A distribution testing oracle separation between QMA and QCMA

Anand Natarajan[1] and Chinmay Nirkhe[2]

[1]Massachusetts Institute of Technology
[2]IBM Quantum Cambridge

It is a long-standing open question in quantum complexity theory whether the definition of *non-deterministic* quantum computation requires quantum witnesses (QMA) or if classical witnesses suffice (QCMA). We make progress on this question by constructing a randomized classical oracle separating the respective computational complexity classes. Previous separations [3, 13] required a quantum unitary oracle. The separating problem is deciding whether a distribution supported on regular un-directed graphs either consists of multiple connected components (yes instances) or consists of one expanding connected component (no instances) where the graph is given in an adjacency-list format by the oracle. Therefore, the oracle is a distribution over n-bit boolean functions.

1 Introduction

There are two natural *quantum* analogs of the computational complexity class NP. The first is the class QMA in which a quantum polynomial-time decision algorithm is given access to a $\mathsf{poly}(n)$ *qubit* quantum state as a witness for the statement. This class is captured by the QMA-complete local Hamiltonian problem [18] in which the quantum witness can be interpreted as the ground-state of the local Hamiltonian. The second is the class QCMA in which the quantum polynomial-time decision algorithm is given access instead to a $\mathsf{poly}(n)$ *bit classical* state. While it is easy to prove that $\mathsf{QCMA} \subseteq \mathsf{QMA}$ as the quantum witness state can be immediately measured to yield a classical witness string, the question of whether $\mathsf{QCMA} \stackrel{?}{=} \mathsf{QMA}$, first posed by Aharonov and Naveh [4], remains unanswered. If $\mathsf{QCMA} = \mathsf{QMA}$, then every local Hamiltonian would have an efficient classical witness of its ground energy; morally, this can be thought of as an efficient classical description of its ground state. The relevance of local Hamiltonians to condensed matter physics makes this question a central open question in quantum complexity theory [2].

Anand Natarajan: anandn@mit.edu, This work was partially completed while a participant in the Simons Institute for the Theory of Computing program *The Quantum Wave in Computing: Extended Reunion*. Natarajan thanks Elizabeth Crosson, Aram Harrow, Zhiyang He, Robin Kothari, Yupan Liu, and Mehdi Soleimanifar for helpful discussions.

Chinmay Nirkhe: nirkhe@ibm.com, Some of the initial ideas of this work were done while affiliated with the University of California, Berkeley. This work was partially completed while a participant in the Simons Institute for the Theory of Computing program *The Quantum Wave in Computing: Extended Reunion*. Nirkhe thanks Srinivasan Arunachalam, Andrew Childs, Yi-Kai Liu, William Kretschmer, Kunal Marwaha, Umesh Vazirani, and Elizabeth Yang for helpful discussions.

Because $\mathsf{P} \subseteq \mathsf{QCMA} \subseteq \mathsf{QMA} \subseteq \mathsf{PSPACE}$, any unconditional separation of the two complexity classes would imply $\mathsf{P} \neq \mathsf{PSPACE}$ and seems unlikely without remarkably ingenious new tools. A more reasonable goal is an oracle separation between the two complexity classes. The first oracle separation, by Aaronson and Kuperberg [3], showed that there exists a black-box unitary problem for which quantum witnesses suffice and yet no polynomial sized classical witness and algorithm can solve the problem with even negligible success probability. A second black-box separation was discovered a decade later by Fefferman and Kimmel [13]. The Fefferman and Kimmel oracle is a completely positive trace perseving (CPTP) map called an "in-place" permutation oracle. Both oracles [3, 13] are inherently quantum[1]. Whereas, the "gold-standard" of oracle separations — namely black-box function separations (also known as classical oracle separations) — only require access to a *classical function* that can be queried in superposition[2].

1.1 Graph oracles

The major result of this work is to prove that there exists a distribution over black-box function problems separating QMA and QCMA. Each black-box function corresponds to the adjacency list of a $N \stackrel{\text{def}}{=} 2^n$ vertex constant-degree colored graphs[3] $G = (V, E)$. Roughly speaking, a graph is a YES instance if the second eigenvalue of its normalized adjacency matrix is 1 (equivalently, if it has at least two connected components) and a graph is a NO instance if it second eigenvalue is at most $1-\alpha$ for some fixed constant α (equivalently, the graph has one connected component and is expanding). We call this problem the *expander distinguishing problem*.

Distribution oracles A distribution over functions (equivalently, a distribution over graphs) is a YES instance if it is entirely supported on YES graphs and a distribution over functions is a NO instance if it is entirely supported on NO graphs.

In this work, we construct, for every n, families of YES and NO distributions over graphs such that following hold for the promise problem of distinguishing a graph sampled from a YES distribution from a graph sampled from a NO distribution.

1. There is a QMA proof system that solves this problem, where the verifier runs in quantum polynomial time and has black-box query access to the sampled graph, and the honest prover's (quantum) witness depends only on the distribution, not on the specific sample.

2. No QCMA proof system can solve this problem, provided the prover's (classical) witness is only allowed to depend on the distribution, and not on the sample.

Our work is not the first to consider oracles that sample from distributions over functions. The in-place oracle separation of [13] between QMA and QCMA used oracles that sampled random permutations. For a somewhat different problem, of separating

[1] It might be reasonable to wonder if the unitary oracles can be converted into classical oracles by providing oracle access to the exponentially long classical descriptions of the respective matrices. This is not known to be true because it is unclear how to use access to the classical description to solve the QMA problem.

[2] One reason this model is natural is that if we were given a circuit of size C to implement this classical function, then we would automatically get a quantum circuit of size C to implement the oracle, simply by running the classical circuit coherently. This is not true for the "in-place" permutation oracle model, assuming that one-way functions exist.

[3] A similar problem was previously conjectured to be an oracle separation for these complexity classes by Lutomirski [22].

Authors	Separating black box object	Proof techniques used
Aaronson & Kuperberg [3]	n-qubit unitaries	Adversary method
Fefferman & Kimmel [13]	n-qubit CPTP maps	Combinatorial argument, Adversary method
This work	Distributions over n-bit boolean functions	Combinatorial argument, Adversary method, Polynomial method
Conjectured	n-bit boolean function	?

Figure 1: List of known oracle separations

bounded-depth quantum-classical circuits, [8] introduced a related notion called a "stochastic oracle"—the main difference between this and our model is that a stochastic oracle resamples an instance every time it is queried.

Comparison with previous oracle separations between QMA and QCMA Figure 1 summarizes our work in relation to previous oracle separations. In terms of results, we take a further step towards the standard oracle model—all that remains is to remove the randomness from our oracle. In terms of techniques, we combine the use of counting arguments and the adversary method from previous works with a BQP lower bound for a similar graph problem, due to [6]. This lower bound was shown using the polynomial method. We view the judicious combination of these lower bound techniques—as simple as it may seem—as one of the conceptual contributions of this paper.

Intuition for hardness The expander distinguishing problem is a natural candidate for a separation between QMA and QCMA because it is an "oracular" version of the sparse Hamiltonian problem, which is complete for QMA [10, Problem H-4]. To see this, we recall some facts from spectral graph theory. The top eigenvalue of the normalized adjacency matrix A for regular graphs is always 1 and the uniform superposition over vertices is always an associated eigenvector. If the graph is an expander (the NO case of our problem), the second eigenvalue is bounded away from 1, but if the graph is disconnected (the YES case of our problem), then the second eigenvalue is exactly 1. Thus, our oracle problem is exactly the problem of estimating the minimum eigenvalue of $\mathbb{I} - A$ (a sparse matrix for a constant-degree graph), on the subspace orthogonal to the uniform superposition state. Viewing $\mathbb{I} - A$ as a sparse Hamiltonian, we obtain the connection between our problem and the sparse Hamiltonian problem.

One reason to show oracle separations between two classes is to provide a *barrier* against attempts to collapse the classes in the "real" world. We interpret our results as confirming the intuition that any QCMA protocol for the sparse Hamiltonian must use more than just black-box access to entries of the Hamiltonian: it must use some nontrivial properties of the ground states of these Hamiltonians. In this sense, it emulates the original quantum adversary lower bound of [9] which showed that any BQP-algorithm for solving NP-complete problems must rely on some inherent structure of the NP-complete problem as BQP-algorithms cannot solve unconstrained search efficiently.

Naturalness of the randomized oracle model Some care must be taken whenever one proves a separation in a "nonstandard" oracle model—see for instance the "trivial" example in [1] of a randomized oracle separating MA_1 from MA. We believe that our randomized oracle model is natural for several reasons. Firstly, as mentioned above, randomization was used in the quantum oracle of [13] for essentially the same reason: to impose a restriction on the witnesses received from the prover. Secondly, it is consistent

with our knowledge that our oracle separates QMA from QCMA even when the randomness is removed (and indeed we conjecture this is the case, as described below.) Thirdly, the randomization still gives the prover access to substantial information about the graph: in particular, the prover knows the full connected component structure of the graph. As we show, this information is enough for the prover to give a *quantum* witness state, that in the YES case convinces the verifier with certainty. Our result shows that even given full knowledge of the component structure, the prover cannot construct a convincing classical witness—we believe this sheds light on how a QMA witness can be more powerful than a QCMA witness.

1.2 Overview of proof techniques

Quantum witnesses and containment in oracular QMA A quantum witness for any YES instance graph is any eigenvector $|\xi\rangle$ of eigenvalue 1 that is orthogonal to the uniform superposition over vertices. The verification procedure is simple: project the witness into the subspace orthogonal to the uniform superposition over vertices, and then perform one step of a random walk along the graph, by querying the oracle for the adjacency matrix in superposition. Verify that the state after the walk step equals $|\xi\rangle$. This is equivalent to a 1-bit phase estimation of the eigenvalue. If a graph is a NO instance, then there does not exist any vector orthogonal to the uniform superposition (the unique eigenvector of value 1) that would pass the previous test.

Whenever, the graph has a connected component of $S \subsetneq V$, then an eigenvector orthogonal to the uniform superposition of eigenvalue 1 exists. When $|S| \ll N$, this eigenvector is very close to $|S\rangle$, the uniform superposition over basis vectors $x \in S$. Notice that this state only depends on the connected component S and not the specific edges of the graph. Furthermore, the state $|S'\rangle$ for any subset S' that approximates S forms a witness that is accepted with high probability.

Lower bound on classical witnesses The difficulty in this problem lies in proving a *lower bound* on the ability for classical witnesses to distinguish YES and NO instances. To prove a lower bound, we argue that any quantum algorithm with access to a polynomial length classical witness must make an exponential number of (quantum) queries to the adjacency list of the graph in order to distinguish YES and NO instances. This, in turn, lower bounds the time complexity of any QCMA algorithm distinguishing YES and NO instances but is actually slightly stronger since we don't consider the computational complexity of the algorithm between queries.

Proving lower bounds when classical witnesses are involved is difficult because the witness could be based on any property of the graph. For example, the classical witness could describe cycles, triangles, etc. contained in the graph — while it isn't obvious why such a witness would be helpful, proving that any such witness is insufficient is a significant challenge. One way to circumvent this difficulty is to first show a lower bound *assuming* some structure about the witness[4], and then "remove the training wheels" by showing that the assumption holds for any good classical witness.

[4]Assuming structure about a witness is a common technique in theoretical computer science and in particular lower bounds for classical witnesses of quantum statements. For example, lower bounds against natural proofs [20]. Another example is the NLTS statement [7] which is about lower bounds for classical witnesses for the ground energy of a quantum Hamiltonian of a particular form: constant-depth quantum circuits.

Lower bound against "subset witnesses" One structure we can assume is that the witness only depends on the set of vertices contained in the connected component S. This is certainly the case for the quantum witness state in eq. (24). Our result shows that any polynomial-length witness only depending on the vertices in S requires an exponential query complexity to distinguish YES and NO instance graphs.

The starting point for this statement is the exponential query lower bound *in the absence of a witness* (i.e. for BQP) for the expander distinguishing problem proven by Ambainis, Childs and Liu [6], using the polynomial method. In [6], the authors define two distributions over constant-degree regular colored graphs: the first is a distribution P_1 over random graphs with overwhelming probability of having a second normalized eigenvalue at most $1 - \epsilon_0$. The second is a distribution P_ℓ over random graphs with overwhelming probability of having ℓ connected components. Since, almost all graphs in P_1 are NO graphs and all graphs in P_ℓ are YES graphs, any algorithm distinguishing YES and NO instances must be able to distinguish the two distributions. We first show that a comparable query lower bound still holds even when the algorithm is given a witness consisting of polynomially many random points F from any one connected component.

Next, we show that if there were a QCMA algorithm where the optimal witness depends only on the set of vertices S in one of the connected components, by a counting argument, there must exist a combinatorial *sunflower* of subsets S that correspond to the same witness string. A *sunflower*, in this context, is a set of subsets such that each subset contains a core $F \subset V$ and every vertex of $V \setminus F$ occurs in a small fraction of subsets. This implies that there exists a BQP algorithm which distinguishes YES instances corresponding to the sunflower from all NO instances. Next, we show using an adversary bound [5], a quantum query algorithm cannot distinguish the distribution of YES instances corresponding to the sunflower from the uniform distribution of YES instances such that the core F is contained in a connected component (the ideal sunflower).

This indistinguishability, along with the previous polynomial method based lower bound, proves that QCMA algorithm — whose witness only depends on the vertices in the connected component — for the expander distinguishing problem must make an exponential number of queries to the graph.

Removing the restriction over witnesses Our proof, thus far, has required the restriction that the witness only depends on the vertices in the connected component. In some sense, this argues that there is an oracle separation between QMA and QCMA if the prover is restricted to being "near-sighted": it cannot see the intricacies of the edge-structure of the graph, but can notice the separate connected components of the graph. If the near-sighted prover was capable of sending quantum states as witnesses, then she can still aid a verifier in deciding the expander distinguishing problem, whereas if she could only send classical witnesses, then she cannot aid a verifier.

