On the Complexity of Unique Quantum Witnesses and Quantum Approximate Counting

Anurag Anshu
Harvard University, Cambridge, MA, USA

Jonas Haferkamp
Saarland University, Saarbrücken, Germany

Yeongwoo Hwang
Harvard University, Cambridge, MA, USA

Quynh T. Nguyen
Harvard University, Cambridge, MA, USA

Abstract

We study the long-standing open question on the power of unique witnesses in quantum protocols, which asks if UniqueQMA, a variant of QMA whose accepting witness space is 1-dimensional, contains QMA under quantum reductions.

This work rules out any black-box reduction from QMA to UniqueQMA by showing a quantum oracle separation between $\mathsf{BQP}^{\mathsf{UniqueQMA}}$ and QMA. This provides a contrast to the classical case, where the Valiant-Vazirani theorem shows a black-box randomized reduction from UniqueNP to NP, and suggests the need for studying the structure of the ground space of local Hamiltonians in distilling a potential unique witness. Via similar techniques, we show, relative to a quantum oracle, that $\mathsf{QMA}^{\mathsf{QMA}}$ cannot decide quantum approximate counting, ruling out a quantum analogue of Stockmeyer's algorithm in the black-box setting. Our results employ a subspace reflection oracle, previously considered in [4, 3, 51], but we introduce new tools which allow us to exploit the unique witness constraint. We also show a strong "polarization" behavior of QMA circuits, which could be of independent interest in studying quantum polynomial hierarchies.

We then ask a natural question; what structural properties of the local Hamiltonian problem can we exploit? We introduce a physically motivated candidate by showing that the ground energy of local Hamiltonians that satisfy a computational variant of the eigenstate thermalization hypothesis (ETH) can be estimated through a UniqueQMA protocol. Our protocol can be viewed as a quantum expander test in a low energy subspace of the Hamiltonian and verifies a unique entangled state across two copies of the subspace. This allows us to conclude that if UniqueQMA is not equivalent to QMA, then QMA-hard Hamiltonians must violate ETH under adversarial perturbations (more accurately, further assuming the quantum PCP conjecture if ETH only applies to extensive energy subspaces). Under the same assumption, this also serves as evidence that chaotic local Hamiltonians, such as the SYK model may be computationally simpler than general local Hamiltonians.

2012 ACM Subject Classification Theory of computation → Quantum computation theory

Keywords and phrases Quantum complexity, approximate counting, Valiant-Vazirani, eigenstate thermalization hypothesis

Digital Object Identifier 10.4230/LIPIcs.ITCS.2026.10

Related Version *Full Version*: `https://arxiv.org/abs/2410.23811`

Funding This work was done in part while the authors were visiting the Simons Institute for the Theory of Computing, supported by DOE QSA grant number FP00010905. AA and QTN acknowledge support through the NSF Award No. 2238836.
Anurag Anshu: AA acknowledges support through the NSF award QCIS-FF: Quantum Computing & Information Science Faculty Fellow at Harvard University (NSF 2013303).
Jonas Haferkamp: JH acknowledges support through the Harvard Quantum Initiative postdoctoral fellowship.

Yeongwoo Hwang: YH is supported by the National Science Foundation Graduate Research Fellowship under Grant No. 2140743. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors(s) and do not necessarily reflect the views of the National Science Foundation.
Quynh T. Nguyen: QTN acknowledges support through the Harvard Quantum Initiative PhD fellowship.

Acknowledgements In an earlier version of this paper, the oracle separation between UniqueQMA and QMA was given via a quantum oracle. We thank Chinmay Nirkhe and Anand Natarajan for pointing out that our proof did not in fact need quantumness in this case. This version has been updated to provide a classical oracle separation. We thank Soonwon Choi, Sam Garratt, Yingfei Gu, Rahul Jain and Shivaji Sondhi for insightful discussions, and especially thank Soonwon Choi for sharing an example which shows that the computational ETH assumption is incorrect if the Gaussian prescription is applied at high energies (we only applying it at low energies).

1 Introduction

One of the central aims of quantum complexity theory is to explore the structure of ground, and more generally low-energy, subspaces in quantum many-body systems. A key question in this field is whether these low-energy subspaces of local Hamiltonians are computationally simpler than generic quantum subspaces.[1] A long-standing problem which formalizes this inquiry is the relationship between UniqueQMA and QMA [9]. The complexity class UniqueQMA captures verification tasks in which positive instances have a 1-dimensional accepting subspace (unique quantum witness) while QMA places no restriction on the dimension of the accepting subspace. The open question is to determine the relationship between these classes: are they equivalent under quantum reductions, or does the uniqueness promise strictly reduce computational power? On one hand, an equivalence would imply that a polynomial-time quantum algorithm can *uniquely* verify some fixed state in a low-energy subspace of local Hamiltonians, a task that intuitively seems unattainable for generic quantum subspaces. On the other, a separation would formalize the physics intuition that "natural" Hamiltonians with a non-degenerate ground space and inverse polynomial spectral gap are computationally simpler than QMA-hard Hamiltonians and thus amenable to analysis.

This question is also motivated by the insights produced by studying the class UniqueNP. Classically, the celebrated Valiant-Vazirani isolation lemma [59] was used to show that UniqueNP contains NP under *randomized* reductions, i.e., $\mathsf{NP} \subseteq \mathsf{RP}^{\mathsf{UniqueNP}}$ (where RP is a one-sided error version of BPP), suggesting that the uniqueness promise does not significantly reduce the computational power of NP, at least under randomized reductions. The isolation lemma has underlaid various developments in theoretical computer science, including approximate counting [57], Toda's theorem [58], parallel computation [43], one-way functions [26], and average-case complexity [15]. One could hope that studying the possibility of a "quantum isolation lemma" would sharpen our understanding in a similarly wide range of quantum problems.

Limitations of unique quantum witnesses. There has been a series of works attempting to resolve this question. Aharonov et al. [9] (first appearing on arxiv in 2008) initiated the study of UniqueQMA and possible quantum analogues of the isolation lemma. There, the authors

[1] From an information-theoretic point of view, the ground space of a local Hamiltonian is known to be simpler as it has a polynomial-sized tensor network approximation [50]. However this tensor network is not known to be computationally tractable.

extended the hashing-based arguments in the Valiant-Vazirani isolation lemma to the classes MA and QCMA. However, they noted the failure of a similar strategy for QMA, citing the fact that two random quantum states have exponentially small overlap. A later work [32] sought a completely different approach to obtain a *deterministic* isolation procedure in a special case, but also fell short of achieving a unique witness protocol for general QMA problems as their strategy worked only for accepting subspaces of polynomial dimension. This led [32] to conjecture that a *quantum* reduction is required to obtain a unique verifier, suggesting the question of whether the inclusion $\mathsf{QMA} \subseteq \mathsf{BQP}^{\mathsf{UniqueQMA}}$ holds.[2] This discussion brings us to our first result.

▶ **Theorem 1.** *There exists a* quantum *oracle relative to which* $\mathsf{QMA} \not\subseteq \mathsf{BQP}^{\mathsf{UniqueQMA}}$.

This oracle separation implies there is no black-box quantum isolation procedure for quantum witnesses, unlike its classical counterpart [59], and explains the failures of all approaches in [9, 32]. It may also be viewed as some evidence for the inequality between UniqueQMA and QMA. Furthermore, it suggests the need to exploit the physical structure of quantum low-energy spaces to construct a UniqueQMA verifier for local Hamiltonians (see our result below).

Our Theorem 1 is based on a quantum oracle reflecting about an unknown subspace, and the task is to determine whether the subspace is non-empty (of some large dimensionality). Although this result in a sense is a continuation of a series of works on quantum oracle separations and unitary property testing [4, 31, 51, 6], the techniques used in our work differ in that they make careful use of the structure of the quantum witness. Previous separations used techniques which were not sensitive to the presence of a witness (e.g. replacing the witness with the maximally mixed state), which would undermine the uniqueness promise in our case. We hope that the techniques presented in this work will help develop similar query lower bounds on other witness-restricted quantum classes, such as QMA(2) where constructing oracle separations have been challenging.

Along the way to Theorem 1 we also obtain a *classical* oracle separation between UniqueQMA and QMA (Corollary 12). In fact, we can even separate NP from $\mathsf{P}^{\mathsf{UniqueQMA}}$ under the same classical oracle. However, this separation is perhaps not as interesting as Theorem 1, for one would a priori expect randomized or quantum reductions were needed to reduce QMA to UniqueQMA, in analogy to the classical case. Regardless, we note that *derandomizing* the Valiant-Vazirani lemma is a long-standing open question by itself [13, 37], and this intermediate result rules out such a derandomization in the black-box setting. Thus, we believe the relationship between $\mathsf{BQP}^{\mathsf{UniqueQMA}}$ and QMA is the proper question through which to study the power of unique quantum witnesses.

On classical vs. quantum oracles. Both of our separations Theorem 1 and Theorem 2 are by quantum oracles. In our approach in Theorem 1, the quantumness is necessary as otherwise the classical version of our oracle problem is easily shown to be contained in $\mathsf{NP} \subseteq \mathsf{BQP}^{\mathsf{UniqueQMA}}$. A similar observation can be made for Theorem 2. Historically, when the class or resource is classical, a classical separation can be viewed as the "gold-standard," as evidenced by the long series of works aiming to separate QCMA and QMA with a classical oracle (see the exposition in [60]). However, this decision should be made on a case-by-case basis. In our context, the subspace reflection oracle captures a way in which an algorithm

[2] In this work, is implicit that the oracularized complexity classes refers to the promise version (i.e. UniqueQMA actually refers to PromiseUniqueQMA).

might probe the ground space of a Hamiltonian. Naively, the best algorithmic way to "access" this subspace is to use quantum phase estimation to, say, implement reflections about it. There is a priori no known reason to suspect that this subspace will be well structured for worst-case Hamiltonians and thus we instantiate the ground space with a subspace in a randomly chosen basis. Another consideration when using and interpreting quantum oracle separations is that they only rule out quantum relativizing proof strategies. Indeed, [1, 7] already point out that the proof technique of explicitly representing quantum amplitudes is classically relativizing but quantumly not (and, to our knowledge, this is the only known proof technique with such property). However, this technique has only been used used to show containments in very powerful classical classes like EXP and PSPACE. Thus, quantum oracle separations are still useful in ruling out black-box reductions at lower complexity regimes like $\mathsf{BQP}^{\mathsf{UniqueQMA}}$ and QMA.

