from the Tata Institute of Fundamental Research, Mumbai, India in 2003. His Ph. D. advisor was Jaikumar Radhakrishnan. He was a postdoctoral fellow for two years at the University of California, Berkeley (2004-2006) and for two years at the Institute for Quantum Computing (IQC), University of Waterloo, Canada (2006-2008). In 2008, he joined NUS as an Assistant Professor in the Computer Science Department with a cross appointment with CQT. His research interests are in the areas of information theory, quantum computation, cryptography, communication complexity, and computational complexity theory.

IORDANIS KERENIDIS received his Ph. D. in 2004 from the Computer Science department at the University of California, Berkeley. His advisor was Umesh Vazirani. After a two-year postdoctoral position at the Massachusetts Institute of Technology, he moved to France, where he now holds a Senior Researcher CNRS position, based at the Université Paris-Diderot. Since 2009, he has also been a long-term visiting scholar at the Centre for Quantum Technologies, Singapore. His research interests lie in the intersection of quantum cryptography and complexity theory.

GREG KUPERBERG received a bachelor's degree from Harvard University (1987) and a Ph. D. in geometric topology and quantum algebra from University of California, Berkeley (1991). His advisor was Andrew Casson. Both of his parents are mathematicians, and every subset of the three have authored at least one paper, including the empty subset if one allows other coauthors. He has compiled a computer-assisted survey of complexity classes called "Complexity Zoology."

MIKLOS SANTHA received his Diploma in mathematics in 1979 from Eötvös University in Budapest, and his Ph. D. in mathematics in 1983 from the Université Paris 7. His advisor was Jacques Stern. Since 1988 he has been a CNRS researcher, currently at the Université Paris Diderot, LIAFA. He is also a principal investigator at CQT in Singapore.

OR SATTATH received his B. S. in Physics and Computer Science in 2005, and his M. S. in Computer Science in 2009, both from the Hebrew University. His Ph. D. advisors are Dorit Aharonov and Julia Kempe. He is the proud father of Nadav, his newly born son. His main research interest is quantum complexity theory.

SHENGYU ZHANG received his B. S. in Mathematics at Fudan University in 1999, his M. S. in Computer Science at Tsinghua University under the supervision of Mingsheng Ying in 2002, and his Ph. D. in Computer Science at Princeton University under the supervision of Andrew Chi-Chih Yao in 2006. After working at NEC Laboratories America for a summer, and at the California Institute of Technology for two years as a postdoctoral researcher, he joined The Chinese University of Hong Kong as an assistant professor. His main research interests are complexity theories in various randomized and quantum models.