It now remains to remove the restriction that the witness can only depend on the vertices in the connected component. We do this by introducing *randomness* into the oracle, precisely designed to "blind" the prover to the local structure of the graph. In the standard oracle setting, the verifier and prover both get access to an oracle $x \in \{0,1\}^N$, and the prover provides either a quantum witness, $|\xi(x)\rangle \in (\mathbb{C}^2)^{\otimes \mathsf{poly}(n)}$ or a classical witness, $\xi(x) \in \{0,1\}^{\mathsf{poly}(n)}$. The verifier then runs an efficient quantum algorithm V^x which takes as input $|\xi(x)\rangle$ (or $\xi(x)$, respectively) and consists of quantum oracle gates applying the unitary transform defined as the linear extension of

$$|i\rangle \mapsto (-1)^{x_i} |i\rangle \text{ for } i \in [N]. \tag{1}$$

We now extend modify this setup slightly. Instead of a single oracle x, we consider a distribution $\mathcal{B}$ over oracles. The prover constructs a quantum witness $|\xi(\mathcal{B})\rangle$ (or a classical witness $\xi(\mathcal{B})$, respectively) based on the distribution $\mathcal{B}$. The verifier then samples a classical oracle $x \leftarrow \mathcal{B}$ from the distribution, and then runs the verification procedure V^x which takes as input $|\xi(\mathcal{B})\rangle$ (or $\xi(\mathcal{B})$, respectively) and applies quantum oracle gates corresponding to x. The success probability of the verifier is taken over the distribution $\mathcal{B}$ and the randomness in the verification procedure.

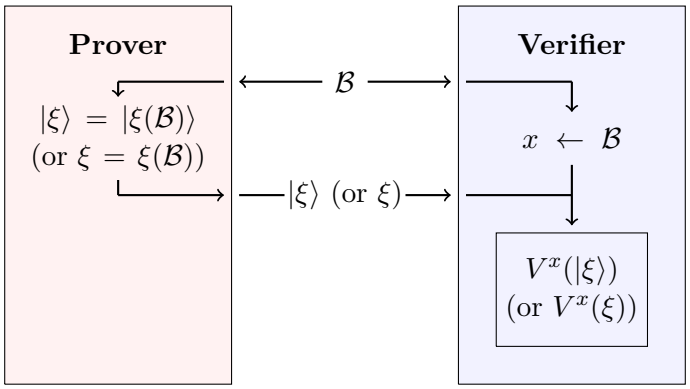

Figure 2: *Cartoon of interaction between Prover and Verifier for a distribution over classical boolean functions.*

From our previous observations, graphs with the same connected component S have the same ideal witness state (given in eq. (24)). So, if the distribution $\mathcal{B}$ is supported on all graphs with the same connected component S, then the witness state from eq. (24) suffices. Furthermore, in the case of the classical witness system, the witness can only depend on S and the previously stated lower bound applies. This motivates the oracle problem of distinguishing distributions, marked either YES or NO, over 2^n bit strings (or equivalently n-bit functions).

1.3 Statement of the result

Theorem 1. *For every sufficiently large integer n that is a multiple of 200, there exist distributions over 100-regular 100-colored graphs on $N = 2^n$ vertices labeled either* YES *or* NO *such that*

- *Each* YES *distribution is entirely supported on* YES *instances of the expander-distinguishing problem and, likewise, each* NO *distribution is entirely supported on* NO *instance of the expander-distinguishing problem.*
- *There exists a* poly(n) *time quantum algorithm V_q taking a witness state $|\xi\rangle$ as input and making $O(1)$ queries to the quantum oracle such that*
  1. *For every* YES *distribution $\mathcal{B}$, there exists a quantum witness $|\xi\rangle \in (\mathbb{C}^2)^{\otimes n}$ such that*
  $$\mathop{\mathbf{E}}_{x\leftarrow\mathcal{B}} \mathbf{Pr}[V_q^x(|\xi\rangle) \text{ accepts}] \geq 1 - O(N^{-3}). \tag{2}$$
  2. *For every* NO *distribution $\mathcal{B}$, for all quantum witnesses $|\xi\rangle \in (\mathbb{C}^2)^{\otimes n}$,*
  $$\mathop{\mathbf{E}}_{x\leftarrow\mathcal{B}} \mathbf{Pr}[V_q^x(|\xi\rangle) \text{ accepts}] \leq 0.01. \tag{3}$$

- *Any quantum algorithm V_c accepting a classical witness of length $q(n)$ satisfying the following two criteria either requires $q(n)$ to be exponential or must make an exponential number of queries to the oracle.*

  1. *For every* YES *distribution $\mathcal{B}$, there exists a classical witness $\xi = \xi(\mathcal{B}) \in \{0,1\}^{q(n)}$*
  $$\mathop{\mathbf{E}}_{x\leftarrow\mathcal{B}} \mathbf{Pr}[V_c^x(\xi) \text{ accepts}] \geq 0.99. \tag{4}$$
  2. *For every* NO *distribution $\mathcal{B}$, for all classical witnesses $\xi \in \{0,1\}^{q(n)}$,*
  $$\mathop{\mathbf{E}}_{x\leftarrow\mathcal{B}} \mathbf{Pr}[V_c^x(\xi) \text{ accepts}] \leq 0.01. \tag{5}$$

Although our main theorem is formulated as a query lower bound, it can be converted to a separation between the relativized classes of QMA and QCMA via a standard diagonalization argument. Similarly, it was pointed out to us [14] that it proves a separation between the relativized classes of BQP/qpoly and BQP/poly, following the technique of [3].

1.4 Implications and future directions

There are several future questions raised by this work that we find interesting:

Oracle and communication separations The most natural question is, of course, whether the oracle's randomness can be removed to obtain a separation in the standard model. We conjecture that our problem yields such a separation, but a new technique seems necessary to prove it. See Section 9 for more details on the technical barriers to derandomizing our construction.

Another natural question is to show a *communication complexity* separation between QMA and QCMA. This has been shown for one-way communication by Klauck and Podder [19] but their problem does not yield a separation for two-way communication. Could our query separation be lifted to the communication world by use of the appropriate gadget?

The class QMA(2) is another relative of QMA which is perhaps even more enigmatic than QCMA. In QMA(2), the witness state is promised to be an unentangled between the first and second half of the qubits. We do not even know of a quantum (unitary) oracle separation between QMA(2) and QMA, nor do we have a natural candidate problem. Could we at least formulate such a candidate by considering "oracular" versions of QMA(2)-complete problems, in analogy to what we do in this work for QCMA.

Search-to-decision In [16], Irani, Natarajan, Nirkhe, Rao and Yuen studied the complexity of generating a witness to a QMA problem (equivalently, generating a ground state of a local Hamiltonian) when given *oracle access* to a QMA oracle. This paradigm, called *search-to-decision*, is commonplace in classical complexity theory (for example, P, NP, MA, etc. all have search-to-decision reductions) and yet [16] gives evidence that QMA likely does not exhibit a search-to-decision reduction. They prove this by showing an oracle relative to which QMA search-to-decision reductions are provably impossible. The oracle used is identical to that of Aaronson and Kuperberg [3] to separate QMA and QCMA. [16] acknowledge this noncoincidence and conjecture whether *any* QMA and QCMA separating oracle yields a QMA search-to-decision impossibility result. Similar to the reasons for why the gold-standard of oracle separation between QMA and QCMA is a n-bit boolean function, the ideal oracle for proving QMA search-to-decision impossibility is also a n-bit boolean function. Does the oracle presented here also yield a search-to-decision impossibility?

Implications for Quantum PCPs The quantum PCP conjecture [4] is one of the biggest open questions in quantum complexity theory. In a recent panel [25] on the quantum PCP conjecture and the NLTS theorem [7], an interesting question was posed of whether MA or QCMA (lower or upper) bounds can be placed on the complexity of the promise-gapped local Hamiltonian problem. We recommend [24] for an introduction to the subject. Because the oracle presented in this result corresponds to a sparse Hamiltonian with a problem of deciding if the second eigenvalue of the Hamiltonian is 1 or $< 1 - \alpha/d = 1 - \Omega(1)$, one might wonder if this provides oracular evidence that quantum PCPs are at least QCMA-hard. Unfortunately, to the best of our knowledge, this is not a reasonable conclusion. While we give evidence that the promise-gapped *sparse* Hamiltonian problem is likely QCMA-hard, the reduction from the sparse Hamiltonian problem to the local Hamiltonian problem does not imply that the promise-gapped local Hamiltonian problem is likely QCMA-hard. The only algorithm known for checking a witness for the sparse Hamiltonian problem is Hamiltonian simulation on the witness which is not a local algorithm.

Connections to Stoquastic Hamiltonians Since the oracles studied in this work correspond to the adjacency lists of graphs, they can be viewed as sparse access to a Hamiltonian H which is the Laplacian of a graph (recall that if the adjacency matrix is A, then the Laplacian is $\mathbb{I} - A/d$). Such Hamiltonians have a special structure not present in general Hamiltonians: they are *stoquastic*, meaning that the off-diagonal entries are non-positive. The local Hamiltonian (LH) problem for stoquastic Hamiltonians is significantly easier than the general LH problem, and in some cases is even contained in MA as shown by Bravyi and Terhal [11]. It is worth noticing why this is not in tension with our result—in particular, why this does not imply that our oracle problem is contained in oracular MA.

- Crucially, the MA-containment for stoquastic LH holds *only* for the ground state: this is because of the Perron-Frobenius theorem, which implies that ground states of such Hamiltonians have nonnegative coefficients. However, in our case, we want the first excited state: the state of minimum energy for H restricted to the subspace orthogonal to the uniform superposition. It was shown by [17] that all excited state energies are QMA-hard to calculate for a stoquastic Hamiltonian.
- The MA containment also uses the locality of the Hamiltonian, which in turn imposes a strong structure on the adjacency matrix of the graph. The random graphs we consider will not have this structure. (While it was shown by [12] showed an AM algorithm for calculating the ground energy *stoquastic and sparse* Hamiltonians, again this does not apply to higher excited states.)
- At an intuitive level, in graph language, the LH problem for stoquastic Hamiltonians is to find a component of the graph where the average value of some *potential function* (given by the diagonal entries of H) is minimized. An MA verifier can solve this by executing a random walk, given the right starting point by Merlin. In contrast, our problem is to determine whether the graph as a whole is connected—a global property which an MA verifier cannot determine.

2 Organization of the paper

The remainder of the paper is the proof of Theorem 1. The proof is divided into smaller components and these intermediate results are joined together in Section 8. In Section 3,

we state some basic definitions and formally define the expander distinguishing problem. In Section 4, we describe the distributions over graphs that constitute YES and NO instances. In Section 5, we prove that there is an efficient QMA algorithm for the expander distinguishing problem. In particular, there is a single quantum witness that serves all the graphs in each of the YES distributions. In Section 6, we use the adversary method and counting arguments to prove that any QCMA algorithm for the expander distinguishing problem for the constructed distributions implies a BQP algorithm for distinguishing YES instances with a connected component corresponding to an ideal sunflower from a generic NO instance. In Section 7, we argue using the polynomial method that such an algorithm is impossible without an exponential query complexity. In Section 9, we present some concluding remarks about our construction and its relation to other notions of computational complexity. Appendices A and B consist of omitted proofs.

3 Preliminaries

3.1 Notation and quantum information basics

We will assume that the reader is familiar with the basics of quantum computing and quantum information. We will use $N \stackrel{\text{def}}{=} 2^n$ throughout this paper and we will only consider graphs of N vertices. The adjacency list of a d-regular d-colored graph on N vertices takes dnN bits to describe. For any m, we abbreviate the set of integers $\{1, 2, \ldots m\}$ as $[m]$. For a set $A \subseteq [N]$, we will use $|A\rangle$ to denote the state $\frac{1}{\sqrt{A}} \sum_{j \in A} |j\rangle$, the subset state corresponding to A. Unless otherwise, specified we assume $\|\cdot\|$ is the Euclidean norm $\|\cdot\|_2$ for a vector, and the spectral norm for a matrix, which is the largest singular value.

3.2 Expander graphs

Definition 2. *A graph G is a* spectral α-expander *(equiv. is α-expanding) if the second highest eigenvalue λ_2 of the normalized adjacency matrix of G satisfies $\lambda_2 \leq 1 - \alpha$. We say that a connected component S of the graph is α-expanding if the restricted graph to the vertices of S is α-expanding.*

Lemma 3. *Let G be a d-regular α-expander. Consider the random walk that starts in any distribution over the vertices, and at each time step, stays in place with probability $1/2$, and moves along an edge of the graph with probability $1/2$. Then for any vertex v, after ℓ steps, the probability $\mathbf{Pr}[v]$ that the walk is in v satisfies*

$$\left|\mathbf{Pr}[v] - \frac{1}{N}\right| \leq \left(1 - \frac{\alpha}{2}\right)^{\ell}. \tag{6}$$

In particular, when $\ell = O(c \log N/\alpha)$ we can get the RHS to be $1/N^c$.

Proof. Let the normalized adjacency matrix of G be A, and let $A' = \frac{1}{2}(\mathbb{I} + A)$ be the transition matrix of the random walk. If G is a d-regular α-expander then $(1/N)\mathbf{1}$ is the unique eigenvector of A (and A') with eigenvalue 1. Moreover, since $\|A\| \leq 1$, all eigenvalues of A' are nonnegative. Since G is an α-expander, the second eigenvalue of A is at most $1 - \alpha$, and thus the second eigenvalue of A' is at most $1 - \alpha/2$.