Improved quantum approximate counting lowerbounds. Classically, the question of uniqueness is closely related to approximate counting. In [59], a random hash is used to isolate a single solution to an NP circuit. The celebrated Stockmeyer algorithm [57] uses similar tools to give a $\mathsf{BPP}^{\mathsf{NP}}$ algorithm for estimating the number of accepting witnesses to multiplicative precision. A quantum analogue of this is to multiplicatively estimate the dimensionality of the accepting subspace of a QMA circuit [18, 53, 16]. The complexity class capturing this quantum approximate counting problem is QXC. As usual, one should be careful when defining quantum analogues of classical classes. However, this problem is well motivated and shown in [17] to be connected to a number of physical relevant computational problems, such as computing quantum partition functions and estimating local observables of Gibbs states. QXC was further studied in [16] and shown to contain the problem of estimating the Kronecker coefficients of the symmetric group. Stockmeyer's approximate counting algorithm has led to many important applications, including the sampling-counting equivalence [33] and even the theoretical foundations for quantum advantage experiments [28]. Thus, it is natural to ask whether there exists a quantum analogue: Could quantum approximate counting be solved in (quantum) polynomial time with access to an oracle solving QMA decision problems? In essence, do we have $\mathsf{QXC} \subseteq \mathsf{BQP}^{\mathsf{QMA}}$?

The phrase "quantum approximate counting" has appeared elsewhere in the literature. In [39, 2], the authors study SBQP, the quantum analogue of SBP (where SBP is the multiplicatively precise analogue of BPP), which captures the task of estimating the acceptance probability of a BQP machine to multiplicative precision. This is not known to be equivalent to the definition of QXC (unlike SBP which is equivalent to classical approximate counting). We do not consider SBQP in this work. Additionally, in [3], "quantum approximate counting" is used to refer to the task of designing quantum (BQP or QMA) algorithms for classical approximate counting, given quantum queries to a Boolean oracle. [51] consider a quantum generalization of the problem in [3], called the quantum approximate *dimension* problem; the task here is to decide whether an unknown subspace has dimension $\leq w$ or $\geq 2w$, given access to the corresponding reflection oracle. Using the quantum approximate dimension problem, we show

▶ **Theorem 2** (Corollary 22, restated)**.** *There exists a* quantum *oracle relative to which* $\mathsf{QXC} \not\subseteq \mathsf{BQP}^{\mathsf{QMA}}$*, and in fact,* $\mathsf{QXC} \not\subseteq \mathsf{QMA}^{\mathsf{QMA}}$*.*

Our Theorem 2 rules out the existence of a quantum analogue of Stockmeyer's theorem in the black-box setting. This is a strengthening of a result of [51], where the authors show a QMA query lower bound against this task via a reduction to the classical problem in [3].

Again, the quantum oracle is necessary in our approach, because the classical version of our oracle is the classical approximate counting problem, which is contained in $\mathsf{BPP}^{\mathsf{NP}} \subseteq \mathsf{BQP}^{\mathsf{QMA}}$. The polynomial methods used in [3, 51] do not transfer to the relativized setting in the presence of QMA-oracle calls. Moreover, prior works [2, 6] dealing with quantum versions of the polynomial hierarchy are all based on variants of the switching lemma, which are not known to be compatible with quantum oracles. Therefore, we need to develop new techniques to handle quantum oracle calls in "stacked" complexity classes like $\mathsf{QMA}^{\mathsf{QMA}}$.

Along the way we also prove a *classical* oracle separation $\mathsf{QXC} \not\subseteq \mathsf{QMA}$, in fact, $\mathsf{SBP} \not\subseteq \mathsf{P}^{\mathsf{QMA}}$, thus giving a much simpler and elementary proof of the QMA lower-bound against classical approximate counting of [3] (who showed $\mathsf{SBP} \not\subseteq \mathsf{QMA}$). Our separation is perhaps stronger as $\mathsf{P}^{\mathsf{QMA}}$ likely strictly contains QMA as it contains coQMA.

Unique witness from a well-connected ground space. Theorem 1 underscores the importance of gaining a deeper understanding of the structure of quantum ground spaces to extract a unique quantum witness. For this, it is crucial to identify classes of Hamiltonians for which we know how to verify a unique low-energy witness. However, the only such classes of Hamiltonian known so far are classical Hamiltonians [59], local Hamiltonians with ground states that can be prepared by a polynomial-sized quantum circuit [9], and local Hamiltonians with a polynomial sized low-energy subspace [32]. Our final result identifies a new class of Hamiltonians for which we can verify a unique low energy state. Our starting point is a simple observation already hinted at in the previous paragraphs: An algorithm that only uses subroutines treating the *entire* Hamiltonian H as a single operator cannot distinguish between states of the same energy as they are indistinguishable with respect to H. For example, quantum phase estimation, which only uses time evolutions of H, falls into this category. Thus, any quantum algorithm that aims to verify a unique low-energy state must use operators which are more fine-grained than H itself, namely the *individual* local terms in H.

The action of local operators on Hamiltonian eigenstates is in itself a challenging problem, with few general results known [12]. Despite this, there are families of local Hamiltonians for which useful statements can be made. A particularly interesting such family is the set of local Hamiltonians that satisfy the Eigenstate Thermalization Hypothesis (ETH) from quantum statistical mechanics [56]. This hypothesis roughly says that within an energy window, eigenstates are well-connected in the sense that local operators induce transitions from one eigenstate to many others. We adopt the following informal definition from [19].

> *ETH (informal).* There is a collection of polynomially many local observables $A_1, \ldots A_m$ such that for every eigenstate $|\psi\rangle$ with energy roughly E, the vector $A_i |\psi\rangle$ roughly stays within a small energy window around E, and has a fairly good overlap with all the eigenstates in this window.

See Section 3.3.1 for the formal definition of ETH (Hypothesis 24) and a discussion on its physical origins. From the computer science perspective, a well-connected low-energy eigenspace should have features of an expanding graph and hence allow one to single out a unique "fixed point". We solidify this intuition by generalizing the quantum expander test for verifying the maximally entangled state [10] and show the following theorem:

▶ **Theorem 3** (Informal restatement of Theorem 25)**.** *There is a* UniqueQMA *protocol for estimating the ground energy of a local Hamiltonian satisfying ETH.*

The protocol verifies a unique quantum state in two copies of the subspace with energy around E, if ETH holds in that subspace. Assuming E is the ground energy itself, we can extract a unique ground state witness. If E is larger but sub-extensive (that is, n^c with $c < 1$, as a function of the number of qubits n), we can still obtain a unique witness for the purposes of UniqueQMA protocols due to known reductions concerning the local-Hamiltonian problem.

A likely implication of Theorem 3 is that the complexity of local Hamiltonians satisfying ETH, such as chaotic quantum Hamiltonians including the SYK model [55], is significantly lower than that of general local Hamiltonians. It is unclear if the theorem would help establish that UniqueQMA and QMA are equivalent under quantum reductions, since we do not know if QMA-hard local Hamiltonians would satisfy the ETH assumption.

2 Preliminaries

2.1 Complexity classes

▶ **Definition 4** (QMA). *A promise problem $L = (L_{yes}, L_{no})$ is in the complexity class* Quantum Merlin-Arthur *(or* QMA*) if there exists a polynomial-time classical algorithm yielding a description of a quantum circuit V which implements a quantum operator $U_x : \mathcal{H} \to \mathcal{H}$ with $\mathcal{H} = \mathcal{H}_{R_{\text{in}}} \otimes \mathcal{H}_{R_{\text{aux}}}$, such that*

1. *(Completeness) For all $x \in L_{yes}$, there exists a* witness *$|\phi\rangle \in \mathcal{H}_{R_{\text{in}}}$ such that*

$$\|\Pi_{acc} U_x |\phi\rangle_{R_{\text{in}}} |0\rangle_{R_{\text{aux}}} \|_2 \geq \frac{2}{3}$$

2. *(Soundness) For all $x \in L_{no}$ and $|\phi\rangle \in \mathcal{H}_{R_{\text{in}}}$,*

$$\|\Pi_{acc} U_x |\phi\rangle_{R_{\text{in}}} |0\rangle_{R_{\text{aux}}} \|_2 \leq \frac{1}{3}$$

where Π_{acc} can be taken to be the projector into the $|0\rangle$ on the first qubit.

▶ **Definition 5** (Unique QMA). *A promise problem $L = (L_{yes}, L_{no})$ is in the complexity class* Unique Quantum Merlin-Arthur *(or* UniqueQMA*) if in addition to the above properties of* QMA*, we have the further restriction that in the* completeness *case, there is some accepting state $|\phi\rangle$ and for all states orthogonal to $|\phi\rangle$, the verifier accepts with probability equal to the soundness case. In particular,*

1. *(Completeness) For all $x \in L_{yes}$, there exists a* unique witness *$|\phi\rangle \in \mathcal{H}_{witness}$ such that*

$$\|\Pi_{acc} U_x |\phi\rangle_{R_{\text{in}}} |0\rangle_{R_{\text{aux}}} \|_2 \geq \frac{2}{3}$$

and for all states $|\psi\rangle \perp |\phi\rangle$,

$$\|\Pi_{acc} U_x |\psi\rangle_{R_{\text{in}}} |0\rangle_{R_{\text{aux}}} \|_2 \leq \frac{1}{3}$$

2.2 Quantum oracles

Our first result is a oracle separation between QMA and UniqueQMA. There have been various types of oracles considered in previous works. Although the "gold standard" for oracle separations is arguably a classical oracle (which can be queried in superposition), quantum separations have required more complex oracles, such as the distributional oracles

of [45]. In this work, we work with quantum oracles, as studied in [4]. Thus, in our setting, an oracle $\mathcal{O}$ refers to some arbitrary unitary operation. We will also allow controlled queries to $\mathcal{O}$, so that a call to $\mathcal{O}$ is in fact a call to the unitary,

$$(\mathbb{I} - \Pi) \otimes \mathbb{I} + \Pi \otimes \mathcal{O}\,,$$

where Π is a projector representing a control subspace. Given a family of oracles, we can define a "oracle promise problem,"

▶ **Definition 6** (Oracle promise problem). *A oracle promise problem is defined by two disjoint sets of oracles (unitary operators), Ω_{YES} and Ω_{NO}. We use $\Omega = (\Omega_{YES}, \Omega_{NO})$ to refer to the decision problem itself.*

Then, any L-query algorithm deciding Ω should be able to receive an oracle $\mathcal{O}$ either in the YES or NO case and use L calls to $\mathcal{O}$ to decide which is the case. Formally, a oracular verifier for $\mathsf{UniqueQMA}^{\mathcal{O}}$ consists of alternating unitaries (consisting of local quantum gates) and calls to $\mathcal{O}$. The task is to (uniquely) determine whether the unknown oracle corresponds to a YES case or NO case instance.

▶ **Definition 7** (UniqueQMA Oracle verifier). *Given a n-qubit oracle $\mathcal{O}$ and $m \in \mathsf{poly}(n)$, a m-qubit, L-query* UniqueQMA *oracular verifier $V^{\mathcal{O}}$ is defined over a n-qubit oracle register $R_{\mathcal{O}}$, in a n_{aux} qubit auxiliary register R_{aux}. Additionally, there is a n_{in}-qubit input register R_{in}. Thus, $m = n_{\text{aux}} + n_{\text{in}}$. Given a n_{in}-qubit witness, the verifier then performs a sequence of unitaries on the full m-qubit subspace, interwoven with L queries to an oracle $\mathcal{O}$ on $R_{\mathcal{O}}$, controlled on the remaining system. Finally, the verifier performs the measurement $\{|0\rangle\langle 0|, |1\rangle\langle 1|\}$ on the first qubit in the auxiliary register (which we label as the register R_{out}), rejecting if it obtains $|0\rangle\langle 0|$ the and accepting if obtains $|1\rangle\langle 1|$. See Figure 1. Thus, the measurement corresponding to the "accept" case can be written as*

$$\begin{aligned} V &\triangleq (\mathbb{I} \otimes |1\rangle\langle 1|_{R_{\text{out}}}) U_L \mathcal{O} U_{L-1} \dots U_1 \mathcal{O} U_0 (\mathbb{I} \otimes |0^{n_{\text{aux}}}\rangle\langle 0^{n_{\text{aux}}}|_{R_{aux}}), \\ \Pi_{accept} &\triangleq V^\dagger V. \end{aligned} \tag{1}$$

We say that V uniquely solves *the oracle promise problem $\Omega = (\Omega_{YES}, \Omega_{NO})$ if*

- *If $\mathcal{O} \in \Omega_{YES}$ then there exists a witness $|\psi\rangle_{R_{\text{in}}}$ so that*

$$\text{acc}(V, \psi) = \| V |\psi\rangle |0\rangle_{R_{\text{aux}}} \|_2^2 = \text{tr}\left[\Pi_{accept} |\psi\rangle\langle\psi|_{R_{\text{in}}} \otimes |0\rangle\langle 0|_{R_{\text{aux}}}\right] \geq 1 - \varepsilon\,,$$

  and for any witness $|\psi'\rangle$ where $|\psi'\rangle$ is orthogonal to $|\psi\rangle$, $\text{acc}(V, \psi') \leq \varepsilon$.
- *If $\mathcal{O} \in \Omega_{NO}$ then $\text{acc}(V) \leq \varepsilon$; for all witnesses the verifier accepts with probability at most ε.*

The standard oracle that we consider in this work are variations of an oracle considered in prior works: the subspace reflection oracle $\mathcal{O}^{\mathcal{S}} = \mathbb{I} - 2\Pi_{\mathcal{S}}$, where $\Pi_{\mathcal{S}}$ corresponds to a projection onto a k-dimensional subspace $\mathcal{S}$. In [4], $\mathcal{O}^{\mathcal{S}}$ with $k = 1$ was used to separate QCMA from QMA; in [3], $\mathcal{S}$ was restricted to be a classical subspace to separate QMA from classical approximate counting; and finally [51] quantized the prior work to separate QMA from quantum approximate counting. We borrow notation from these last two works to define the oracle promise problems we study.

▶ **Definition 8** (Quantum approximate dimension). *In the oracle problem $\mathsf{QApxDim}(k_1, k_2, N)$, the problem is to decide whether an n-qubit unknown oracle corresponds to a reflection about a subspace of dimension $\geq k_2$ (the YES case), or a subspace of dimension $\leq k_1$ (the NO case), promised that one of these two is the case. For technical reasons, we restrict $k_2 \leq 2^{2n/3}$.*

If we are further promised that the subspace is in the computational basis, then we have the classical problem $\mathsf{ApxDim}(k_1, k_2, N)$.

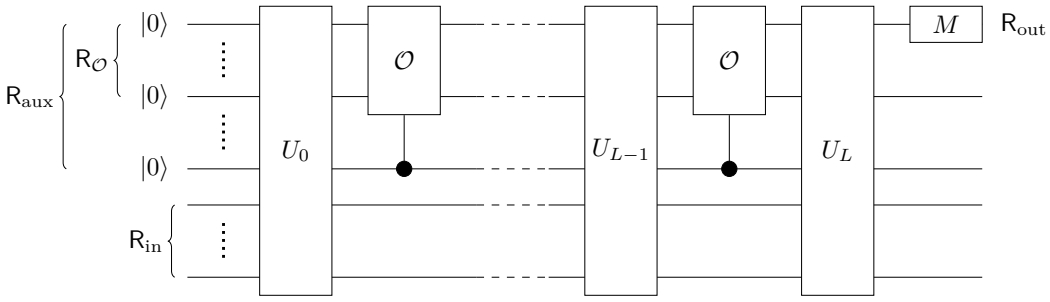

Figure 1 Circuit diagram for the verifier V. $\mathsf{R}_{\mathcal{O}}$ is the oracle register, contained in the auxiliary register R_{aux}, initialized to $|0\rangle_{\mathsf{R}_{\text{aux}}}$. R_{in} is the input register. The final measurement M is performed on R_{out}, which is the first qubit of the auxiliary register.

Variations of this oracle have been used in other works. The case when $\mathcal{S}$ corresponds to a 1-dimensional subspace is exactly the oracle used by [4] to separate QCMA from QMA. When increasing the dimension and restricting the subspace $\mathcal{S}$ to be classical, this becomes the classical oracle used to separate QMA and classical approximate counting [3]; when the subspace is quantum, this is exactly the QApxDim problem and has been shown by [51] to separate QMA and quantum approximate counting. Having said that, our results and proof techniques do not follow from these works in any straightforward way. For example, the [4] oracle would not work in our setting as this problem is easily seen to be in UniqueQMA. A better candidate is the oracle $\mathcal{O}_S$ corresponding to a large subspace S, as there does not seem to be a natural "unique" witness that the prover could provide. However, when constructing our query lower bounds (which are, as usual, the heart of nearly all oracle separations), new techniques are necessary to account for the uniqueness restriction on the witness. Indeed, prior oracle separations involving QMA avoid directly dealing with the witness either by conditioning on a particular classical witness [4], replacing the witness with the maximally mixed state [3, 51], or reducing to a classical circuit lower bound [2, 6]. Instead, our work carefully uses the requirement that the witness be unique.

▶ Remark 9 (Classical oracles). We use the term "classical oracle" to refer to phase oracles $\mathcal{O}^S = \mathbb{I} - 2\Pi_S$, where S is restricted to be in the computational basis. The truly "classical" analogue would be the oracle mapping $|x\rangle|b\rangle \to |x\rangle|b \oplus \mathbf{1}_{x\in S}\rangle$. However, as our separations holds under controlled oracles, these notions are equivalent.

In this work, we also consider oracle separations between relativized classes, e.g. $\mathsf{BQP}^{\mathsf{UniqueQMA}}$. We instantiate verifiers belong to a relativized class ${\mathcal{C}_1}^{\mathcal{C}_2}$ as a verifier for $\mathcal{C}_1$, with access to an black-box oracle $\mathcal{O}$ for a $\mathcal{C}_2$-complete problem. In our case, $\mathcal{C}_2$ is either UniqueQMA or QMA and thus $\mathcal{A}$ is a classical oracle and the corresponding complete problem will be UniqueQCircuitSat or QCircuitSat respectively. However, one difficulty in defining it in this way is that these languages correspond *promise* problems, but there is not guarantee that an algorithm querying $\mathcal{A}$ will do so only on valid instances. To deal with this issue, we take an approach used in prior works [2, 31] and imagine that the outer verifier is given access to an oracle $\mathcal{A}$ for a decision "extension" of the promise problem.

▶ **Definition 10** (Promise oracle extension)**.** *For any oracle given by the partial function $\mathcal{O} : \{0,1\}^n \to \{0,1,\perp\}$, an oracle extension is a total function $\mathcal{A} : \{0,1\}^n \to \{0,1\}$ such that $\mathcal{O}(x) = \mathcal{A}(x)$ for all $x \in Dom(\mathcal{O})$.*

Thus, we take the following definition of a verifier accessing a promise oracle $\mathcal{O}$, as suggested in [2]:

$$V^{\mathcal{O}}(x) \triangleq \begin{cases} 1 & \text{if } V^{\mathcal{A}}(x) = 1 \text{ for every } \mathcal{A} \text{ extending } \mathcal{O}, \\ 0 & \text{if } V^{\mathcal{A}}(x) = 0 \text{ for every } \mathcal{A} \text{ extending } \mathcal{O}, \\ \perp & \text{otherwise.} \end{cases}$$

As stated in [2], it is not obvious that this is the "right" way to define oracle access to a promise problem. Nonetheless, we believe this to be a fairly reasonable definition; when querying an oracle for, e.g., QuantumCircuitSat on an instance falling outside of the promise gap, why should the verifier expect a "reasonable" response from oracle?