Let u be a vector representing a probability distribution over vertices of G (i.e. $u \in \mathbb{R}_+^N$ with $\|u\|_1 = 1$), and let $\mathbf{1}$ be the N-dimensional all-ones vector. Then the statement we wish to prove is equivalent to

$$\left\|(A')^{\ell} u - \frac{1}{N}\mathbf{1}\right\|_{\infty} \leq (1 - \alpha)^{\ell}. \tag{7}$$

Write $u = (1/N)\mathbf{1} + \delta$. By the condition that $1 = \|u\|_1 = \sum_i u_i = 1 + \sum_i \delta_i$, it holds that $\langle \delta_i, \mathbf{1} \rangle = 0$. Then

$$\left\| (A')^\ell u - \frac{1}{N}\mathbf{1} \right\|_\infty = \|(A')^\ell \delta\|_\infty \tag{8a}$$
$$\leq \|(A')^\ell \delta\|_2 \tag{8b}$$
$$\leq (1 - \alpha/2)^\ell \|\delta\|_2 \tag{8c}$$
$$\leq (1 - \alpha/2)^\ell \|\delta\|_1 \tag{8d}$$
$$\leq (1 - \alpha/2)^\ell. \tag{8e}$$

Setting this quantity equal to $1/N^c$ and solving for ℓ, we get

$$(1 - \alpha/2)^\ell = N^{-c} \tag{9a}$$
$$\ell \log(1 - \alpha/2) = -c \log N \tag{9b}$$
$$\ell = -2\frac{\log N}{\log(1 - \alpha/2)} \tag{9c}$$
$$\approx \frac{2c \log N}{\alpha}. \tag{9d}$$

□

3.3 Non-deterministic oracle problems

Definition 4 (Quantum oracle problems)**.** *For a n-bit boolean function $\mathcal{O}$, we say an* oracle *decision problem $\mathcal{L}^{\mathcal{O}}$ is in $\mathsf{QMA}^{\mathcal{O}}(\epsilon)$ if there exists a uniform family of quantum circuits $A^{\mathcal{O}}$ such that*

1. *For every* YES *instance $\mathcal{O}$, there exists a quantum state $|\xi\rangle$ of $\mathsf{poly}(n)$ qubits such that $A^{\mathcal{O}}(|\xi\rangle)$ accepts with probability $\geq 1 - \epsilon$.*
2. *For every* NO *instance $\mathcal{O}$, for all quantum states $|\xi\rangle$ of $\mathsf{poly}(n)$ qubits, $A^{\mathcal{O}}(|\xi\rangle)$ accepts with probability $\leq \epsilon$.*

$\mathsf{QCMA}^{\mathcal{O}}(c, s)$ is defined similarly, except the state $|\xi\rangle$ is promised to be classical. The classes $\mathsf{QMA}^{\mathcal{O}}$ and $\mathsf{QCMA}^{\mathcal{O}}$ are defined as $\mathsf{QMA}^{\mathcal{O}}(1/3)$ and $\mathsf{QCMA}^{\mathcal{O}}(1/3)$, respectively.

We note that due to parallel repetition, $\mathsf{QMA}^{\mathcal{O}}(\epsilon = \frac{1}{2} - 1/\mathsf{poly}(n)) = \mathsf{QMA}^{\mathcal{O}} = \mathsf{QMA}^{\mathcal{O}}(\epsilon = 2^{-\mathsf{poly}(n)})$. Likewise, for $\mathsf{QCMA}^{\mathcal{O}}$. This justifies removing the constant ϵ from the definition. We now define the same problem for oracles equaling distributions over n-bit boolean functions.

Definition 5 (Random classical oracles)**.** *A random oracle $\mathcal{R}$ is a distribution over classical oracles $\{\mathcal{O}\}$. We say an* oracle *decision problem $\mathcal{L}^{\mathcal{R}}$ is in $\mathsf{QMA}^{\mathcal{R}}(\epsilon)$ if there exists a uniform family of quantum circuits $A^{\mathcal{O}}$ such that*

1. *For every* YES *instance $\mathcal{R}$, there exists a quantum state $|\xi\rangle$ of $\mathsf{poly}(n)$ qubits such that*
$$\mathop{\mathbf{E}}_{\mathcal{O}\in\mathcal{R}} \mathbf{Pr}\left[A^{\mathcal{O}}(|\xi\rangle) \text{ accepts}\right] \geq 1 - \epsilon. \tag{10}$$
2. *For every* NO *instance $\mathcal{O}$, for all quantum states $|\xi\rangle$ of $\mathsf{poly}(n)$ qubits,*
$$\mathop{\mathbf{E}}_{\mathcal{O}\in\mathcal{R}} \mathbf{Pr}\left[A^{\mathcal{O}}(|\xi\rangle) \text{ accepts}\right] \leq \epsilon. \tag{11}$$

$\mathsf{QCMA}^{\mathcal{R}}(c,s)$ *is defined similarly, except the state* $|\xi\rangle$ *is promised to be classical.*

Ideally, we would define the classes $\mathsf{QMA}^{\mathcal{R}}$ and $\mathsf{QCMA}^{\mathcal{R}}$ are defined as $\mathsf{QMA}^{\mathcal{R}}(1/3)$ and $\mathsf{QCMA}^{\mathcal{R}}(1/3)$, respectively. However, the parallel repetition argument for boolean function oracles cannot be extended to distributions over boolean functions. This is because the ϵ error that an algorithm is the expectation of the success probability of the algorithm over the distribution. It is possible that the algorithm runs on every instance in the distribution with error ϵ or it is possible that the algorithm succeeds with 0 error on a $1-\epsilon$ fraction of the distribution and fails on the remaining ϵ fraction. In the first case, the success of the algorithm can be improved with parallel repetition while it cannot in the second case[5].

3.4 Graph oracles

Definition 6 (Colored Graphs)**.** *Given a d-colored d-regular graph $G=(V,E)$ on N vertices, we say G contains a triple $(j_1,j_2,\kappa)\in V^2\times[d]$ if the edge (j_1,j_2) exists in G and is colored with color κ.*

Definition 7 (Adjacency graph oracles)**.** *Let G be a d-colored d-regular undirected graph. The graph $G=(V,E)$ can be described by an adjacency function*

$$G: V\times[d]\to V \tag{12}$$

where the output of (j,κ) returns the neighbor of j along the edge colored with κ. Quantum access to the function G is provided by the following oracle *unitary:*

$$|j,\kappa,z\rangle \overset{G}{\mapsto} |j,c,z\oplus G(j,\kappa)\rangle\,. \tag{13}$$

We call the function G the adjacency graph oracle *corresponding to G.*

Definition 8 (Expander distinguishing problem)**.** *The (α,ζ)-expander distinguishing problem is a promise oracle language where the input is an oracle G for a d-colored d-regular undirected graph G on N vertices. The problem is to distinguish between the following two cases, promised that one holds:*

- YES*: the graph G has a connected component S of size at most $|S|\leq\zeta$.*
- NO*: the graph G is an α-expander.*

In this paper, we will think of α as a constant and $\zeta\sim N^{9/10}$. To simplify notation, since the oracles considered in this result always correspond to graphs G, we express the algorithm as $\mathcal{A}^G$ rather than $\mathcal{A}^{\mathcal{O}}$.

4 Random distributions over graphs with many connected components

In this subsection, we describe distributions over graphs where the graphs with high probability consist of ℓ connected components. It should not be surprising that the distribution is almost identical to the distribution used by Ambainis, Childs, and Liu [6] in their proof that the expander distinguishing problem requires an exponential number of quantum queries

[5]We note that this subtlety is overlooked in Fefferman and Kimmel [13] but we believe that their result without parallel repetition is correct. Furthermore, the adversary bounds used in [13] do not address this issue but can be rectified using the adversary bound stated in Theorem 13 which deals with *average-case* distinguishing.

for any quantum query algorithm in the absence of a proof. This is because we will reduce any QCMA algorithm to an efficient query algorithm for some expander distinguishing problem.

The lower bound in [6] is crucially a lower bound on the *polynomial degree* of any polynomial that distinguishes two graph distributions. From there, it isn't too much to argue that these graph distributions are very close to YES and NO instances as prescribed in the expander distinguishing problem; therefore any algorithm solving the expander distinguishing problem must be able to distinguish these two graph distributions. Our first goal is to amplify the argument of [6] to a more restricted class of graphs.

[6] Graphs The goal of the construction is a distribution which depends on an integer ℓ and a subset $F \subset V$. The integer ℓ will roughly correspond to the number of connected components (henceforth denoted $C_1, \ldots, C_\ell$) in the graph and we insist that $F \subset C_1$. Every $v \in V \setminus F$ appears in each subset C_i with equal probability of $1/\ell$. The actual construction will be slightly more complicated than this but, morally, this is what we hope to achieve from the distribution.

Formal construction Let N be an integer and for integer $M \geq N$, integer ℓ dividing M and a subset $F \subset V$ define the distribution $P_{M,\ell}(F)$ over graphs on N vertices as follows:

1. Start by constructing a graph G' on M vertices: Partition V' into ℓ equally sized sets of vertices $V_1, \ldots, V_\ell$. On each subset V_k, create a random *colored* subgraph by randomly choosing d perfect matchings (each with a different color $1, \ldots, d$) and taking their union.

2. To construct the graph G on N vertices: We first choose an injective map $\iota : V \hookrightarrow V'$. First, we pick a function $k : V \to [\ell]$. We pick k as a uniformly random function conditioned on the fact that $k(j) = 1$ for each $j \in F$. Let $\iota(v)$ be a random vertex from $V_{k(j)}$ *without replacement* to satisfy injectivity. If all vertices from $V_{k(j)}$ have been selected with replacement, output the graph on N vertices with no edges (i.e. abort).

3. Induce a graph G on V from G' and the map ι — i.e. an edge $(j_1, j_2, \kappa) \in V^2 \times [d]$ exists if $(\iota(j_1), \iota(j_2), \kappa) \in V'^2 \times [d]$ is an edge.

4. For a vertex j and a color κ, if the previous induced edges did not introduce a κ-colored edge from j, then add edge (j, j, κ).

5. The distribution over graphs G is henceforth called $P_{M,\ell}(F)$; when $F = \emptyset$, we write it as $P_{M,\ell}$.

Notationally, for edges $e = (u, v, \kappa) \in G$, we will use $\iota(e) = (\iota(u), \iota(v), \kappa) \in G'$. Furthermore, we will extend ι naturally to subgraphs and subsets of vertices and edges.

Remark 9. *For any F, $P_{M,1}(F) = P_{M,1}(\emptyset) \stackrel{\text{def}}{=} P_{M,1}$.*

4.1 Setting of constants

The lower bounds we prove for the QCMA algorithm are by no means tight (up to constants). We make no attempt to perfect the choice of constants as our only goal is to prove an exponential lower bound on the size of the any quantum witness or the number of queries required to solve the expander distinguishing problem. For this reason, we pick the following constants:

Chosen constants The degree of the graph G' is set to be $d = 100$. We assume $\ell = N^{1/10}$, $\gamma = N^{-1/10}$ and $M = (1+\gamma)N$.

Induced constants In Definition 8, we define an (α, ζ)-expander distinguishing problem. We will only consider $\alpha = 1/(2 \cdot 10^8)$ (which is a consequence of Lemma 10 and the chosen constants). Notationally, we will use $z \stackrel{\text{def}}{=} N/\ell = N^{9/10}$. We will use $\zeta = (1+\gamma)z = M/\ell$.

Conventions Typically, we will assume (for the purposes of contradiction) that $|F| \leq N^{1/100}$ but as that is a term we wish to bound, we explicitly state it each time. Anytime a set S is described, it will be of size ζ, but we will also state this.

4.2 Concentration bounds for random distributions over graphs

We will need the following concentration lemma about the generated distributions. The lemma proves that $P_{M,1}$ is approximately a YES instance and that $P_{M,\ell}(F)$ is approximately a NO instance. Overall, this lemma proves that any algorithm solving the expander distinguishing problem must do very well on identifying the distribution $P_{M,1}$ as a NO instance and identifying the distributions $P_{M,\ell}(F)$ as a YES instance. The proof of this lemma is provided in Appendix A.

Lemma 10 (Adaptation of Lemma 16 of [6])**.** *Assume $|F| \leq N^{1/100}$. Then with probability at least $\geq 1 - O(N^{-3})$, a graph drawn from distribution $P_{M,\ell}(F)$ consists of exactly ℓ connected components each α-expanding and consisting of between $(1-\gamma)z$ and $(1+\gamma)z$ vertices.*

Likewise, the probability that a graph drawn from the distribution $P_{M,1}$ is α-expanding is $\geq 1 - O(N^{-3})$.

Note that being expanding necessarily implies connectivity. Note that when $F = \emptyset$ or $\ell = 1$, there are simpler proofs with tighter bounds but the bound proven here for the general statement is sufficient for our result.

The second concentration lemma that we will use is that $P_{M,\ell}$ is approximately equal to sampling a set F of size $\leq N^{1/100}$ and then sampling a graph from $P_{M,\ell}(F)$. The proof of this lemma is also provided in Appendix A.

Lemma 11. *Let m be $\leq N^{1/100}$. Let $\mathcal{D}_1$ be the distribution on pairs (G, F) obtained by sampling $G \sim P_{M,\ell}$, choosing a uniformly random vertex $v \in G$, and then choosing F to be a uniformly random subset of the connected component of G containing v of size m. Let $\mathcal{D}_2$ be the distribution on pairs (G, F) obtained by first choosing F to be a uniformly random subset of V with size m, and then sampling $G \sim P_{M,\ell}(F)$. Then these distribution are close in statistical distance:*

$$\|\mathcal{D}_1 - \mathcal{D}_2\| \leq 3N^{-9/200}. \tag{14}$$

5 QMA protocol

In this section we show that the expander distinguishing problem (over a fixed graph — i.e., no distribution) can be solved with a polynomial number of queries (indeed, with just two queries) if a quantum witness is provided. Our algorithm has the added benefit of being time-efficient, so we have shown that this problem is contained in QMA^G. In Section 8, we prove that there still exists a QMA protocol if we consider distribution oracles.