Lastly, we remark that whenever we have a verifier of the form ${\mathcal{C}_1}^{\mathcal{C}_2}$, we assume that the oracle for $\mathcal{C}_2$ acts on instances of size polynomial in the outer problem instance. This is required technically as many of our arguments go through union bounding over possible inputs to the oracle for $\mathcal{C}_2$. Moreover, as noted in [3], there is a distinguisher between (classical) subspaces of dimension k and $2k$ with a witness of size roughly $\mathcal{O}(k)$. Thus, without a bound on the witness register, this distinguisher would preclude a $\mathsf{P}^{\mathsf{QMA}}$ query lower bound for approximate counting.

3 Results

3.1 Separating QMA from $\mathsf{BQP}^{\mathsf{UniqueQMA}}$

We consider circuits with access to the unknown subspace reflection oracle $\mathcal{O}_S$, as well as an oracle $\mathcal{A}$ for UniqueQCircuitSat (more precisely, their "oracularized" versions), the natural complete problems for UniqueQMA. Rather than directly working with ApxDim and QApxDim, we consider a slightly different oracle promise problem, QApxDimExact with parameters k_1 and k_2 so that the YES case corresponds to subspace dimension exactly k_2, and the NO case corresponds to dimension exactly k_1. Then, the specific problems we consider are,

$$\mathsf{EmptyNonEmpty} \triangleq \mathsf{ApxDimExact}(0,1,N) \cup \mathsf{ApxDimExact}(0,k,N) \text{ and}$$
$$\mathsf{QEmptyNonEmpty} \triangleq \mathsf{QApxDimExact}(0,1,N) \cup \mathsf{QApxDimExact}(0,k,N)\,.$$

where the union is taking within the NO case and YES case oracles individually (so that the NO case is still the identity oracle, and YES case corresponds to subspaces of dimension 1 and k). Since $\mathsf{EmptyNonEmpty} \subseteq \mathsf{ApxDim}(0,1,N)$ (and $\mathsf{QEmptyNonEmpty} \subseteq \mathsf{QApxDim}(0,1,N)$), we could've considered the harder problem instead, but we use EmptyNonEmpty since this includes all subspace dimensions needed to make the proof go through.

The overall idea for the separation is to use a hybrid argument and separately consider steps of the $\mathsf{BQP}^{\mathsf{UniqueQMA}}$ verifier where a query is made to the circuit sat oracle $\mathcal{A}$ and steps where the verifier directly queries the oracle $\mathcal{O}_S$. Therefore, our separation requires two ingredients.

1. First, we argue that the verifier's calls to $\mathcal{A}$ are either on circuits $\mathtt{C}$ which are themselves good verifiers for the oracle distinguishing task *or* are nearly useless. We show this by establishing a *polarization* property of oracular quantum circuits.
2. Second, we establish our oracle separation between UniqueQMA and QMA. This allows us to rule out the first case in the item above, as well as to argue that the verifier's direct calls to $\mathcal{O}_S$ are not too helpful.

Interestingly, for the second step, the separating oracle can actually be made classical and we show a UniqueQMA query lower bound for $\mathcal{O}_S$ restricted to a *computation basis* subspace.

▶ **Theorem 11** (UniqueQMA query lower bound, EmptyNonEmpty). *Let V be a L-query, m-qubit verifier which uniquely decides EmptyNonEmpty with error $\varepsilon \leq 2^{-n}$ both on $\mathcal{O}_{No}$ and on at least $(1-2^{-n})$-fraction of YES instances. Then, $L \in \Omega(\sqrt{k})$. In particular, if we take $k \in 2^{\Omega(n)}$, then V requires* exponentially-*many queries in n.*

Standard diagonalization arguments yield the following.

▶ **Corollary 12.** *There exists a classical oracle $\mathcal{O}$ such that* $\mathsf{UniqueQMA}^{\mathcal{O}} \subsetneq \mathsf{QMA}^{\mathcal{O}}$*, and in fact,* $\mathsf{P}^{\mathsf{UniqueQMA}^{\mathcal{O}}} \not\subseteq \mathsf{NP}^{\mathcal{O}}$.

In a sense the proof of the above theorem is where the uniqueness property is crucially used. Imagine we have two classical subspaces S and T, where T is a subspace *extending* S, so that $T = S \oplus \Delta$. We will pick the dimension of S and T to be on the order superpoly(n); we will show that as $\dim(S), \dim(T) \ll 2^n$, a UniqueQMA verifier using poly(n) oracle queries cannot distinguish S and *random* extensions T. When S is taken from as a YES instance of an oracle distinguishing task with a unique witness ϕ_S, this observation essentially says that the "same" witness works for T. Flipping the order of quantification, we can equivalently say that for a random T, *nearly all* of its "roots" S have the same witness. But since $\dim(T) \gg \dim(S)$, the S's can be picked so that they are orthogonal. Finally, applying a variation of the hybrid method shows that the same witness must overlap too many orthogonal subspaces, yielding a contradiction.

However, when moving to Item 1, the use of quantum oracle is seemingly necessary. When considering the $\mathsf{BQP}^{\mathsf{UniqueQMA}}$ verifier's direct calls to $\mathcal{O}_S$, the correlation between the verifier's state and the oracle is naturally captured by the overlap of some intermediate state and the projector onto the oracle subspace; we can get a lot of mileage by studying quantities of the form "$\mathrm{tr}[\Pi_S \psi]$" and averaging over subspaces S. The calls to the circuit sat oracles $\mathcal{A}$ require more careful handling. The challenge is, for each oracular circuit $\mathsf{C}^{\mathcal{O}_S}$ on which the outer verifier is called, we would like to show that the answer $\mathcal{A}(\mathsf{C}^{\mathcal{O}_S})$ cannot help tell apart the yes and no case oracles (with high probability). We establish this via a concentration argument, enabled by our use of quantum oracles. In our case, we use a version of Levy's lemma for the Grassmannian (the space of k-dimensional subspaces of $\mathbb{C}^N$), due to [25].

▶ **Lemma 13** (Levy's lemma on the Grassmannian (informal)). *Let Δ be uniform over $G_{N,k}$ and Π_Δ be corresponding projector. Given any quantum state $|\psi\rangle \in \mathbb{C}^{N'}$, with $N' \geq N$, we have that*

$$\Pr_{\Delta \sim G_{N,k}} [\mathrm{tr}[\Pi_\Delta |\psi\rangle\langle\psi|] \geq 2N^{-1/3}] \leq \exp\left(-\mathsf{poly}(N)\right) .$$

Via Lemma 13 we show that the acceptance probabilities of oracular quantum polynomial time circuits must *concentrate* when averaged over the random quantum subspace oracles. A similar observation is made in [31] to rule out the possibility of a search-to-decision reduction for QMA. We use this observation somewhat differently. Whereas [31] used concentration to argue that on across *yes* instances the circuit behaves similarly, we instead want to argue the same thing across yes *and no* instances. This requires a different and more intricate argument considering the average acceptance probabilities in the yes and no cases individually. In the specific form used in this work, we call this property "polarization".

▶ **Lemma 14** (Polarization of UniqueQMA oracle circuits). *Let V be an polynomial-time oracular verifier and $k = N^\alpha$, with $\alpha \leq 2/3$. Then for any function $f(n) \in \exp\left(-N^{1/4}\right)$:*

$$\Pr_{\mathcal{T} \sim \mathrm{Haar}_{N,k}} [\mathcal{A}^{\mathcal{T}}(V) \neq \mathcal{A}^{\emptyset}(V)] \leq f(n) .$$

The final obstacle when designing oracle separations for relativized (or "stacked") classes is the fact that the natural complete problem for UniqueQMA, quantum circuit sat, is defined as a *promise problem*. As mentioned above, we address this issue by defining an appropriate "decision-extension" of the promise oracle.

▶ **Theorem 15** (Relativized Oracle Separation). *Let $V^{\mathcal{T}}$ be any L-query* BQP *verifier with gates implementing controlled calls to the subspace reflection oracle $\mathcal{O}^S$, as well as an oracle for* $\textsc{UniqueQCircuitSat}^{\mathcal{O}^S}$. *Suppose $V^{\mathcal{T}}$ solves* QEmptyNonEmpty*, with error at most $\varepsilon \in \exp(-n)$. Then, if $k \geq N^{\alpha}$ for some $\alpha \in \Omega(1)$, then $L \in \Omega(N/k)$.*

▶ **Corollary 16.** *There exists a quantum oracle $\mathcal{O}$ such that* $\mathsf{BQP}^{\mathsf{UniqueQMA}^{\mathcal{O}}} \not\subseteq \mathsf{QMA}^{\mathcal{O}}$.

3.2 Separating QXC from $\mathsf{QMA}^{\mathsf{QMA}}$

In this section, we use the techniques from the previous result to obtain an oracle separation between QMA and quantum approximate counting. This definition was introduced by [17].

▶ **Definition 17** (Quantum approximate counting, QXC). *Let V be a m-sized* QMA *verification circuit which takes a n-qubit state $|\psi\rangle$ as input, with $m \in \mathsf{poly}(n)$. Then, given thresholds $0 < b < a \leq 1$ with $a - b \geq \frac{1}{\mathsf{poly}(n)}$, and an error parameter $\varepsilon \geq \frac{1}{\mathsf{poly}(n)}$, the approximate counting problem* QXC *is to compute an estimate D such that $(1-\varepsilon)D_a \leq D \leq (1+\varepsilon)D_b$, where D_a is defined as the dimension of the witness subspace for which V accepts with probability at least a, respectively for D_b.*

We formalize the above as oracle promise problem $\mathsf{ApxDim}(k_1, k_2, N) = (\Omega_{\text{Yes}}, \Omega_{\text{No}})$ with $k_2 \leq N$ and $k_1 \leq k_2(1 - 1/\mathsf{poly}(n))$ where,

- Each $\mathcal{O} \in \Omega_{\text{Yes}}$ corresponds to a reflection over a subspace $S \subseteq \mathbb{C}^N$ of dimension k_1, with $\mathcal{O} = \mathbb{I} - 2\Pi_S$.
- Each $\mathcal{O} \in \Omega_{\text{No}}$ corresponds to a reflection over a subspace $T \subseteq \mathbb{C}^N$ of dimension k_2, with $\mathcal{O} = \mathbb{I} - 2\Pi_T$.

Intuitively, the different parameters can be attributed to the fact that in UniqueQMA, the unique witness condition affords structure even within the YES case instances (thus, both subspaces of dimension $\approx 2^{k_1 n}$ and $\approx 2^{k_2 n}$ can be treated as YES cases). For QMA, we need to treat the dimension $\geq 2^{k_2 n}$ as a NO case in order to carry through similar arguments. As in the case of UniqueQMA, we will combine an (classical) oracle separation between QMA and QXC as well a version of the polarization lemma for QMA verifiers.

▶ **Theorem 18** (QMA query lower bound for ApxDim). *Let $k_1 \in o(N)$ and $k_2 = 2k_1$. Let $V^{\mathcal{O}}$ be an oracular* QMA *verifier deciding $\mathsf{ApxDim}(k_1, k_2, N)$ with error $\varepsilon \leq \exp(-n)$. Then, $L \in \Omega(\sqrt{N/k_1})$.*

▶ **Lemma 19** (Polarization of QMA oracle circuits). *Let $\mathcal{C}$ be any $n' = \mathsf{poly}(n)$-qubit* QMA *circuit which makes $L \in o(\sqrt{N/k_1})$ queries to a n qubit oracle $\mathcal{O}_{\mathcal{S}} = \mathbb{I} - 2\Pi_{\mathcal{S}}$. Let $\mathcal{S}$ be a dimension k_1 subspace. Let $\mathcal{T}$ be a $k_2 < N^{2/3}$ dimensional extension of $\mathcal{S}$. Then for some decision-extension $\mathcal{A}$ of an oracle for $\mathsf{QCiruitSAT}(\eta, 1-\eta)$,*

$$\Pr_{\mathcal{T} \sim \mathcal{Q}_{\mathcal{S}\uparrow, k_2}} [\mathcal{A}(\mathcal{C}^{\mathcal{S}}) \neq \mathcal{A}(\mathcal{C}^{\mathcal{T}})] \leq 2\gamma\,,$$

where $\gamma \leq \exp\left(-N^{1/3}\right)$.

▶ Remark 20. A similar lower bound where dimension k_1 corresponds to the YES case and dimension k_2 corresponds to the NO case is obtained in [51, 3] but through much more intricate methods utilizing a version of the polynomial method for Laurent polynomials. Our proof is much simpler and gets the same quantitative bound, which is shown to be tight in [3, Theorem 5]. However, their techniques are better in that their lower bound is symmetric (it also works when the YES case is the larger subspace, and the NO case is smaller). Intuitively, this is because their proof goes via the inclusion of QMA in SBQP, the exponentially precise analogue of BQP, and thus is not sensitive to a witness.

In both separations, the polarization property puts strong limitations on the way the outer verifiers can interact with the circuit sat oracles. As such, we hope these techniques will generalize beyond our setting to yield lower bounds against verifiers querying quantum oracles, such as those capturing constant stacks of QMA (the "QMA hierarchy").

Finally, we make a remark on our (classical) oracle separation of QMA and QXC. One might be tempted to leverage an argument similar to [4] to remove the outer witness. This does not work, as the witness in our case is quantum and would require conditioning on a event with probability $\leq \exp\left(-2^{\mathsf{poly}(n)}\right)$. Instead, we use a hybridization argument similar to the UniqueQMA proof. When extending a (yes case) small subspace S to a (no cases) larger subspace $S \oplus \Delta$, the only way the QMA algorithm changes its behavior is that the witness overlaps with Δ. Thus, choosing many orthogonal extensions Δ leads to a contradiction. This recovers the QMA lower bound for classical approximate counting (with quantum queries) of [3], albeit by a much more straightforward proof without involving the machinery of Laurent polynomials used in [3].

▶ **Theorem 21** (Query lower bound for $\mathsf{QMA}^{\mathsf{QMA}}$)**.** *Let V be any $m = \mathsf{poly}(n)$-qubit, L-query oracular* QMA *verifier which purports to solve the n-qubit oracle promise problem $\mathsf{QApxDim}(k_1, k_2, N)$ with error $\varepsilon \leq \exp(-n)$ and $k_1 < k_2 \leq N^{2/3}$. Suppose V has gates implementing controlled calls to a n qubit subspace reflection oracle $\mathcal{O}^{\mathcal{S}}$, as well as an $n_o \leq n'$ qubit oracle $\mathcal{A}$ for* $\textsc{QCircuitSat}^{\mathcal{O}_{\mathcal{S}}}$*. Then $L \in \Omega(N^{1/3})$.*

▶ **Corollary 22.** *There exists a quantum oracle $\mathcal{O}$ such that $\mathsf{QXC}^{\mathcal{O}} \not\subseteq \mathsf{QMA}^{\mathsf{QMA}^{\mathcal{O}}}$.*

3.3 Protocol for unique verification

The oracle separations in the previous section suggest that we need to use the structure of QMA problems if we want to have a chance to design a unique verifier. In this section, we describe a UniqueQMA verification protocol for local Hamiltonians that are sufficiently chaotic in the sense of the eigenstate thermalization hypothesis (ETH) – an influential hypothesis in quantum statistical mechanics.

3.3.1 Eigenstate thermalization hypothesis

Thermodynamics expects the evolution of a generic many-body system to be ergodic, meaning that the system evolves to eventually explore all possible configurations and thermalize. However, the evolution of a closed quantum system is governed by the Schrodinger's equation which is reversible and also leaves eigenstates invariant. Why do individual energy eigenstates of a large, interacting quantum system look thermal when probed by "simple" observables (like local operators), even though the full wavefunction evolves unitarily without randomness? A widespread explanation to this seeming paradox is the eigenstate thermalization hypothesis (ETH) [23, 56], which has produced many predictions in agreements with thermalization on many physical systems [21]. ETH has many slight variants, but they are typically a set of

assumptions about generic Hamiltonians that, informally, predicts that the eigenstates locally resemble a Gibbs state, and furthermore, local observables obey certain random matrix behaviors in the Hamiltonian energy eigenbasis inspired by quantum chaos and random matrix theory.

The first mathematically clean and consistent description of ETH was proposed by Srednicki [56], which is in turn inspired by Wigner's pioneering random matrix theory (RMT) framework of analyzing quantum systems using statistical methods. Srednicki's ETH proposes a "chaotic" behavior of local operators in the eigenbasis of a local Hamiltonian. It is hard to characterize precisely how a local operator A may act on an energy eigenstate $|\nu_i\rangle$, so ETH leverages randomness to state that the matrix elements of a local operator A "looks like" $\langle\nu_j| A |\nu_i\rangle = c_i\delta_{i,j} + e^{-S/2}R_{ij}$, where c_i is some quantity capturing $\langle\nu_i| A |\nu_i\rangle$ and $|\nu_i\rangle, |\nu_j\rangle$ are eigenstates with nearby eigen-energies (a local operator cannot connect far away eigenstates - see [12]) and S is the entropy within this window. Crucially, R_{ij} are chosen to be i.i.d random Gaussians. More formally, we consider the following formulation closely following [19, Section IV], which is arguably more mathematically concrete than [56] to work with at a rigorous level.

Let us first introduce some notations. Fix an energy $e_0 > n^{0.99}$ and $\Delta = 1/\mathrm{poly}(n)$ throughout this section[3]. Denote by $\mathcal{I}_\Delta$ the energy window of width $\Delta/2$ around e_0, $\mathcal{I}_\Delta = [e_0 - \frac{\Delta}{2}, e_0 + \frac{\Delta}{2}]$. Denote by Π_Δ the projector onto the subspace of eigenstates with energy $e \in \mathcal{I}_\Delta$. We will always think of the eigenstates as ordered by non-decreasing energies.

▶ **Hypothesis 23** (Srednicki's ETH ansatzes). *Consider an n-qubit local Hamiltonian H and a narrow energy window of width Δ around an energy e_0. The corresponding "ETH ansatz ensemble", denoted $\mathcal{G}$, is a distribution over Hermitian operators $A \sim \mathcal{G}$ with parameters $\{\mu_\alpha\}_\alpha$ and $\{f_{\alpha,\beta}\}_{\alpha,\beta}$, whose entries within the said window have the form*

$$\langle\nu_\alpha| A |\nu_\beta\rangle = \mu_\alpha\delta_{\alpha,\beta} + \frac{f_{\alpha,\beta}}{\sqrt{\mathrm{Tr}(\Pi_\Delta)}}g_{\alpha,\beta}, \tag{2}$$

whenever the energies of the eigenstates $|\nu_\alpha\rangle$, $|\nu_\beta\rangle$ are in the interval $\mathcal{I}_\Delta$. In the off-diagonal entries, $g_{\alpha,\beta} = g^*_{\beta,\alpha}$ are complex i.i.d. Gaussians with mean 0 and variance 1. The real coefficients $f_{\alpha,\beta}$ represent flexibility in the variance of the Gaussians, subject to $f_{\alpha,\beta} = f_{\beta,\alpha}$ due to the Hermitian condition and $f \leq f_{\alpha,\beta} \leq 1$, where f is a constant. The diagonal entries $\mu_\alpha = \langle\nu_\alpha| A |\nu_\alpha\rangle$ are non-random, but smoothly vary within the window $\mathcal{I}_\Delta$ (as a function of the eigenenergy). The normalization factor $\mathrm{Tr}(\Pi_\Delta)$ is simply the number of eigenstates within this window, as Π_Δ denotes the projector onto them. All these prescriptions are subject to the (global) constraints $A = A^\dagger$ and $\|A\| = 1$. [4]

We say that H and a collection of local [5] *observables $A_1, \ldots, A_m$ satisfy ETH within the $\mathcal{I}_\Delta$ window around energy e_0 if these operators "look like" randomly drawn from the ETH ensemble.*

A quick objection to the ETH ansatz is - where is the randomness coming from? And what does "look like" mean? Clearly the Hamiltonian is fixed and the statement is being made for all eigenstates in the energy window. The view is that we make a choice of $g_{\alpha\beta}$ by

[3] ETH is usually defined for "extensive" energies, those that scale linearly with the system size. Our threshold $n^{0.99}$ is just a way to capture this without introducing another parameter to set the extensive threshold.