Lemma 12. *There is a* QMA^G *protocol* $\mathcal{A}_{\mathsf{QMA}}$ *that solves the* (α, ζ)*-expander distinguishing problem with the following properties:*

1. ***Query complexity:*** *the algorithm makes two queries to* G*.*
2. ***Completeness:*** *In the* YES *case, there exists a witness state that the verifier accepts with certainty.*
3. ***Soundness:*** *In the* NO *case, no witness state is accepted by probability greater than* $1 - \alpha/4$*.*
4. ***Nice witnesses:*** *In the* YES *case, if* $S \subsetneq V$ *is a connected component of the graph* G*, then the state*

$$|S\rangle = \frac{1}{\sqrt{|S|}} \sum_{v \in S} |v\rangle \tag{15}$$

is accepted with probability at least $1 - \sqrt{|S|/N}$*. In particular, since there exists a connected component of size at most* ζ*, there is a state of this form that is accepted with probability* $1 - \sqrt{\zeta/N}$*.*

Proof. First, we note the following fact: given access to the adjacency graph oracle for a d-regular graph, we can implement the unitary U_{walk} defined by

$$U_{\text{walk}} |j, \kappa\rangle |\text{ancilla}\rangle \mapsto |G(j, \kappa), \kappa\rangle |\text{ancilla}\rangle , \tag{16}$$

where $G(j, \kappa)$ is the κ-th neighbor of j (which is guaranteed to exist by the d-regularity condition). To see this, prepare an ancilla in the $|0\rangle$ state, and apply

$$|j, \kappa, 0\rangle \overset{G}{\mapsto} |j, \kappa, G(j, \kappa)\rangle \overset{\text{swap registers}}{\mapsto} |G(j, \kappa), \kappa, j\rangle \overset{G}{\mapsto} |G(j, \kappa), \kappa, 0\rangle \tag{17}$$

Here, we have used the fact that for a validly colored graph, if $G(j_1, \kappa) = j_2$, then $G(j_2, \kappa) = j_1$. Moreover, using controlled queries to G, we can also implement the controlled version of U_{walk}.

Let us also define the following states:

$$|\overline{0}_V\rangle := \frac{1}{\sqrt{N}} \sum_{v \in V} |j\rangle , \qquad |\overline{0}_d\rangle := \frac{1}{\sqrt{d}} \sum_{\kappa=1}^{d} |\kappa\rangle . \tag{18}$$

Let $|\psi\rangle$ be the witness state received from the prover. The verifier performs the following operation:

1. First, the verifier prepares an ancillas in the state $|+\rangle$ and $\frac{1}{\sqrt{d}} \sum_{\kappa=1}^{d} |\kappa\rangle$. The total state at this point is

$$|+\rangle_{\text{control}} \otimes |\psi\rangle_{\text{witness}} \otimes \left(\frac{1}{\sqrt{d}} \sum_{\kappa=1}^{d} |\kappa\rangle \right)_{\text{color}} . \tag{19}$$

   Now, the verifier applies U_{walk} to the witness and color registers controlled on the control register.

2. Next, the verifier measures the control register in the $\{|+\rangle , |-\rangle\}$ basis, and the color register using the two-outcome measurement $\{M_0^C, M_1^C\}$ where $M_0^C = |\overline{0}_d\rangle$. The verifier proceeds to the next step if and only if the control measurement yields $+$ and the color measurement yields 0. Otherwise, it rejects.

3. Finally, it performs the two-outcome measurement $\{M_0^V, M_1^V\}$ on the witness register where $M_0^V = |\overline{0}_V\rangle\langle\overline{0}_V|$. If the outcome 0 is obtained, it rejects. Otherwise, it accepts.

The associated quantum circuit for this verifier is given in Figure 3.

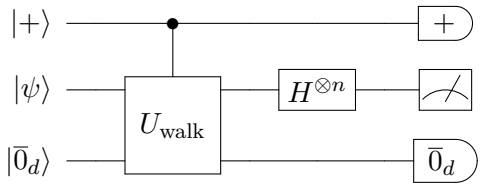

Figure 3: The QMA^G verifier.

Query complexity From the description of the algorithm it is clear that only one query to U_{walk}, and thus two queries to G are made.

Analysis To analyze the verification algorithm, let us write the witness as

$$|\psi\rangle = \sum_j \alpha_j |j\rangle . \tag{20}$$

Then the state after the controlled gate is

$$\frac{1}{\sqrt{2}} |0\rangle |\psi\rangle |\overline{0}_d\rangle + \frac{1}{\sqrt{2}} |1\rangle \sum_j \sum_{\kappa\in[d]} \frac{\alpha_j}{\sqrt{\kappa}} |G(j,\kappa)\rangle |i\rangle . \tag{21}$$

If we apply the projector of the third register onto $|\overline{0}_d\rangle$, we get the resulting un-normalized state is

$$\frac{1}{\sqrt{2}} |0\rangle |\psi\rangle + \frac{1}{\sqrt{2}} |1\rangle \sum_j \sum_{\kappa\in[d]} \frac{\alpha_j}{d} |G(j,\kappa)\rangle = \frac{1}{\sqrt{2}} |0\rangle |\psi\rangle + \frac{1}{\sqrt{2}} |1\rangle A |\psi\rangle \tag{22}$$

where A is the normalized adjacency matrix of the graph. The probability of measuring the control register as $|+\rangle$ on this un-normalized state is equal to to the probability that the control register and the color both yield accepting outcomes, and can be calculated to be

$$\begin{aligned}
\mathbf{Pr}[+,\overline{0}_d] &= \left| (\langle +| \otimes I) \left(\frac{1}{\sqrt{2}} |0\rangle \otimes |\psi\rangle + \frac{1}{\sqrt{2}} |1\rangle \otimes A |\psi\rangle \right) \right|^2 && \text{(23a)} \\
&= \left| \frac{1}{2} |\psi\rangle + \frac{1}{2} A |\psi\rangle \right|^2 && \text{(23b)} \\
&= \frac{1}{4} \langle\psi|\psi\rangle + \frac{1}{4} \langle\psi| A^\dagger A |\psi\rangle + \frac{1}{2} \operatorname{Re} \langle\psi| A |\psi\rangle && \text{(23c)} \\
&= \frac{1}{4} + \frac{1}{4} \langle\psi| A^2 |\psi\rangle + \frac{1}{2} \langle\psi| A |\psi\rangle && \text{(23d)} \\
&\leq \frac{1}{2} + \frac{1}{2} \langle\psi| A |\psi\rangle . && \text{(23e)}
\end{aligned}$$

where in the penultimate line we used that A is a real symmetric matrix, and in the last line we used that A has operator norm at most 1.

Completeness In the YES case, let S be a connected component and T be its complement. Then the prover will send the state

$$|\xi_S\rangle \stackrel{\text{def}}{=} \sqrt{\frac{|T|}{N}}\,|S\rangle - \sqrt{\frac{|S|}{N}}\,|T\rangle\,. \tag{24}$$

It is easy to check that $|\xi_S\rangle$ is a $+1$ eigenvector of A, and that it is orthogonal to $|\overline{0}_V\rangle$. Therefore, by eq. (23d), it is accepted by the verifier with probability 1.

Soundness Assume that for some NO case graph, the witness $|\psi\rangle$ is accepted with probability $1-\gamma$ for $\gamma < \alpha/4$. Note the second eigenvalue λ satisfies $|\lambda| \leq 1$, so $\delta = 1-\lambda \leq 2$, and hence $1-\gamma > 1/2$. Since step 2 of the verifier must accept with at least this probability, by eq. (23e),

$$1 - 2\gamma \leq \langle\psi|A|\psi\rangle\,. \tag{25}$$

Since the NO case graph has a single connected component and a unique eigenvector of eigenvalue 1, M_0^V is the projector onto the 1-eigenspace of A. Furthermore, since the graph is α-expanding, every other eigenvector has eigenvalue at most $1-\alpha$. Since A commutes with M_0^V,

$$\begin{aligned}
1 - 2\gamma &\leq \langle\psi|A|\psi\rangle && \text{(26a)}\\
&= \langle\psi|M_0^V A M_0^V|\psi\rangle + \langle\psi|(\mathbb{I}-M_0^V)A(\mathbb{I}-M_0^V)|\psi\rangle && \text{(26b)}\\
&\leq \left\|M_0^V\,|\psi\rangle\right\|^2 + (1-\alpha)\left(1 - \left\|M_0^V\,|\psi\rangle\right\|^2\right) && \text{(26c)}\\
&= \alpha\left\|M_0^V\,|\psi\rangle\right\|^2 + (1-\alpha) && \text{(26d)}
\end{aligned}$$

Solving this equation tells us that

$$\left\|M_0^V\,|\psi\rangle\right\|^2 \geq 1 - \frac{2\gamma}{\alpha}. \tag{27}$$

Therefore, the measurement of the witness register in step 3 of the verifier will reject with probability $1 - 2\gamma/\alpha > 1/2$, causing the algorithm to reject with probability greater than $1/2$, a contradiction. Therefore, $\gamma \geq \alpha/4$.

Nice witnesses Grilo, Kerenedis and Sikora noted in [15] that the witness for a QMA problem can always be assumed to be a subset state. We show that the same property holds for our oracular witnesses. The overlap between the ideal witness and $|S\rangle$ is

$$\langle\xi_S|S\rangle = \sqrt{\frac{|T|}{N}}. \tag{28a}$$

Therefore, the trace distance between these states satisfies

$$\|\,|\xi_S\rangle\langle\xi_S| - |S\rangle\langle S|\,\|_1 \leq \sqrt{1 - \frac{|T|}{N}} = \sqrt{\frac{|S|}{N}}. \tag{29}$$

Observe that the entire operation of the verifier given G can be modeled by a single two-outcome measurement $\{M_0^G, M_1^G\}$, where 0 corresponds to rejection and 1 to acceptance.

Therefore, by a standard trace distance fact [23, Equation 9.22], if G is a YES instance, the probability that $|S\rangle$ is accepted is at least

$$\mathbf{Pr}[|S\rangle \text{ accepted}] \geq 1 - \sqrt{\frac{|S|}{N}}. \tag{30}$$

□

6 Adversary method

In this section, we use the adversary method of Ambainis to argue that any successful QCMA algorithm implies a BQP algorithm for distinguishing the distribution $P_{M,\ell}(F)$ and $P_{m,1}$ for $|F| \leq N^{1/100}$. In Section 6.1 we state the adversary method result and in the following sections we prove the statement.

6.1 Ambainis' proof of the adversary method

The adversary method of Ambainis [5] is a convenient way of arguing lower bounds on the query complexity of oracular quantum algorithms. The adversary method lower bounds the complexity of any algorithm which (with high probability) computes $f(a)$ for a function $f : \{0,1\}^N \to \{0,1\}$. The quantum algorithm is allowed access to $a \in \{0,1\}^N$ by a oracle gate O which applies linearly the transform $|i\rangle \mapsto (-1)^{a_i} |i\rangle$ for $i \in \{0,1\}^n$ (here $N = 2^n$). In doing so, the adversary method is a convenient way of producing BQP (query) lower-bounds.

To use it in our distributional setting, we make two modifications to the adversary bound. The first is to relax the notion of correctness. The lower bound of Ambainis is for a lower-bound for any algorithm which, for *each* $a \in \{0,1\}^N$, outputs $f(a)$ correctly with probability $1-\epsilon$ for $\epsilon < \frac{1}{2}$. We instead consider an *average-case* notion of success in which

$$\mathop{\mathbf{E}}_{a\in\{0,1\}^N} \mathop{\mathbf{Pr}}_{\mathcal{A}} [\mathcal{A}^a \neq f(a)] \leq \epsilon^2. \tag{31}$$

By Markov's inequality, this implies

$$\mathop{\mathbf{Pr}}_{a\in\{0,1\}^N} \left[\mathop{\mathbf{Pr}}_{\mathcal{A}} [\mathcal{A}^a \neq f(a)] \geq \epsilon \right] \leq \epsilon, \tag{32}$$

or in other words, most a are (with high probability) correctly identified.

The second modification is to *restrict* the set of locations that the algorithm is allowed to query the oracle. The reason for this is somewhat subtle. Essentially, the original lower bound of Ambainis was designed for decision problems with deterministic oracles, and relies on constructing a relation between two *disjoint* sets of oracle instances, one consisting only of YES instances and the other only of NO instances. However, in our setting, we are interested in distinguishing two distributions over oracles that may have overlapping support. In order to define disjoint YES and NO sets of instances even when the distributions overlap, we add to each oracle string a a set of flag bits b that indicate which of the two distributions the string a was sampled from. Naturally, any reasonable model cannot permit the algorithm to query the flag bits: otherwise, it would be easy to distinguish even two statistically close distributions with few queries.

More formally, we consider a generalization where the oracle string is a tuple $(a, b) \in \{0,1\}^N \times \{0,1\}^M$ and $f : \{0,1\}^{N+M} \to \{0,1\}$ but the algorithm can only query positions of a. In this model, with the average-case notion of success defined above, we obtain the following adversary lower bound for distributions:

Theorem 13. *Let $f : \{0,1\}^{N+M} \to \{0,1\}$ be a function and let $X, Y \subset \{0,1\}^N \times \{0,1\}^M$ be two subsets such that $X \subset f^{-1}(0)$ and $Y \subset f^{-1}(1)$. Let $R \subset X \times Y$ be a relation such that*

1. *For every $x \in X$, let $R_x \subset Y$ equal $R_x = \{y : (x,y) \in R\}$ such that $\underline{m} \leq |R_x| \leq \overline{m}$.*
2. *For every $y \in Y$, let $R_y \subset X$ equal $R_y = \{x : (x,y) \in R\}$ such that $\underline{m'} \leq |R_x| \leq \overline{m'}$.*
3. *For every $x = (a,b) \in X$ and $i \in [N]$, let $\ell_{x,i}$ be the number of $y = (c,d) \in Y$ such that $(x,y) \in R$ and $a_i \neq c_i$. Likewise, for every $y = (c,d) \in Y$ and $i \in [N]$, let $\ell_{y,i}$ be the number of $x = (a,b) \in X$ such that $(x,y) \in R$ and $a_i \neq c_i$. Let $\ell_{\max}$ be the maximum product $\ell_{x,i}\ell_{y,i}$ over $(x,y) \in R$ and $i \in [N]$ such that $a_i \neq c_i$.*

Then any quantum algorithm $\mathcal{A}$ which only queries the first N bits of the oracle and computes f such that

$$\mathop{\mathbf{E}}_{x=(a,b)\in X} \Pr_{\mathcal{A}}[\mathcal{A}^a \neq 0] \leq \epsilon^2 \quad \text{and} \quad \mathop{\mathbf{E}}_{y=(c,d)\in Y} \Pr_{\mathcal{A}}[\mathcal{A}^c \neq 1] \leq \epsilon^2 \tag{33}$$

uses

$$\geq \left(1 - 2\sqrt{\epsilon(1-\epsilon)}\right)\sqrt{\frac{(\underline{m} - 2\epsilon\overline{m})\left(\underline{m'} - 2\epsilon\overline{m'}\right)}{\ell_{\max}}} \quad \text{queries.} \tag{34}$$

Corollary 14. *Let X and Y be two subsets of $\{0,1\}^{N+M}$ satisfying the three conditions listed in Theorem 13. Then, any query algorithm $(1-\delta)$-distinguishing the uniform distributions over X and Y, must use eq. (34) queries for $\epsilon = 2\delta$.*

The proofs of both statements are presented in Appendix B.

6.2 Setup from QCMA algorithm

In this subsection, we show that if there is a QCMA algorithm for solving the *expander distinguishing problem* then there exists a sunflower ❀ (defined below) of YES instances which correspond to the same optimal witness $\mathsf{wt}^\star$. If we hardcode $\mathsf{wt}^\star$ into the QCMA algorithm, we generate a quantum query algorithm that, with no access to a prover, accepts instances corresponding to ❀ and rejects all NO instances.

Definition 15 (Sunflower)**.** *A collection of subsets ❀ $\subset \binom{V}{\zeta}$ is* (μ, ζ, t)-sunflower *if there exists a subset $F \subset V$ with $|F| \leq t$ satisfying the following two conditions:*

1. *For all $S \in$ ❀, $F \subseteq S$.*
2. *For all $x \in \left(\bigcup_{S\in❀} S\right) \setminus F$, the $\Pr_{S\in❀}[x \in S] \leq \left(\frac{\zeta}{N}\right)^{1-\mu}$.*

We call the set F the core of the sunflower.

YES instances corresponding to subsets For any graph G and subset S of size ζ, define $G \triangleleft S$ if G has a connected component $C_i \subseteq S$. Let $\mathsf{S}_\triangleleft$ be the set of G such that $G \triangleleft S$. For each subset S of size ζ, define B_S to be the restriction of the distribution $P_{M,\ell}$ to graphs in $\mathsf{S}_\triangleleft$. The intuition is that the witness $|\xi_S\rangle$ from eq. (24) will be a good witness for B_S since the connected components of $P_{M,\ell}$ are of a size concentrated around z.

There is a small complexity, which we address now, in that the distribution $P_{M,\ell}$ is not a uniform distribution over a set of graphs. To rectify this, we can always assume that the oracle corresponding to a graph G sampled to $P_{M,\ell}$ consists of a queryable component corresponding to the adjacency list of G and a non-queryable component corresponding to the random coins r_G that were flipped in order to generate G according to $P_{M,\ell}$. We will also define B_S as the restriction of the extended oracle. Therefore, both $P_{M,\ell}$ and B_S are uniform distributions over some support.

Lastly, the distributions B_S are not exactly YES distributions since their support is not *entirely* on YES graphs of the expander distinguishing problem. However, similar to Lemma 10, we will show that B_S is almost entirely supported on YES graphs. Therefore, it suffices to use B_S as a *proxy* for YES instances until the very end where we handle this subtlety.

Corollary 16. *For every graph $G \in \mathsf{S}_\triangleleft$ such that G has a connected component C_i with $|C_i| \geq (1-\gamma)z$,*

$$\mathbf{Pr}[\mathcal{A}^G_{\mathsf{QMA}}(|S\rangle) = 1] \geq 1 - 3\sqrt{\gamma}. \tag{35}$$

Proof. Since, $|C_i|/|S| = |C_i|/\zeta \geq (1-\gamma)/(1+\gamma) \geq 1 - 2\gamma$, then

$$\||S\rangle - |C_i\rangle\| \leq \sqrt{2\left(1 - \sqrt{\frac{|C_i|}{|S|}}\right)} \leq \sqrt{2(1-\sqrt{1-2\gamma})} \leq 2\sqrt{\gamma}. \tag{36}$$

By Lemma 12, $|C_i\rangle$ is accepted with probability $1 - \sqrt{(1-\gamma)z/N}$, then $|S\rangle$ is accepted with probability $1 - 3\sqrt{\gamma}$. □

Corollary 17. *Let $M = (1+\gamma)N$ and $\gamma = N^{-1/10}$ and $\ell = N^{-1/10}$. For every S of size $(1+\gamma)z$, the distribution B_S is a YES instance and*

$$\mathop{\mathbf{E}}_{G \leftarrow B_S}\left[\mathbf{Pr}\left[\mathcal{A}^G_{\mathsf{QMA}}(|S\rangle) = 1\right]\right] \geq 1 - 3\sqrt{\gamma} - O(N^{-3}). \tag{37}$$

Proof. By Lemma 10, with all but $O(N^{-3})$ probability, a graph G drawn from the distribution $P_{M,\ell}$ consists of ℓ connected components each with size $\in [(1-\gamma)z, (1+\gamma)z]$. Then every connected component of the graph G is contained in some connected component S of size $(1+\gamma)z$. By the symmetry of the distribution $P_{M,\ell}$ under permutations, it follows that the distribution $\widetilde{P}_{M,\ell}$ formed by sampling a subset S of size $(1+\gamma)z$ and then sampling a graph from the distribution B_S, is $O(N^{-3})$ close to the distribution $P_{M,\ell}$. Therefore, with probability all but $O(N^{-3})$, for any S, the distribution B_S will consist of ℓ connected components each with size $\in [(1-\gamma)z, (1+\gamma)z]$. By the previous corollary, the algorithm with witness $|S\rangle$ accepts with probability $\geq 1 - 3\sqrt{\gamma}$. A union bound completes the proof. □

QCMA algorithm implies a quantum low-query algorithm for some sunflower ❀

Lemma 18. *For some $\epsilon > 0$, assume there exists a k-query non-deterministic quantum algorithm which accepts a q-length* classical *witness and accepts every distribution B_S for subset S of size ζ with probability $1-\epsilon$ and accepts any* NO *distribution B_{NO} with probability at most ϵ. Then for $\mu > 0$, there exists a $(\mu, \zeta, 2q/(\mu \log \ell))$-sunflower ❀ and a k-query quantum algorithm that accepts every distribution B_S for $S \in ❀$ and accepts any* NO *distribution B_{NO} with probability at most ϵ.*

Proof. Assume such a non-deterministic algorithm $\mathcal{A}_{\mathsf{QCMA}}$ exists. Let the optimal witness (for algorithm $\mathcal{A}_{\mathsf{QCMA}}$) for oracle B_S be $\mathsf{wt}(B_S)$; since the oracles are in bijection with subsets $S \in \binom{V}{\zeta}$, we can think of wt as function

$$\mathsf{wt} : \binom{V}{\zeta} \to \{0,1\}^q. \tag{38}$$

Formally this means that for every S, there exists a $k(n)$-query quantum algorithm $\mathcal{A}_{\mathsf{QCMA}}$ and a witness $\mathsf{wt}(S)$ such that

$$\mathop{\mathbf{E}}_{G \leftarrow B_S} \left[\mathbf{Pr} \left[\mathcal{A}^G_{\mathsf{QCMA}}(\mathsf{wt}(S), 1^n) = 1 \right] \right] \geq 1 - \epsilon. \tag{39}$$

Let $\mathsf{wt}^\star \in \{0,1\}^q$ be the most popular witness (one associated with the largest number of subsets S) and let $\Sigma \stackrel{\mathsf{def}}{=} \mathsf{wt}^{-1}(\mathsf{wt}^\star)$. The size $|\Sigma|$ is by a counting argument at least $2^{-q}\binom{N}{\zeta}$. Notice that since $|S| = \zeta$, then for a uniformly random $S \in \binom{V}{\zeta}$, $\mathbf{Pr}_S[j \in S] = \zeta/N$ for all $j \in V$. Ideally, if S were instead sampled from Σ, we would like that $\mathbf{Pr}_{S \in \Sigma}[j \in S] \sim \zeta/N$ for all $j \in V$. Of course, this is too good to be true as Σ could be the set of all S such that 0^n is contained in S (for example). Instead, we build a sunflower ❀ $\subset \Sigma$ with the following greedy strategy inspired by [13]. For ❀ $= \Sigma$ (initially), whenever there exists a j such that

$$\mathop{\mathbf{Pr}}_{S \in ❀}[j \in S] \geq \left(\frac{\zeta}{N}\right)^{1-\mu}, \tag{40}$$

then we restrict ❀ $\leftarrow$ ❀ $\cap \{S : j \in S\}$. By construction, after each restriction, the size of ❀ is at least its size before restriction multiplied by $(\zeta/N)^{1-\mu}$. If we continue this process and each time add j to a set F, then

$$|❀| \geq \left(\frac{\zeta}{N}\right)^{(1-\mu)|F|} |\Sigma| \geq \left(\frac{\zeta}{N}\right)^{(1-\mu)|F|} \cdot 2^{-q(n)} \cdot \frac{N!}{\zeta!(N-\zeta)!}. \tag{41}$$

On the other hand, after selecting the set F, each element of $S \in ❀$ necessarily contains $F \subset S$ and therefore

$$(S \setminus F) \subseteq \binom{V \setminus F}{\zeta - |F|}. \tag{42}$$

Therefore,

$$|❀| \leq \frac{(N-|F|)!}{(\zeta-|F|)!(N-\zeta)!}. \tag{43}$$

Combining eq. (41) and eq. (43), we get that

$$\left(\frac{\zeta}{N}\right)^{(1-\mu)|F|} \cdot 2^{-q(n)} \leq \frac{\zeta!}{N!} \cdot \frac{(N-|F|)!}{(\zeta-|F|)!} = \frac{\zeta}{N} \cdot \frac{\zeta-1}{N-1} \cdots \cdots \frac{\zeta-|F|+1}{N-|F|+1} \leq \left(\frac{\zeta}{N}\right)^{|F|}. \tag{44}$$

Equivalently,

$$|F| \leq \frac{q(n)}{\mu \log(N/\zeta)} \leq \frac{2q(n)}{\mu \log \ell}. \tag{45}$$

The end result is that ❀ is $(\mu, \zeta, \frac{2q(n)}{\mu \log \ell})$-sunflower. Let $\mathcal{A}$ be the algorithm $\mathcal{A}_{\mathsf{QCMA}}$ with the witness $\mathsf{wt}^\star$ hard-coded into the algorithm. Since ❀ $\subset \Sigma$, for every $S \in$ ❀,

$$\mathop{\mathbf{E}}_{G \leftarrow B_S} \left[\mathbf{Pr} \left[\mathcal{A}^G(1^n) = 1 \right] \right] \geq 1 - \epsilon. \tag{46}$$

And since the algorithm $\mathcal{A}_{\mathsf{QCMA}}$ accepts every NO distribution B_{NO} with probability at most ϵ irrespective of proof, then $\mathcal{A}$ also does them same. □

6.3 Query lower bound for distinguishing sunflowers and fixed distributions

Let Ⓕ $\stackrel{\mathsf{def}}{=} \binom{V}{\zeta} \cap \{S : F \subseteq S\}$. This is the *ideal sunflower* with a core of F. We will show by an adversary bound that the sunflower ❀ and the ideal sunflower Ⓕ are indistinguishable by quantum query algorithms with few queries. Consider the distribution $H_{❀}$ defined by sampling an $S \in$ ❀ and then sampling a graph from B_S. Similarly, define the distribution $H_{Ⓕ}$ but by first sampling an $S \in$ Ⓕ. We want to show that any quantum query algorithm requires exponentially many queries to distinguish $H_{❀}$ and $H_{Ⓕ}$. The main result of this subsection is the following lemma.

Lemma 19. *For $\delta < 1/4$, any quantum query algorithm $(1-\delta)$-distinguishing the distributions $H_{❀}$ and $H_{Ⓕ}$ where ❀ is a (μ, ζ, t)-sunflower and F is the corresponding core requires*

$$\geq \frac{1}{2}\left(1 - 2\sqrt{2\delta(1-2\delta)}\right)(1-4\delta) \cdot \sqrt{\left(\frac{N}{\zeta}\right)^{1-\mu}} \text{ queries.} \tag{47}$$

6.3.1 A warmup lemma for distinguishing graphs

The main challenge in proving Lemma 19 is the complicated structure inherent in graphs. However, if we work instead directly with the *sets* S, the problem is much simpler, and was already solved in [13, Lemma 11]. They showed that given membership query access (equivalently, the indicator function for the set), it requires exponentially many quantum queries to distinguish a sample from ❀ from a sample from Ⓕ.

We will work up to the result we wish to prove by gradually adding more structure to the objects being queried until we reach graphs. We will start by working with *permutations* that map the set S to a known set, and show that any algorithm with query access to the permutation and its inverse requires exponentially many queries.

To be precise, let $U = [\zeta]$. Let $\Pi_{❀}$ be the set of all permutations and inverses (π, π^{-1}) such that $\pi(S) = U$ for some $S \in$ ❀. Similarly, define $\Pi_{Ⓕ}$. We shall abuse notation and also use $\Pi_{❀}$ and $\Pi_{Ⓕ}$ to refer to the uniform distributions over these sets of permutations. We first claim that no quantum query algorithm can distinguish the distributions $\Pi_{❀}$ and $\Pi_{Ⓕ}$ without an exponential number of queries. Note that the algorithm is allowed to query both the permutation and its inverse[6].

[6]This can be equivalently modeled by having a separate in-place oracle for the permutation and its inverse, or having a single "standard" oracle for the permutation.

Lemma 20. *Any quantum query algorithm $(1-\delta)$-distinguishing the distributions $\Pi_{❀}$ and $\Pi_{Ⓕ}$ where ❀ is a (μ,ζ,t)-sunflower and F is the corresponding core requires*

$$\geq \left(\frac{1}{2} - 2\sqrt{2\delta(1-2\delta)}\right)(1-4\delta)\cdot\sqrt{\left(\frac{N}{\zeta}\right)^{1-\mu}} \text{ queries.} \tag{48}$$

Proof. We will use the adapted adversary bound (Theorem 13 and Corollary 14) proved in Appendix B. To do so we need to construct a relation $R \subset \Pi_{❀} \times \Pi_{Ⓕ}$. To build this relation, for every pair $(S_x, S_y) \in ❀ \times Ⓕ$, pick permutations χ_{xy} and ψ_{xy} such that

1. $\chi_{xy}(S_x) = \psi_{xy}(S_y) = U$.
2. For all $j \in (S_x \cap S_y) \cup (V \setminus (S_x \cup S_y))$, $\chi_{xy}(j) = \psi_{xy}(j)$.
3. For all $j_1 \in S_x \setminus S_y$, there exists a $j_2 \in S_y \setminus S_x$ such that $\chi_{xy}(j_1) = \psi_{xy}(j_2)$ and $\chi_{xy}(j_2) = \psi_{xy}(j_1)$.