[4] This can be guaranteed by the countering fluctuations in the entries outside of the window $\mathcal{I}_\Delta$.

[5] The locality of A_i is typically allowed up to $o(n)$, but A_i should be simple (like Pauli, or efficiently implementable). For this work, it suffices to think of A_i as constant-local Pauli operators.

drawing from iid Gaussian and then fix them for the entire calculation (and any predictions are taken "with high probability"). Given that H is a complicated object, the belief is that we would be unable to tell a difference (with high probability over the ETH ensemble) if we did calculations using the precise entries of the A's in the energy eigenbasis. Such behaviors have been experimentally and numerically verified in many physical quantum systems [48], as well as generic, random Hamiltonians, such as the famous SYK model, where ETH is postulated to hold, see [21, Section 4.3].

In this work, we take this computational indistinguishability viewpoint to design a unique verification protocol for ETH Hamiltonians, but we hope the formulation below could be useful for other applications of ETH in complexity theory.

▶ **Hypothesis 24** (Computational ETH)**.** *Setup:*

(The Hamiltonian) *Consider an n-qubit local Hamiltonian H and a fixed energy window $\mathcal{I}_\Delta$ around e_0. Suppose that H has non-clustering of eigenstates in the window $\mathcal{I}_\Delta$.*

(The Experiment) *Consider a polynomial-time algorithm $\mathcal{N}$ that takes in an input state $|\xi\rangle$ promised to be inside $\mathcal{I}_\Delta$ and runs a polynomial-sized circuit (possibly with ancillas) that ends with a binary POVM $\{\mathsf{Acc}, \mathrm{I}-\mathsf{Acc}\}$. Furthermore, there is a set of $m = \mathrm{poly}(n)$ local Pauli operators $\mathtt{Alg} = \{A_1, \ldots, A_m\}$, such that $\mathcal{N}$ consists of gates that depend directly as a black-box on $\{A_1, \ldots, A_m\}$ and the individual local terms in H (which for examples could be time evolutions of any of these operators, or measurements, etc.)*

(The ETH Ensemble) *Let $\mathcal{N}'$ be an algorithm obtained by the following procedure: (1) iid sample a set of operators $\mathtt{Alg}' = \{A'_1, \ldots, A'_m\}$ from the operator ensemble defined in Hypothesis 23, (2) replace the components related to $\mathtt{Alg}$ in $\mathcal{N}$ by their $\mathtt{Alg}'$ versions* [6]*. More precisely, within the window $\mathcal{I}_\Delta$, each A'_i has the form*

$$\langle \nu_\alpha | A'_i | \nu_\beta \rangle = \mu^i_\alpha \delta_{\alpha,\beta} + \frac{f_{\alpha,\beta}}{\sqrt{\mathrm{Tr}(\Pi_\Delta)}} g^i_{\alpha,\beta}, \tag{3}$$

where the coefficients and random variables are as described in Hypothesis 23. Note that the diagonal entries μ^i_α are non-random and thus depend on A_i, so that distinct A'_i is sampled from a distinct distribution. The $g^i_{\alpha,\beta}$ are normal complex Gaussian variables sampled iid across different i.

Then, we say H and $\mathcal{N}$ satisfy the computational ETH within the energy window $\mathcal{I}_\Delta$ if it holds, with high probability over the randomness of the ETH ensemble, that

$$\mathrm{Tr}((\mathcal{N}(\xi) - \mathcal{N}'(\xi))\mathsf{Acc}) < \varepsilon_{\mathrm{ETH}} = \mathrm{negl}(n), \qquad \forall \text{ states } \xi \text{ of energy } \in \mathcal{I}_\Delta. \tag{4}$$

The above computational version is an attempt to capture in the computer science language the common physicists' interpretations of ETH. Above, we allow the "experiment" to be beyond BQP: it can be given a proof (untrusted, except for the fact that its energy is within $\mathcal{I}_\Delta$ which can be efficiently verified). This allows us to work in the context of QMA. The $\varepsilon_{\mathrm{ETH}}$ is similar to the distinguishability advantage encountered in cryptography. Here we take it to be $\mathrm{negl}(n)$, but we actually only need $\varepsilon_{\mathrm{ETH}}$ to be a sufficiently small constant to guarantee the performance of our UniqueQMA verifier in this section.

[6] These A' may require exponential circuit complexity, but we think of them as a black box, and these replacements only appears in our calculations.

3.3.2 Verifying a unique state in an energy window satisfying ETH

Given a local Hamiltonian H on n qubits, in this section, we give a protocol that uniquely accepts a specific state in a known energy window around e_0 satisfying certain properties that will be instantiated by ETH.

The intuition behind the protocol comes from tests of entanglement via quantum expanders [10]. More precisely, the EPR state $|\Omega\rangle = \frac{1}{\sqrt{D}} \sum_{i=1}^{D} |ii\rangle$ can be tested locally using unitaries of the form $U \otimes \overline{U}$. We briefly explain the idea: Clearly, $U \otimes \overline{U}|\Omega\rangle = |\Omega\rangle$ for any unitary $U \in SU(D)$ so any averaged operator over the actions $U_i \otimes \overline{U}_i$ for unitaries $U_i \in SU(D), i = 1, \ldots, M$ satisfies $\frac{1}{M} \sum_{i=1}^{M} U_i \otimes \overline{U}_i |\Omega\rangle = |\Omega\rangle$. A quantum expander is roughly a collection of unitaries U_i, such that the second highest eigenvalue of $\frac{1}{M} \sum_{i=1}^{M} U_i \otimes \overline{U}_i$ is bounded away from 1. Such quantum expander families exist with surprisingly small M and constant gap between second highest eigenvalue and 1 (see e.g. [27, 29]). Another common approach, used e.g. for self-testing, is to pick the local Pauli operators, which results in an inverse polynomial gap between the second highest eigenvalue and 1 (see e.g. [46]). The expectation value $\frac{1}{M} \sum_{i=1}^{M} \langle\psi| U_i \otimes \overline{U}_i |\psi\rangle$ can be efficiently estimated in many settings and suggests a simple test: If the above expectation is close to 1, then $|\psi\rangle$ is close to $|\Omega\rangle$. Otherwise, $|\psi\rangle$ and $|\Omega\rangle$ are far from each other. In particular, $|\Omega\rangle$ is uniquely accepted by this text. A plausible strategy is therefore to use this principle to uniquely verify the maximally entangled state on two copies of the space spanned by the eigenstates in an energy window. It will turn out that we cannot actually guarantee that the uniquely accepted state is the maximally entangled state on these subspaces. But for our purposes it will suffice that the Perron-Frobenius theorem guarantees the existence of a non-degenerate highest eigenvalue for an eigenstate with non-negative entries. Then, Merlin can simply send this state and Arthur can uniquely verify the ground state energy via a quantum expander.

Two obstacles prevent us from applying this naively: 1) A general unitary will move not respect the subspace spanned by eigenstates with energies in a window and 2) even if we can apply such unitaries we need to bound the gap of the averaging operator to guarantee uniqueness. We solve 1) by considering unitaries of the form $e^{-\iota\varepsilon A}$ for sufficiently small, but constant $\varepsilon > 0$, which approximately preserve subspaces. We then use phase estimation to prepare measurements of the energy window and effectively apply the projector onto the respective subspaces. In order to approach 2) we invoke the ETH: By choosing the A_i to be local (think local Pauli matrices), their matrix entries in the energy window are indistinguishable from the Gaussians in Hypothesis 23. We then use probabilistic techniques to show with high probability over the Gaussians A_i' that the operators $\mathbb{E}_i e^{-\iota\varepsilon A_i} \otimes e^{\iota\varepsilon A_i}$ are sufficiently gapped. In particular, for the Gaussian A_i' we obtain a gap and, consequently, unique verification. The ETH then guarantees that any quantum algorithm using the correct A_i will have the same behavior.

Finally, we assume there is a known energy e_0 and interval $\mathcal{I}_\Delta$ such that the above properties holds.

▶ **Assumption 1** (Existence of "good" Energy). *There exists a known energy $e_0 \in (0, 1)$ and value Δ such that Hypothesis 24 holds. Furthermore, e_0 can be represented by $O(\log n)$ bits.*

Given the existence of this energy window, we use quantum phase estimation to approximately implement the projector Π_Δ onto this subspace. In the algorithm, we use Π to indicate this projector, with subscripts indicating the registers on which the projector acts.

Algorithm 1 Energy subspace test $\mathtt{Alg}$.

Input: state $|\psi\rangle$ on the registers $\mathsf{R}^{(1)}_{\mathrm{sys}}, \mathsf{R}^{(2)}_{\mathrm{sys}}$.