Such permutations are easy to find by picking a permutation χ_{xy} and then choosing the unique ψ_{xy} that satisfies the constraints. Notice that every permutation mapping S_x to U can be expressed as $\tau \circ \chi_{xy}$ such that $\tau(U) = U$. Likewise, every permutation mapping S_y to U can be expressed as $\tau \circ \psi_{xy}$ for $\tau(U) = U$. Construct the relation R by adding all pairs defined by the same τ:

$$R \stackrel{\text{def}}{=} \left\{ \left(\left(\tau \circ \chi_{xy}, \chi_{xy}^{-1} \circ \tau^{-1}\right), \left(\tau \circ \psi_{xy}, \psi_{xy}^{-1} \circ \tau^{-1}\right) \right) \ : \ (S_x, S_y) \in ❀ \times Ⓕ, \tau(U) = U \right\}. \tag{49}$$

Consider an element $(\pi, \pi^{-1}) \in \Pi_{❀}$ such that $\pi(S_x) = U$. For any $S_y \in Ⓕ$, write $\pi = \tau \circ \chi_{xy}$. Then $(\tau \circ \psi_{xy}, \psi_{xy}^{-1} \circ \tau^{-1}) \in \Pi_{Ⓕ}$ forms a neighbor of (π, π^{-1}) along R. Therefore, the degree m of every element of $\Pi_{❀}$ is $|Ⓕ|$ and, analogously, the degree m' of every element of $\Pi_{Ⓕ}$ is $|❀|$.

Consider now any $\left((\sigma_x, \sigma_x^{-1}), (\sigma_y, \sigma_y^{-1})\right) \in R$ such that $\sigma_x(S_x) = \sigma_y(S_y) = U$. To apply Theorem 13, we need to calculate $\ell_{x,j}$ and $\ell_{y,j}$ where $\ell_{x,j}$ is the number of y' such that $\left((\sigma_x, \sigma_x^{-1}), (\sigma_{y'}, \sigma_{y'}^{-1})\right) \in R$ and $(\sigma_x, \sigma_x^{-1})$ and $(\sigma_y, \sigma_y^{-1})$ differ at index j. $\ell_{y,j}$ is defined analogously (see Theorem 13). There are two cases to consider: (A) either an index j such that $\sigma_x(j) \neq \sigma_y(j)$ or (B) an index j such that $\sigma_x^{-1}(j) \neq \sigma_y^{-1}(j)$.

- Case (A): By construction, in order for $\sigma_x(j) \neq \sigma_y(j)$, either (A1) $j \in S_x \setminus S_y$ or (A2) $j \in S_y \setminus S_x$.
  - (A1): A simple upper bound for $\ell_{x,j}$ is $|Ⓕ|$. To bound $\ell_{y,j}$, notice that since $j \notin S_y$, then in order for $\sigma_{x'}(j) \neq \sigma_y(j)$, $j \in S_{x'} \setminus S_y$. The number of such x' is equal to the number of sets $S_{x'}$ in ❀ that contain the point $\sigma_y(j)$. Since $F \subset S_y$, it follows that the point $\sigma_y(j)$ is *not* contained in F. Therefore, since ❀ is a sunflower, it holds that the number of $S_{x'} \in ❀$ satisfying the condition is at most
  $$|❀| \cdot \mathop{\mathbf{Pr}}_{S \in ❀}[\sigma_y(j) \in S] \leq |❀| \cdot \left(\frac{\zeta}{N}\right)^{1-\mu}. \tag{50}$$
  - (A2): A simple upper bound for $\ell_{y,j}$ is $|❀|$. To bound $\ell_{x,j}$, notice that since $j \notin S_x$, then in order for $\sigma_{y'}(j) \neq \sigma_x(j)$, $j \in S_{y'} \setminus S_x$. Since $F \subset S_x$, then the number of such y' is at most $\frac{\zeta - |F|}{N - |F|} \cdot |Ⓕ| \leq 2\left(\frac{\zeta}{N}\right)|Ⓕ|$.

- Case (B): By construction, in order for $j' \stackrel{\text{def}}{=} \sigma_x^{-1}(j) \neq \sigma_y^{-1}(j)$, either (B1) $j' \in S_x \setminus S_y$ or $j' \in S_y \setminus S_x$.
  - (B1): Follows (A1) but for j'.
  - (B2): Follows (A2) but for j'.

Since this exhausts all cases,

$$\ell_{\max} \stackrel{\text{def}}{=} \max_{\substack{((\sigma_x,\sigma_x^{-1}),(\sigma_y,\sigma_y^{-1}))\in R \\ \text{differing at } j}} \ell_{x,j}\ell_{y,j} \leq \left(\frac{\zeta}{N}\right)^{1-\mu} |\maltese||\textcircled{F}|. \tag{51}$$

Then by direct application of Corollary 14, the algorithm must use

$$\geq \left(1 - 2\sqrt{2\delta(1-2\delta)}\right)(1-4\delta) \cdot \sqrt{\left(\frac{N}{\zeta}\right)^{1-\mu}} \text{ queries.} \tag{52}$$

□

A short corollary of Lemma 20 is that there is a similar query lower bound for distinguishing distributions over graphs. Let $\mathcal{G}$ be a distribution over graphs with a connected component of $U = [\zeta]$. Let $\mathcal{G}_{\maltese}$ be the distribution over graphs formed by sampling a permutation pair (π, π^{-1}) from $\Pi_{\maltese}$, a graph G from $\mathcal{G}$ and outputting the graph $\pi^{-1}(G)$. By construction, $\mathcal{G}_{\maltese}$ is a distribution over graphs with a connected component of S for $S \in \maltese$. Likewise, define the distribution $\mathcal{G}_{\textcircled{F}}$.

Corollary 21. *For $\delta < 1/4$, any quantum query algorithm $(1-\delta)$-distinguishing the distributions $\mathcal{G}_{\maltese}$ and $\mathcal{G}_{\textcircled{F}}$ where $\maltese$ is a (μ, ζ, t)-sunflower and F is the corresponding core requires*

$$\geq \frac{1}{2}\left(1 - 2\sqrt{2\delta(1-2\delta)}\right)(1-4\delta) \cdot \sqrt{\left(\frac{N}{\zeta}\right)^{1-\mu}} \text{ queries.} \tag{53}$$

Proof. Any algorithm $\mathcal{A}$ for distinguishing $\mathcal{G}_{\maltese}$ and $\mathcal{G}_{\textcircled{F}}$, can be used as a subroutine in a (not necessarily time-efficient) algorithm $\mathcal{A}'$ for distinguishing $\Pi_{\maltese}$ and $\Pi_{\textcircled{F}}$ in twice as many queries.

To motivate the algorithm, observe that if G is a random graph drawn from $\mathcal{G}$, then the graph $\pi^{-1}(G)$ is a random graph drawn from $\mathcal{G}_{\maltese}$ or $\mathcal{G}_{\textcircled{F}}$, depending on whether $\pi \in \Pi_{\maltese}$ or $\Pi_{\textcircled{F}}$. Moreover, a graph oracle query to the graph $\pi^{-1}(G)$ can be performed using two oracle queries to π, π^{-1}.

Now we can specify the algorithm $\mathcal{A}'$: first, it samples a graph G from $\mathcal{G}$—this step is not time-efficient, but it makes no oracle queries. Next, it runs $\mathcal{A}$ on the graph $\pi^{-1}(G)$, simulating oracle queries to this graph as described above. It answers according to the outcome of $\mathcal{A}$. If $\mathcal{A}$ successfully distinguishes $\mathcal{G}_{\maltese}$ and $\mathcal{G}_{\textcircled{F}}$, then $\mathcal{A}'$ distinguishes $\Pi_{\maltese}$ and $\Pi_{\textcircled{F}}$ with the same probability, and using twice as many oracle queries, as claimed. Thus, by Lemma 20, we obtain the claimed query bound for $\mathcal{A}$. □

6.3.2 Improving to more general permutations

While Lemma 20 and Corollary 21 are simple enough to prove, they are insufficient at proving indistinguishability for the graph distributions $H_{\maltese}$ and $H_{\textcircled{F}}$ defined at the start of this section. This is because, unlike the distribution $\mathcal{G}_{\maltese}$, the distribution $H_{\maltese}$ cannot be

defined in terms of independently sampling a graph G and a set S. For one, the sizes of the connected components in $H_{❀}$ do not exactly equal z; instead, the concentrate tightly around z. It was precisely the independence of the graphs and sets that made Corollary 21 easy to prove.

To fix the argument, we prove the following variations of Lemma 20 and Corollary 21. For a sunflower ❀ with core F and any k such that $|F| \leq k \leq \zeta$, let $\Pi_{❀}^{(k)}$ be the distribution formed by the following procedure:

1. Sample a set S from ❀.
2. Sample uniformly randomly a subset $C \subset S$ of size k.
3. Sample uniformly randomly a permutation $\pi : V \to V$ such that $\pi(C) = [k]$.
4. Output (π, π^{-1}).

Define the distribution $\Pi_{Ⓕ}^{(k)}$ similarly where we change the first step to sampling from Ⓕ.

Lemma 22. *Any quantum query algorithm $(1-\delta)$-distinguishing the distributions $\Pi_{❀}^{(k)}$ and $\Pi_{Ⓕ}^{(k)}$ where ❀ is a (μ, ζ, t)-sunflower and F is the corresponding core requires*

$$\geq \left(1 - 2\sqrt{2\delta(1-2\delta)}\right)(1-4\delta) \cdot \sqrt{\left(\frac{N}{\zeta}\right)^{1-\mu}} \text{ queries.} \tag{54}$$

Proof. This proof is equivalent to that of Lemma 20 except we use $U = [k]$. Note, the listed bound has no dependence on k; this is because $k \leq \zeta$ and we express here the weaker bound with ζ. □

Likewise, a short corollary of Lemma 22 is the following. Construct the distribution $\mathcal{G}_{❀}^{(k)}$ by the following procedure:

1. Sample a graph G from the restriction of the distribution $P_{M,\ell}$ to graphs with a connected component of exactly $[k]$.
2. Sample a permutation (π, π^{-1}) from $\Pi_{❀}^{(k)}$.
3. Output $(\pi^{-1}(G), r_G)$ where r_G is the random coin flips that would have generated G when sampling according to $P_{M,\ell}$. The oracle will be divided into a queryable component of $(\pi^{-1}(G))$ and a un-queryable component of r_G.

Corollary 23. *For $\delta < 1/4$, any quantum query algorithm $(1-\delta)$-distinguishing the distributions $\mathcal{G}_{❀}^{(k)}$ and $\mathcal{G}_{Ⓕ}^{(k)}$ where ❀ is a (μ, ζ, t)-sunflower and F is the corresponding core requires*

$$\geq \frac{1}{2}\left(1 - 2\sqrt{2\delta(1-2\delta)}\right)(1-4\delta) \cdot \sqrt{\left(\frac{N}{\zeta}\right)^{1-\mu}} \text{ queries.} \tag{55}$$

Proof. The corollary follows from Lemma 22 via a reduction from permutations to graphs exactly as in the proof of Corollary 21 from Lemma 20. □

6.3.3 Completing the proof

Proof of Lemma 19. Notice that for any $k \neq k'$, the support of $\mathcal{G}_{\text{❀}}^{(k)}$ is disjoint from the support of $\mathcal{G}_{\text{❀}}^{(k')}$, and likewise for $\mathcal{G}_{\textcircled{F}}^{(k)}$ and $\mathcal{G}_{\textcircled{F}}^{(k')}$. Let us again abuse notation and use $\mathcal{G}_{\text{❀}}^{(k)}$ to denote the support of the corresponding distribution. For each k, the lower bound from Corollary 23 is shown via an adversary bound with a relation R_k, and parameters $m, m', \ell_{\max}$, and moreover these parameters are the same for all k. Thus, we may construct a relation R between $\bigcup_k \mathcal{G}_{\text{❀}}^{(k)}$ and $\bigcup_k \mathcal{G}_{\textcircled{F}}^{(k)}$ by simply taking the union $R = \bigcup_k R_k$. This relation maintains the same parameters $m, m', \ell_{\max}$ due to the disjointness of supports for different k. Lastly, notice that $\bigcup_k \mathcal{G}_{\text{❀}}^{(k)}$ is equal to the support of $H_{\text{❀}}$ as described in the statement of Lemma 19. Likewise, for $H_{\textcircled{F}}$. Since $H_{\text{❀}}$ and $H_{\textcircled{F}}$ are uniform distributions over their support, by Corollary 14 using the relation R that we have constructed, the distributions are indistinguishable without the stated number of queries.

□

6.4 Statistical indistinguishability between random distributions

The final step of this section is to show that no algorithm can distinguish the distributions $H_{\textcircled{F}}$ and $P_{M,\ell}(F)$ with more than a negligible probability. This will be because these distributions are statistically close and this can be proven by a Chernoff tail bound.

Lemma 24. *The statistical distance between $H_{\textcircled{F}}$ and $P_{M,\ell}(F)$ is $O(N^{-3})$.*

Proof. Notice that the distribution $H_{\textcircled{F}}$ is equivalent to sampling a graph from $P_{M,\ell}(F)$ conditioned on consisting of ℓ connected components each with size $\in [(1-\gamma)z, (1+\gamma)z]$. By Lemma 10, with all but $O(N^{-3})$ probability, a graph from $P_{M,\ell}(F)$ satisfies this condition. Therefore, the statistical distance between these distributions is bounded by $O(N^{-3})$. □

7 Polynomial method lower bound

In this section, we prove that any quantum query algorithm cannot distinguish the graph distributions $P_{M,1}$ and $P_{M,\ell}(F)$. When $F = \emptyset$, this is equivalent to the problem studied by [6] in their quantum query lower bound:

Theorem 25 (Restatement of Theorem 2 of [6])**.** *For any sufficiently small constant $\epsilon_1 > 0$, any* deterministic *quantum query algorithm $\mathcal{A}$ distinguishing the distributions $P_{M,1}$ and $P_{M,\ell}$ for any $1 < \ell < N^{1/4}$ by probability ϵ_1. I.e.*

$$\mathop{\mathbf{E}}_{G \leftarrow P_{M,1}} \left[\Pr_{\mathcal{A}} \left[\mathcal{A}^G = 1 \right] \right] - \mathop{\mathbf{E}}_{G \leftarrow P_{M,\ell}} \left[\Pr_{\mathcal{A}} \left[\mathcal{A}^G = 1 \right] \right] \geq \epsilon_1 \tag{56}$$

must make at least $\Omega(N^{1/4}/\log N)$ queries. Here the Ω notation hides a dependence on ϵ_1.

The proof used in that result is very technical and builds on the polynomial method. Fortunately, we can show our query lower bound via a *reduction* to the [6] result. The reduction requires taking a short walk which mixes well by the expander mixing lemma.

Lemma 26. *Suppose there exists some F_0 and a q_1-query quantum algorithm that ϵ_1-distinguishes the distributions $P_{M,1} = P_{M,1}(F_0)$ and $P_{M,\ell}(F_0)$ for $\ell > 1$. Then there exists a q_2-query quantum algorithm that ϵ_2-distinguishes the distributions $P_{M,1}$ and $P_{M,\ell}$ with $q_2 = q_1 + O(N^{3/100})$ and $\epsilon_2 = \epsilon_1 - O(N^{-9/200})$.*

Intuitively, what this lemma says is that the set of points F_0 (which are in the same connected component) is not a helpful witness. Concretely, such a witness is negligibly more helpful than no witness at all. This is because, in the case of $P_{M,1}$ or $P_{M,\ell}$, the connected components are expanding and therefore the verifier can easily select a random subset of the points from a single connected component without any assistance from the prover. This can be shown via an application of the expander mixing lemma. Therefore, if a query algorithm exists for distinguishing $P_{M,1}$ and $P_{M,\ell}(F)$, it can be used as a subroutine for distinguishing $P_{M,1}$ and $P_{M,\ell}$ without any witness.

Furthermore, due to Ambainis, Childs, and Liu [6], we know Theorem 25 — i.e. that distinguishing the distributions without witnesses has a query lower bound. Therefore, the problem has a query lower bound even when a set of points F from a connected component are provided:

Corollary 27. *For any F_0 with $|F_0| \leq N^{1/100}$, any sufficiently small constant ϵ_1, and any ℓ with $1 < \ell < N^{1/4}$, any quantum query algorithm to ϵ_1-distinguish $P_{M,1}$ and $P_{M,\ell}(F_0)$ must make $\Omega(N^{1/4}/\log N)$ queries.*

Proof of Corollary 27. Suppose an algorithm making $q = o(N^{1/4}/\log N)$ queries existed. Then by Lemma 26 there exists an algorithm making $q' = q + O(N^{3/100}) = o(N^{1/4}/\log N)$ queries that distinguishes between $P_{M,1}$ and $P_{M,\ell}$ as well. However, this is impossible by Theorem 25. □

The remainder of this section is the proof of Lemma 26.

Let $\mathcal{A}_0$ be the hypothesized algorithm making q_1 queries to ϵ_1-distinguish $P_{M,1}$ and $P_{M,\ell}(F_0)$. We first claim that for any F with $|F| = |F_0|$, there exists an algorithm $\mathcal{A}_1$ that, given as classical input a list of all the vertices in F, and as oracle input an oracle G where G is a sample from either $P_{M,1}$ or $P_{M,\ell}(F)$, can ϵ_1-distinguish between these two cases using q_1 queries to G. The algorithm $\mathcal{A}_1$ is as follows:

1. Given F, compute a permutation π on V that maps F to F_0. (This step is not efficient in terms of runtime, but makes no queries to the oracle G.)
2. Run $\mathcal{A}_0$ with every query to G replaced by a query to $\pi(G)$. Return the answer given by $\mathcal{A}_0$.

The correctness of the algorithm follows from the fact that π maps the distribution $P_{M,\ell}(F)$ exactly to $P_{M,\ell}(F_0)$. Therefore, for all F such that $|F| = |F_0|$,

$$\mathop{\mathbf{E}}_{G \leftarrow P_{M,\ell}(F)} \left[\mathop{\mathbf{Pr}}_{\mathcal{A}_1}[\mathcal{A}_1(F,G) = 1] \right] - \mathop{\mathbf{E}}_{G \leftarrow P_{M,1}} \left[\mathop{\mathbf{Pr}}_{\mathcal{A}_1}[\mathcal{A}_1(F,G) = 1] \right] \geq \epsilon_1. \tag{57}$$

As this holds for all such F,

$$\mathop{\mathbf{E}}_{F} \mathop{\mathbf{E}}_{G \leftarrow P_{M,\ell}(F)} \left[\mathop{\mathbf{Pr}}_{\mathcal{A}_1}[\mathcal{A}_1(F,G) = 1] \right] - \mathop{\mathbf{E}}_{F} \mathop{\mathbf{E}}_{G \leftarrow P_{M,1}} \left[\mathop{\mathbf{Pr}}_{\mathcal{A}_1}[\mathcal{A}_1(F,G) = 1] \right] \geq \epsilon_1. \tag{58}$$

Next, we will show that the input of F can be removed from the algorithm: given just access to G, it is possible to compute a suitable F without making too many queries to the oracle. Specifically, we define the algorithm $\mathcal{A}_2$ to distinguish between $P_{M,1}$ and $P_{M,\ell}$ given only oracle access to G.

1. For a choice of t to be defined later, construct a set F_1 by starting at a random vertex v_0 and taking a $100t \cdot N^{1/100}$-step random walk along the graph as described in Lemma 3. If $|F_1| \geq |F|$, pick the first $|F|$ points from F_1 as the set F'. If not, output 0 (i.e. abort).

2. Run $\mathcal{A}_1$ on input F' with oracle access to G.

We will argue that for an appropriately chosen t, this algorithm achieves the success probability and query complexity claimed in the theorem. To do so, we will argue in two stages.

1. First, we argue that the distribution of F' chosen by random walk is very close to F' chosen uniformly at random from subsets of a connected component of G. This analysis uses the expander mixing lemma.

2. Second, we argue that the distribution over pairs (G, F') obtained after the first step of $\mathcal{A}_2$ is statistically indistinguishable from the distribution over pairs (G, F) sampled by first choosing a uniformly random $F \subseteq V$ and then choosing a random $G \leftarrow P_{M,\ell}(F)$. This will make use of Lemma 11, shown in Appendix A. By eq. (58), the algorithm $\mathcal{A}_1$ can ϵ_1-distinguish inputs distributed in this manner, and thus the second step of $\mathcal{A}_2$ can ϵ_2-distinguish inputs of $P_{M,1}$ and $P_{M,\ell}$ for ϵ_2 just slightly smaller than ϵ_1.

In our analysis, we will denote probabilities over the distribution of (G, F') generated by $\mathcal{A}_2$ by $\mathop{\mathbf{Pr}}_{\mathcal{A}_2}[\cdot]$ and probabilities over the distribution of (G, F) obtained by first sampling $F \subseteq V$, and then sampling $G \leftarrow P_{M,\ell}(F)$ by $\mathop{\mathbf{Pr}}_{F \text{ then } G}[\cdot]$. The notation $\mathop{\mathbf{Pr}}_{\text{unif}}[\cdot]$ denotes the distribution over F obtained by first picking a uniformly random vertex v in G, and then picking F to be a uniformly random subset of the connected component of G containing v with size $|F_0|$.

7.1 From random walk sampling to uniform sampling

Henceforth, define *expander walk sampling* as the sampling procedure of selecting a uniformly random vertex as the initial vertex v_1, then subsequently taking t steps of a lazy random walk (as defined in the expander mixing lemma, Lemma 3) to choose v_2, and so forth. In this case, the graph and the integer t will be clear from context.

We start by showing a sequence of claims that establish that if the expander walk sampling procedure for generating F' starts in a connected component C of G with size $|C| = K$ and expansion α, then the distribution over sets F' generated by the random walk is close to uniformly sampling points from C. Our main result here will be Claim 30.

Claim 28. *Let $\delta = (1 - \alpha/2)^t$ and let r be a natural number with $rK\delta < 1$. The for any sequence of r vertices $v_1, \ldots, v_r$, the probability $\mathop{\mathbf{Pr}}_{\text{unif}}[\cdot]$ that this sequence was obtained by iid random sampling and the probability $\mathop{\mathbf{Pr}}_{\text{walk}}[\cdot]$ that it was obtained by expander walk sampling differ by*

$$\left| \mathop{\mathbf{Pr}}_{\text{unif}}[v_1, \ldots, v_r] - \mathop{\mathbf{Pr}}_{\text{walk}}[v_1, \ldots, v_r] \right| \leq \left(\frac{1}{K}\right)^r \cdot \left(rK\delta + (rK\delta)^2 \frac{1}{1 - rK\delta}\right). \tag{59}$$

Proof. The proof is by direct calculation and application of the expander mixing lemma. First we show one side of the bound.

$$\mathop{\mathbf{Pr}}_{\text{unif}}[v_1, \ldots, v_r] = \left(\frac{1}{K}\right)^r \tag{60a}$$

$$\mathop{\mathbf{Pr}}_{\text{walk}}[v_1, \ldots, v_r] \leq \left(\frac{1}{K} + \delta\right)^r \tag{60b}$$

$$= \left(\frac{1}{K}\right)^r \cdot (1 + K\delta)^r \tag{60c}$$

By hypothesis $K\delta < 1$. Using this, we have the estimate

$$(1 + K\delta)^r = 1 + rK\delta + \sum_{j=2}^{r} \binom{r}{j} (K\delta)^j \tag{61a}$$

$$\leq 1 + rK\delta + \sum_{j=2}^{r} (rK\delta)^j \tag{61b}$$

$$\leq 1 + rK\delta + (rK\delta)^2 \frac{1}{1 - (rK\delta)} \tag{61c}$$

Substituting this bound into Equation (60c), we obtain

$$\mathop{\mathbf{Pr}}_{\text{walk}}[v_1, \ldots, v_r] - \mathop{\mathbf{Pr}}_{\text{unif}}[v_1, \ldots, v_r] \leq \left(\frac{1}{K}\right)^r \cdot \left(rK\delta + (rK\delta)^2 \frac{1}{1 - rK\delta}\right). \tag{62}$$

Now for the other side. Again, applying the expander mixing lemma,

$$\mathop{\mathbf{Pr}}_{\text{walk}}[v_1, \ldots, v_r] \geq \left(\frac{1}{K} - \delta\right)^r \tag{63a}$$

$$= \left(\frac{1}{K}\right)^r (1 - K\delta)^r \tag{63b}$$

$$\geq \left(\frac{1}{K}\right)^r (1 - rK\delta), \tag{63c}$$

where in the last step we have used Bernoulli's inequality and the assumption that $K\delta < 1$. □

The following claim will be used to bound the probability that the expander walk sampling procedure aborts, by instead bounding the probability that iid sampling fails to generate enough distinct points.

Claim 29. *The probability that $T \geq 100|F|$ iid samples from C contain fewer than $|F|$ distinct vertices is at most* $\exp(-T/16)$.

Proof. Let X_v be the event that vertex $v \in C$ is sampled. It is clear that

$$\mathbf{E}[X_v] = 1 - \left(1 - \frac{1}{K}\right)^T \tag{64a}$$

$$= \frac{T}{K} - \sum_{j=2}^{T} \binom{T}{j} \cdot \left(\frac{-1}{K}\right)^j \tag{64b}$$

$$\geq \frac{T}{K} - \sum_{j=2}^{T} \left(\frac{T}{K}\right)^j \tag{64c}$$

$$\geq \frac{T}{K} - \left(\frac{T}{K}\right)^2 \frac{1}{1 - T/K}. \tag{64d}$$

In our setting $K \geq (1-\gamma)z \geq N^{0.5}$ and $T = 100|F| = 100N^{1/100}$, so for sufficiently large N we have

$$\frac{T}{K} \geq \mathbf{E}[X_v] \geq \frac{T}{K} - \frac{2 \cdot 10^4}{K \cdot N^{0.48}} \geq \frac{T-1}{K}. \tag{65}$$

So the expected number of vertices that are sampled is

$$\mu \stackrel{\text{def}}{=} \mathbf{E}[N_{\text{sampled}}] = \sum_v \mathbf{E}[X_v] \tag{66a}$$

$$\geq T - 1. \tag{66b}$$

Moreover, each even X_v is independent, so the number of sampled vertices concentrates well around its mean. By a Chernoff bound we have

$$\mathbf{Pr}[N_{\text{sampled}} \leq (1-\epsilon)\mu] < \exp\left(-\frac{\epsilon^2 \mu}{2}\right). \tag{67}$$

Setting $\epsilon = 0.5$ we have

$$\mathbf{Pr}[N_{\text{sampled}} \leq 50|F| - 1] < \exp\left(-\frac{(T-1)}{8}\right) \leq \exp(-T/16). \tag{68}$$

□

We now combine these two claims and apply them to our setting. Define the event $[F' \leftarrow G]$ if F' is the set of vertices selected from the graph G. Define the distribution $\mathbf{Pr}_{\text{unif}}[\cdot]$ corresponding to first choosing a connected component C with probability proportional to $|C|$, taking r uniform iid samples from the connected component C and setting F' to be the first $|F|$ distinct sampled points. Likewise, define the distribution $\mathbf{Pr}_{\mathcal{A}_2}$ corresponding to choosing a random vertex v in G, taking C to be the connected component containing v, taking r samples according to an expander random walk in C initialized at v with t steps between samples, and then setting F' to be the first $|F|$ distinct sampled points. Set K to be the maximum size of a connected component in G and let $\delta = (1+\alpha/2)^t$. Then we have the following distance bound between the distributions.