1. Prepare and append $|\mu\rangle_{\mathsf{R}^{(1)}_{\mathrm{aux}}} |\mu\rangle_{\mathsf{R}^{(2)}_{\mathrm{aux}}}$ to $|\psi\rangle_{\mathsf{R}^{(1)}_{\mathrm{sys}},\mathsf{R}^{(2)}_{\mathrm{sys}}}$, where $|\mu\rangle = \frac{1}{\sqrt{L}} \sum_{k=0}^{L-1} |k\rangle$.
2. Measure $\{\Pi_{\mathsf{R}^{(1)}_{\mathrm{sys}}\mathsf{R}^{(1)}_{\mathrm{aux}}}, \mathrm{I} - \Pi_{\mathsf{R}^{(1)}_{\mathrm{sys}}\mathsf{R}^{(1)}_{\mathrm{aux}}}\}$, $\{\Pi_{\mathsf{R}^{(2)}_{\mathrm{sys}}\mathsf{R}^{(2)}_{\mathrm{aux}}}, \mathrm{I} - \Pi_{\mathsf{R}^{(2)}_{\mathrm{sys}}\mathsf{R}^{(2)}_{\mathrm{aux}}}\}$ and reject if $\mathrm{I} - \Pi_{\mathsf{R}^{(1)}_{\mathrm{sys}}\mathsf{R}^{(1)}_{\mathrm{aux}}}$ or $\mathrm{I} - \Pi_{\mathsf{R}^{(2)}_{\mathrm{sys}}\mathsf{R}^{(2)}_{\mathrm{aux}}}$ is obtained.
3. Measure $\mathsf{R}^{(1)}_{\mathrm{aux}}$ in the basis $\{\Pi_{\mathsf{R}^{(1)}_{\mathrm{aux}}}, \mathrm{I} - \Pi_{\mathsf{R}^{(1)}_{\mathrm{aux}}}\}$, $\mathsf{R}^{(2)}_{\mathrm{aux}}$ in the basis $\{\Pi_{\mathsf{R}^{(2)}_{\mathrm{aux}}}, \mathrm{I} - \Pi_{\mathsf{R}^{(2)}_{\mathrm{aux}}}\}$ and reject if $\mathrm{I} - \Pi_{\mathsf{R}^{(1)}_{\mathrm{aux}}}$ or $\mathrm{I} - \Pi_{\mathsf{R}^{(2)}_{\mathrm{aux}}}$ is obtained.
4. Prepare a register $\mathsf{R}_{\mathrm{control}}$ and define the state $\frac{1}{\sqrt{2\ell}} \sum_i (|i,0\rangle_{\mathsf{R}_{\mathrm{control}}} + |i,1\rangle_{\mathsf{R}_{\mathrm{control}}})$. Controlled on $|i,0\rangle\langle i,0|_{\mathsf{R}_{\mathrm{control}}}$, apply $e^{\iota\varepsilon A_i}$ on $\mathsf{R}^{(1)}_{\mathrm{sys}}$ and $e^{-\iota\varepsilon \overline{A}_i}$ on $\mathsf{R}^{(2)}_{\mathrm{sys}}$. Controlled on $|i,1\rangle\langle i,1|_T$, apply $e^{-\iota\varepsilon A_i}$ on $\mathsf{R}^{(1)}_{\mathrm{sys}}$ and $e^{\iota\varepsilon \overline{A}_i}$ on $\mathsf{R}^{(2)}_{\mathrm{sys}}$.
5. Repeat Steps 2,3.
6. Measure $\mathsf{R}_{\mathrm{control}}$ to see if the register is in $\frac{1}{\sqrt{2\ell}} \sum_i (|i,0\rangle_{\mathsf{R}_{\mathrm{control}}} + |i,1\rangle_{\mathsf{R}_{\mathrm{control}}})$. Reject if it is not.

We show that the above algorithm indeed succeeds in verifying the existence of a unique ground state.

▶ **Theorem 25** (Energy Subspace Test). *Fix a parameter e_0 and consider a Hamiltonian H. There exists an algorithm $\mathtt{Alg} = \mathtt{Alg}(H, e_0, L)$ (where $L = \mathrm{poly}(n)$ is a precision parameter) such that*

1. *If $\lambda_0(H) > e_0 + \frac{3}{L}$ then $\mathtt{Alg}$ accepts with probability at most $\frac{1}{3}$.*
2. *If $\lambda_0(H) \leq e_0$ and e_0 satisfies Assumption 1 w.r.t. H, then there exists a unique state $|\Psi\rangle$ such that $\mathtt{Alg}$ accepts on input $|\Psi\rangle$ with probability at least $\frac{2}{3}$, and on states orthogonal to $|\Psi\rangle$, $\mathtt{Alg}$ accepts with probability at most $\frac{1}{3}$.*

4 Physical relevance

On the surface, our first result can be viewed as evidence that UniqueQMA is not equal to QMA. An alternative interpretation is via a characterization due to [9]; in that paper, the authors show that the class of local Hamiltonians with an inverse polynomial spectral gap is UniqueQMA complete. Thus, our oracle separations give evidence that the local Hamiltonian problem is easier for inverse polynomially gapped instances[7]. This is interesting from a physics perspective as many classes of physically relevant Hamiltonians are known to have an inverse polynomial spectral gap (e.g., ferromagnetic Heisenberg [38], 1D AKLT [5], Movassagh-Shor [42]). In particular, the presence of a spectral gap in these models immediately makes studying various properties more tractable [47, 40]. Thus our work provides some evidence for the physical intuition that inverse polynomially gapped Hamiltonians are "easy" by separating such Hamiltonians from QMA, under an oracle.

Our third result has implications for the Eigenstate Thermalization Hypothesis. If UniqueQMA and QMA are not equivalent (under quantum reductions), then hard instances of the local Hamiltonian problem cannot satisfy ETH in its low energy window (near-extensive). This violation of ETH is robust against adversarial perturbations of near-extensive strength. Under the quantum PCP conjecture [8], which states that there are local Hamiltonians such that ground energy estimation is QMA-hard even up to extensively scaling accuracy, the

[7] A more recent work [22] also studied the role of the spectral gap in small promise gap regimes.

latter statement can be extended to extensive energy scales. In other words, if UniqueQMA and QMA are not equivalent, then quantum PCP Hamiltonians need to robustly violate ETH. Note that the same assumption implies QMA-hard Hamiltonians are not many-body localized [44] either. Thus, our results suggest that these Hamiltonians may correspond to an interesting phase of matter that is robust against adversarial perturbations.

We expect our unique verification protocol to work for families of random quantum Hamiltonians such as random spin systems [24] or the SYK model [49, 34, 35, 30]. If random quantum Hamiltonians satisfy a sufficiently strong variant of ETH, our result implies that the average-case is considerably easier than the worst-case, under the assumption UniqueQMA and QMA are not equivalent. Indeed, randomized versions of NP-hard problems such as optimizing the Sherrington-Kirkpatrick model [52] have been proven to be average-case easy [41].

5 Open problems

There are multiple avenues to continue this work. Here, we list a couple:

- Mulmuley, Vazirani, and Vazirani [43] introduced a variant of the isolation lemma in [59], by introducing random weights to the edges of a graph to isolate a unique maximally weighted clique. The quantum analogue of this approach is to consider the quantum variant of the Clique problem, which is QMA hard [14]. It seems to suffer with issues analogous to those raised in [9]: introducing poly(n) amount of randomness does not suffice to single out a unique quantum witness. Can the ideas in [43] still be useful in some quantum problems?
- Theorem 2 demonstrates a quantum oracle to which $\mathsf{QXC}^{\mathcal{O}} \not\subseteq \mathsf{QMA}^{\mathsf{QMA}^{\mathcal{O}}}$. We expect our tools can be extended to showing separations for constants stacks of QMA (i.e. for the QMA hierarchy [2])? Extending our argument to, e.g., $\mathsf{QMA}^{\mathsf{QMA}^{\mathsf{QMA}}}$ would require 1) showing polarization for $\mathsf{QMA}^{\mathsf{QMA}}$ circuits, and 2) a two-sided query lower bound, along the lines of [51], but for relativized classes.
- Theorem 3 constructs a quantum algorithm that verifies some low energy state under ETH. Can we verify a more natural low energy state assuming ETH, such as the purification of Gibbs state (called the thermofield double state)? Note that, due to the Feynman-Kitaev circuit-to-Hamiltonian mapping [36], uniquely verifying a state is equivalent to the state being a near ground state of a local Hamiltonian with inverse polynomial spectral gap. Such local Hamiltonians for the thermofield double state have been constructed under physics-motivated assumptions in [20].
- Our results hint at a separation between worst-case and average-case local Hamiltonians. Random spin Hamiltonians [24] or the SYK model [49, 34, 35] are expected to satisfy the ETH and are therefore solvable by a UniqueQMA machine. It is interesting to find more evidence for the easiness of average-case Hamiltonians. Recent work studies the complexity of approximating the maximal energy of the SYK model. Ref. [30] provides a quantum algorithm to prepare a state with energy $c\sqrt{n}$ – a constant fraction of the expected optimum. Ref. [11] finds further evidence that estimating the energy of the SYK might even be quantumly easy. More concretely, they show that the annealed and quenched free energies agree at inverse polynomial temperatures. This is in contrast to classical random spin systems where these quantities disagree when the system is in a "glassy" phase that is algorithmically hard.

- It is interesting to find more connections between quantum many-body physics assumptions such as ETH and quantum computer science. The works [19, 54] show that ETH implies fast preparation of Gibbs quantum states, if there is no bottleneck throughout the eigenvalue distribution and Gibbs distribution. However, the latter assumption cannot be justified on general grounds. On the other hand, our results work perfectly fine without this assumption.

References

1 Scott Aaronson. On perfect completeness for qma. *arXiv preprint arXiv:0806.0450*, 2008.

2 Scott Aaronson, DeVon Ingram, and William Kretschmer. The acrobatics of bqp. In *37th Computational Complexity Conference*, 2022.

3 Scott Aaronson, Robin Kothari, William Kretschmer, and Justin Thaler. Quantum lower bounds for approximate counting via laurent polynomials. In Shubhangi Saraf, editor, *35th Computational Complexity Conference, CCC 2020, Saarbrücken, Germany (Virtual Conference), July 28-31, 2020*, volume 169 of *LIPIcs*, pages 7:1–7:47. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.CCC.2020.7.

4 Scott Aaronson and Greg Kuperberg. Quantum versus classical proofs and advice. *Theory Comput.*, 3(1):129–157, 2007. doi:10.4086/TOC.2007.V003A007.

5 Ian Affleck, Tom Kennedy, Elliott H Lieb, and Hal Tasaki. Rigorous results on valence-bond ground states in antiferromagnets. *Condensed Matter Physics and Exactly Soluble Models: Selecta of Elliott H. Lieb*, pages 249–252, 2004.

6 Avantika Agarwal and Shalev Ben-David. Oracle separations for the quantum-classical polynomial hierarchy. *arXiv preprint arXiv:2410.19062*, 2024. doi:10.48550/arXiv.2410.19062.

7 Avantika Agarwal and Srijita Kundu. A cautionary note on quantum oracles. *arXiv preprint arXiv:2504.19470*, 2025. doi:10.48550/arXiv.2504.19470.

8 Dorit Aharonov, Itai Arad, and Thomas Vidick. Guest column: the quantum pcp conjecture. *Acm sigact news*, 44(2):47–79, 2013. doi:10.1145/2491533.2491549.

9 Dorit Aharonov, Michael Ben-Or, Fernando G. S. L. Brandão, and Or Sattath. The pursuit of uniqueness: Extending valiant-vazirani theorem to the probabilistic and quantum settings. *Quantum*, 6:668, 2022. doi:10.22331/Q-2022-03-17-668.

10 Dorit Aharonov, Aram W. Harrow, Zeph Landau, Daniel Nagaj, Mario Szegedy, and Umesh V. Vazirani. Local tests of global entanglement and a counterexample to the generalized area law. In *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18-21, 2014*, pages 246–255. IEEE Computer Society, 2014. doi:10.1109/FOCS.2014.34.

11 Eric R Anshuetz, Chi-Fang Chen, Bobak T Kiani, and Robbie King. Strongly interacting fermions are non-trivial yet non-glassy. *arXiv preprint arXiv:2408.15699*, 2024.

12 Itai Arad, Tomotaka Kuwahara, and Zeph Landau. Connecting global and local energy distributions in quantum spin models on a lattice. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(3):033301, 2016. doi:10.1088/1742-5468/2016/03/033301.

13 Richard Beigel, Harry Buhrman, and Lance Fortnow. Np might not be as easy as detecting unique solutions. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pages 203–208, 1998. doi:10.1145/276698.276737.

14 Salman Beigi and Peter W. Shor. On the complexity of computing zero-error and holevo capacity of quantum channels, 2008. arXiv:0709.2090.

15 Shai Ben-David, Benny Chor, and Oded Goldreich. On the theory of average case complexity. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 204–216, 1989.

16 Sergey Bravyi, Anirban Chowdhury, David Gosset, Vojtech Havlicek, and Guanyu Zhu. Quantum complexity of the kronecker coefficients. *PRX Quantum*, 5(1):010329, 2024.

17 Sergey Bravyi, Anirban Chowdhury, David Gosset, and Pawel Wocjan. On the complexity of quantum partition functions. *CoRR*, abs/2110.15466, 2021. arXiv:2110.15466.

18 Brielin Brown, Steven T Flammia, and Norbert Schuch. Computational difficulty of computing the density of states. *Physical review letters*, 107(4):040501, 2011.

19 Chi-Fang Chen and Fernando G. S. L. Brandão. Fast Thermalization from the Eigenstate Thermalization Hypothesis. *arXiv preprint arXiv:2112.07646*, 2023.

20 William Cottrell, Ben Freivogel, Diego M. Hofman, and Sagar F. Lokhande. How to build the thermofield double state. *Journal of High Energy Physics*, 2019(2):58, February 2019. doi:10.1007/JHEP02(2019)058.

21 Luca D'Alessio, Yariv Kafri, Anatoli Polkovnikov, and Marcos Rigol. From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics. *Advances in Physics*, 65(3):239–362, 2016.

22 Abhinav Deshpande, Alexey V Gorshkov, and Bill Fefferman. Importance of the spectral gap in estimating ground-state energies. *PRX Quantum*, 3(4):040327, 2022.

23 Josh M Deutsch. Quantum statistical mechanics in a closed system. *Physical review a*, 43(4):2046, 1991.

24 László Erdős and Dominik Schröder. Phase transition in the density of states of quantum spin glasses. *Mathematical Physics, Analysis and Geometry*, 17(3):441–464, 2014.

25 Friedrich Götze and Holger Sambale. Higher order concentration on stiefel and grassmann manifolds. *Electronic Journal of Probability*, 28:1–30, 2023.

26 Joachim Grollmann and Alan L Selman. Complexity measures for public-key cryptosystems. *SIAM Journal on Computing*, 17(2):309–335, 1988. doi:10.1137/0217018.

27 David Gross and Jens Eisert. Quantum margulis expanders. *arXiv preprint arXiv:0710.0651*, 2007.

28 Dominik Hangleiter and Jens Eisert. Computational advantage of quantum random sampling. *Reviews of Modern Physics*, 95(3):035001, 2023.

29 Matthew B Hastings. Random unitaries give quantum expanders. *Physical Review A—Atomic, Molecular, and Optical Physics*, 76(3):032315, 2007.

30 Matthew B Hastings and Ryan O'Donnell. Optimizing strongly interacting fermionic hamiltonians. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 776–789, 2022. doi:10.1145/3519935.3519960.

31 Sandy Irani, Anand Natarajan, Chinmay Nirkhe, Sujit Rao, and Henry Yuen. Quantum search-to-decision reductions and the state synthesis problem. *arXiv preprint arXiv:2111.02999*, 2021. arXiv:2111.02999.

32 Rahul Jain, Iordanis Kerenidis, Greg Kuperberg, Miklos Santha, Or Sattath, and Shengyu Zhang. On the power of a unique quantum witness. *Theory Comput.*, 8(1):375–400, 2012. doi:10.4086/TOC.2012.V008A017.

33 Mark R Jerrum, Leslie G Valiant, and Vijay V Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical computer science*, 43:169–188, 1986. doi:10.1016/0304-3975(86)90174-X.

34 Alexei Kitaev. A simple model of quantum holography (part 1). *Entanglement in strongly-correlated quantum matter*, page 38, 2015.

35 Alexei Kitaev. A simple model of quantum holography (part 2). *Entanglement in strongly-correlated quantum matter*, page 38, 2015.

36 Alexei Y. Kitaev, A. H. Shen, and Mikhail N. Vyalyi. *Classical and Quantum Computation*, volume 47 of *Graduate studies in mathematics*. American Mathematical Society, 2002. URL: https://bookstore.ams.org/gsm-47/.

37 Adam R Klivans and Dieter Van Melkebeek. Graph nonisomorphism has subexponential size proofs unless the polynomial-time hierarchy collapses. In *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*, pages 659–667, 1999. doi:10.1145/301250.301428.

38 Tohru Koma and Bruno Nachtergaele. The spectral gap of the ferromagnetic xxz-chain. *Letters in Mathematical Physics*, 40(1):1–16, 1997.

39 Greg Kuperberg. How hard is it to approximate the jones polynomial? *arXiv preprint arXiv:0908.0512*, 2009. arXiv:0908.0512.

40 Zeph Landau, Umesh Vazirani, and Thomas Vidick. A polynomial time algorithm for the ground state of one-dimensional gapped local hamiltonians. *Nature Physics*, 11(7):566–569, 2015.

41 Andrea Montanari. Optimization of the sherrington-kirkpatrick hamiltonian. *SIAM J. Comput.*, 54(4):S19–1, 2025. doi:10.1137/20M132016X.

42 Ramis Movassagh and Peter W Shor. Power law violation of the area law in quantum spin chains. *arXiv preprint arXiv:1408.1657*, 2014.

43 Ketan Mulmuley, Umesh V. Vazirani, and Vijay V. Vazirani. Matching is as easy as matrix inversion. *Combinatorica*, 7(1):105–113, March 1987. doi:10.1007/BF02579206.

44 Rahul Nandkishore and David A Huse. Many-body localization and thermalization in quantum statistical mechanics. *Annu. Rev. Condens. Matter Phys.*, 6(1):15–38, 2015.

45 Anand Natarajan and Chinmay Nirkhe. A distribution testing oracle separation between qma and qcma. *Quantum*, 8:1377, 2024. doi:10.22331/Q-2024-06-17-1377.

46 Anand Natarajan and Thomas Vidick. Robust self-testing of many-qubit states. *arXiv preprint arXiv:1610.03574*, 2016. arXiv:1610.03574.

47 Román Orús. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of physics*, 349:117–158, 2014.

48 Marcos Rigol, Vanja Dunjko, and Maxim Olshanii. Thermalization and its mechanism for generic isolated quantum systems. *Nature*, 452(7189):854–858, 2008.

49 Subir Sachdev and Jinwu Ye. Gapless spin-fluid ground state in a random quantum heisenberg magnet. *Physical review letters*, 70(21):3339, 1993.

50 Norbert Schuch, Michael M. Wolf, Frank Verstraete, and J. Ignacio Cirac. Computational complexity of projected entangled pair states. *Phys. Rev. Lett.*, 98:140506, April 2007. doi:10.1103/PhysRevLett.98.140506.

51 Adrian She and Henry Yuen. Unitary property testing lower bounds by polynomials. In *Leibniz International Proceedings in Informatics (LIPIcs): 14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.ITCS.2023.96.

52 David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Physical review letters*, 35(26):1792, 1975.

53 Yaoyun Shi and Shengyu Zhang. Note on quantum counting classes, 2009. URL: http://www.cse.cuhk.edu.hk/~syzhang/papers/SharpBQP.pdf.

54 Oles Shtanko and Ramis Movassagh. Preparing thermal states on noiseless and noisy programmable quantum processors, 2023. arXiv:2112.14688.

55 Julian Sonner and Manuel Vielma. Eigenstate thermalization in the sachdev-ye-kitaev model. *Journal of High Energy Physics*, 2017(11):149, November 2017. doi:10.1007/JHEP11(2017)149.

56 Mark Srednicki. The approach to thermal equilibrium in quantized chaotic systems. *Journal of Physics A: Mathematical and General*, 32(7):1163, 1999.

57 Larry Stockmeyer. The complexity of approximate counting. In *Proceedings of the fifteenth annual ACM symposium on Theory of computing*, pages 118–126, 1983.

58 Seinosuke Toda. Pp is as hard as the polynomial-time hierarchy. *SIAM Journal on Computing*, 20(5):865–877, 1991. doi:10.1137/0220053.

59 Leslie G. Valiant and Vijay V. Vazirani. NP is as easy as detecting unique solutions. *Theor. Comput. Sci.*, 47(3):85–93, 1986. doi:10.1016/0304-3975(86)90135-0.

60 Mark Zhandry. Toward Separating QMA from QCMA with a Classical Oracle, 2024. arXiv:2411.01718 [quant-ph]. doi:10.48550/arXiv.2411.01718.