Claim 30. *Suppose r, K, δ are such that $rK\delta \leq 10/11$. For any F' of size $|F'| = |F|$, let $\delta_C(F') = \mathbf{Pr}_{\text{unif}}[F' \leftarrow G] - \mathbf{Pr}_{\mathcal{A}_2}[F' \leftarrow G]$. Then*

$$|\delta_G(F')| \leq \left(rK\delta + (rK\delta)^2 \frac{1}{1-rK\delta}\right) \leq 10rK\delta. \tag{69}$$

Moreover, $\mathop{\mathbf{Pr}}_{\mathcal{A}_2}[\text{abort}] \leq 10rK\delta + \exp(-T/16)$.

Proof. For a sequence $v_1, \ldots, v_r$ of vertices, let $X_{v_1,\ldots,v_r \to F'}$ be the event that F' is the set of the first $|F|$ distinct vertices in $v_1, \ldots, v_r$.

$$|\delta_G(F')| = \left| \sum_{v_1,\ldots,v_r} X_{v_1,\ldots v_r \to F'} \cdot \left(\mathop{\mathbf{Pr}}_{\text{unif}}[v_1, \ldots, v_r] - \mathop{\mathbf{Pr}}_{\mathcal{A}_2}[v_1, \ldots, v_r] \right) \right| \tag{70a}$$

$$\leq \sum_{v_1,\ldots,v_r} X_{v_1,\ldots v_T \to F'} \cdot \left| \mathop{\mathbf{Pr}}_{\text{unif}}[v_1, \ldots, v_r] - \mathop{\mathbf{Pr}}_{\mathcal{A}_2}[v_1, \ldots, v_r] \right| \tag{70b}$$

$$\leq K^r \cdot \left(\frac{1}{K}\right)^r \cdot \left(rK\delta + (rK\delta)^2 \frac{1}{1-rK\delta}\right) \tag{70c}$$

$$= \left(rK\delta + (rK\delta)^2 \frac{1}{1-rK\delta}\right) \tag{70d}$$

$$\leq 10rK\delta. \tag{70e}$$

where we have used Claim 28 to bound the difference in probabilities of each sequence of vertices. Now, to bound the abort probability, let $X_{v_1,\ldots,v_r\to\text{abort}}$ be the event that $v_1,\ldots,v_r$ contain fewer than $|F|$ distinct vertices.

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{\mathcal{A}_2}[\text{abort}] &= \sum_{v_1,\ldots,v_r} X_{v_1,\ldots,v_r\to\text{abort}} \cdot \mathop{\mathbf{Pr}}_{\mathcal{A}_2}[v_1,\ldots,v_r] && \text{(71a)}\\
&\le 10rK\delta + \sum_{v_1,\ldots,v_r} X_{v_1,\ldots,v_r\to\text{abort}} \cdot \mathop{\mathbf{Pr}}_{\text{unif}}[v_1,\ldots,v_r] && \text{(71b)}\\
&\le 10rK\delta + \exp(-99|F|/8), && \text{(71c)}
\end{aligned}$$

where we have used Claim 28 to replace $\mathbf{Pr}_{\mathcal{A}_2}[\cdot]$ by $\mathbf{Pr}_{\text{unif}}[\cdot]$ and then Claim 29 to bound the abort probability of uniform sampling. □

7.2 From $\mathbf{Pr}_{\mathcal{A}_2}[\cdot]$ to $\mathbf{Pr}_{F \text{ then } G}[\cdot]$

We will now proceed to the main argument showing that the pairs (G,F) sampled by $\mathcal{A}_2$ are distributed close to the distribution expected by $\mathcal{A}_1$.

G has expanding components with high probability To start off, first note that by Lemma 10, with probability at least $1-O(N^{-3})$ a graph drawn from $P_{M,\ell}(F)$, for any F of size $\le N^{1/100}$, will consist of ℓ connected-components which are α-expanders and have size between $[(1-\gamma)z,(1+\gamma)z]$. Since ϵ_1 is a constant, for sufficiently large N, we can restrict to the situation that the graph is of this form and account for this factor in the end. Henceforth set $K_0=(1+\gamma)z$; we are guaranteed that every component has size at most K_0.

Relating the probabilities In the case that each connected component is an α-expander, observe that the probability of every valid pair (G,F') is approximately a constant p independent of G and F'. Also recall that the event $F'\leftarrow G$ is the event that F' is the set of vertices selected from G. Moreover, recall the distributions $\mathcal{D}_1$ and $\mathcal{D}_2$ from Lemma 11, and notice that $\mathcal{D}_2$ is exactly the distribution $\mathbf{Pr}_{F \text{ then } G}$ defined above. We define

$$\begin{aligned}
\delta_{1,2}(G,F') &\stackrel{\text{def}}{=} \mathop{\mathbf{Pr}}_{\mathcal{D}_1}[(G,F')] - \mathop{\mathbf{Pr}}_{\mathcal{D}_2}[(G,F')], && \text{(72a)}\\
\delta_G(F') &\stackrel{\text{def}}{=} \mathop{\mathbf{Pr}}_{\mathcal{A}_2}[F'\leftarrow G] - \mathop{\mathbf{Pr}}_{\text{unif}}[F'\leftarrow G]. && \text{(72b)}
\end{aligned}$$

We will now start with the $\mathbf{Pr}_{\mathcal{A}_2}$ distribution and bound its distance from $\mathbf{Pr}_{F \text{ then } G}$.

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{\mathcal{A}_2}[(G,F')] &= \mathop{\mathbf{Pr}}_{P_{M,\ell}}[G] \cdot \mathop{\mathbf{Pr}}_{\mathcal{A}_2}[F'\leftarrow G] && \text{(73a)}\\
&= \mathop{\mathbf{Pr}}_{P_{M,\ell}}[G] \cdot \left(\mathop{\mathbf{Pr}}_{\text{unif}}[F'\leftarrow G] + \delta_G(F')\right) = \mathop{\mathbf{Pr}}_{\mathcal{D}_1}[(G,F')] + \delta_G(F') \cdot \mathop{\mathbf{Pr}}_{P_{M,\ell}}[G] && \text{(73b)}\\
&= \mathop{\mathbf{Pr}}_{\mathcal{D}_2}[(G,F')] + \delta_{1,2}(G,F') + \delta_G(F') \cdot \mathop{\mathbf{Pr}}_{P_{M,\ell}}[G] && \text{(73c)}\\
&= \mathop{\mathbf{Pr}}_{F \text{ then } G}[(G,F')] + \delta_{1,2}(G,F') + \delta_G(F') \cdot \mathop{\mathbf{Pr}}_{P_{M,\ell}}[G]. && \text{(73d)}
\end{aligned}$$

We may now bound the total variational distance between the two sides.

$$\left\|\Pr_{\mathcal{A}_2}[\cdot] - \Pr_{F \text{ then } G}[\cdot]\right\| = \frac{1}{2}\sum_{(G,F')}\left|\Pr_{\mathcal{A}_2}[(G,F')] - \Pr_{F \text{ then } G}[(G,F')]\right| \tag{74a}$$

$$\leq \frac{1}{2}\sum_{(G,F')}\left(|\delta_{1,2}(G,F')| + \Pr_{P_{M,\ell}}[G]\cdot|\delta_G(F')|\right) \tag{74b}$$

$$\leq \|\mathcal{D}_1 - \mathcal{D}_2\| + \frac{1}{2}\binom{K_0}{|F'|}\cdot\max_{G,F'}|\delta_G(F')| \tag{74c}$$

$$\leq 3N^{-9/200} + K_0^{|F|}\cdot(10rK_0\delta) \tag{74d}$$

$$= 3N^{-9/200} + ((1+\gamma)z)^{|F|+1}\cdot(10\cdot(100|F|)\cdot(1-\alpha/2)^t) \tag{74e}$$

$$\leq 3N^{-9/200} + \left(2N^{9/10}\right)^{N^{1/100}+1}\cdot 1000N^{1/100}\cdot(1-\alpha/2)^t \tag{74f}$$

$$= 3N^{-9/200} + 2000\cdot 2^{N^{0.01}}\cdot N^{0.9N^{0.01}+0.91}\cdot(1-\alpha/2)^t. \tag{74g}$$

A total distance bound of $O(N^{-9/200})$ can be achieved if

$$2^{N^{0.01}}\cdot N^{0.9N^{0.01}+0.91}\cdot(1-\alpha/2)^t \leq N^{-9/200} \tag{75a}$$

$$N^{0.01} + (0.9N^{0.01}+0.91)\cdot\log N + t\cdot\log(1-\alpha/2) \leq -\frac{9}{200}\log N \tag{75b}$$

$$\left(N^{0.01}\left(\frac{1}{\log N}+0.9\right)+0.955\right)\frac{\log N}{\log(1/(1-\alpha/2))} \leq t. \tag{75c}$$

So setting $t = \Theta(N^{0.02})$ is sufficient.

Given this choice of t, let us know calculate the chance that the sampling of F' aborts. By Claim 30, this is at most $10rK\delta + \exp(-T/16) = O(N^{-9/200}) + \exp(-100N^{0.01}/16) = O(N^{-9/200})$

Thus, the total error probability of $\mathcal{A}_2$ equals the error probability of $\mathcal{A}_1$ up to

$$\underbrace{O(N^{-3})}_{\text{Sample a bad graph}} + \underbrace{O(N^{-9/200})}_{\text{Changing } \mathcal{A}_2 \text{ to } F \text{ then } G} + \underbrace{O(N^{-9/200})}_{\text{Sampling } F' \text{ aborts}}, \tag{76}$$

yielding $\epsilon_2 = \epsilon_1 - O(N^{-9/200})$ as claimed in theorem. And the total query complexity assuming not aborting can be calculated as follows. Recall that t was chosen to be $\Theta(N^{0.02})$. The total number of additional queries over $\mathcal{A}_1$ is thus the number of steps in the walk which is $100N^{1/100}\cdot t \leq O(N^{0.03})$. Thus, this algorithm has total query complexity $q + O(N^{0.03})$ and distinguishes with probability ϵ_2 as claimed.

8 Wrapping up the proof of Theorem 1

First, we need to note that the distributions B_S and $P_{M,1}$ which we used as proxies for YES and NO instances are not fully supported on YES and NO instance graphs, respectively. However, they are very close. For every $S \subset [N]$ of size ζ, let $\widetilde{B}_S$ be the restriction of the distribution B_S (defined in Section 6.2) to graphs with ℓ connected components each consisting of between $(1-\gamma)z$ and $(1+\gamma)z$ vertices. By Corollary 17, the statistical distance between B_S and $\widetilde{B}_S$ is $O(N^{-3})$. The YES instances for Theorem 1 are the $\{\widetilde{B}_S\}$.

We consider a single NO instance of $\widetilde{P}_{M,1}$ where $\widetilde{P}_{M,1}$ is the restriction of $P_{M,1}$ to graphs which are α-expanders. The statistical distance between these two distributions is $O(N^{-3})$ by Lemma 10.

Furthermore, we can verify that the supports of $\widetilde{P}_{M,1}$ and $\widetilde{B}_S$ are far apart in Hamming distance. Consider graphs G_1 and G_ℓ from either support, respectively. Consider a connected component C from G_ℓ. In the graph G_ℓ, all the edges on C stay within G_ℓ, but since G_1 is an α-expander and also a $1/10^2$-edge expander (see proof of Lemma 10)), then in G_1, a $\geq 1/10^4$ fraction of the edges emanating from C leave C. As this holds for all components C since $|C| \ll N/2$, then the Hamming distance between the adjacency lists of G_1 and G_ℓ is $\Omega(N)$. As this holds for all graphs G_1 and G_ℓ, then the Hamming distance bound between the supports hold.

8.1 QMA algorithm

For completeness, from Corollary 17, we know that the algorithm $\mathcal{A}_{\mathsf{QMA}}$ with witness state $|S\rangle$ answers distribution $\widetilde{B}_S$ with probability at least $\geq 1-O(N^{-1/20})$. For soundness, from Lemma 10, we know that $\widetilde{P}_{M,1}$ is an $1/(2\cdot 10^8)$-expander with probability $\geq 1-O(N^{-3})$. Therefore, by Lemma 12, the algorithm $\mathcal{A}_{\mathsf{QMA}}$ accepts with probability at most

$$\leq 1-\frac{1}{4}\cdot\frac{1}{(2\cdot 10^8)}+O(N^{-3})\leq 1-\frac{1}{9\cdot 10^8}. \tag{77}$$

By parallel repetition $9\cdot 10^6 = O(1)$ times, we yield a quantum algorithm with ≤ 0.01 soundness.

8.2 QCMA algorithm

We argue now that any QCMA algorithm with completeness 0.99 and soundness 0.01 either requires an exponentially long proof or an exponential number of quantum queries. This is done by arguing that any algorithm with a short proof and few queries cannot have such a large completeness and soundness gap. Assume, therefore, that there exists a QCMA algorithm with a

$$q\leq\frac{n\cdot N^{1/100}}{2000}\text{-bit proof and } f=O(N^{1/50}) \text{ quantum queries} \tag{78}$$

and a completeness and soundness gap of ≥ 0.98. By Lemma 18, there exists a

$$\left(\frac{1}{100},\zeta,\frac{2000q}{n}\right)\text{-sunflower } ❀ \tag{79}$$

with core F and a f-query deterministic quantum algorithm $\mathcal{A}$ that accepts each distribution $\widetilde{B}_S$ for $S\in ❀$ with probability ≥ 0.99 and accepts $\widetilde{P}_{M,1}$ with at most ≤ 0.01 probability. It also accepts that accepts each distribution B_S for $S\in ❀$ with probability $\geq 0.99-O(N^{-3})$ and accepts $P_{M,1}$ with at most $\leq 0.01+O(N^{-3})$ probability. Then with the assumed number of queries, we can apply Lemma 19 with $\delta=1/10$ to argue that $\mathcal{A}$ must accepts the distribution H_{Ω_F} with probability $\geq 0.09-O(N^{-3})$. Next, by Lemma 24, $\mathcal{A}$ must accept the distribution $P_{M,\ell}(F)$ with probability $\geq 0.09-2\cdot O(N^{-3})\geq 0.08$. We conclude by applying Corollary 27. Therefore, $\mathcal{A}$ must accept the distribution $P_{M,1}$ with probability >0.02, a contradiction.

9 Concluding remarks

9.1 Relation to the Fefferman and Kimmel [13] construction

One can think of the result stated in this work as applying the QCMA lower bounding techniques developed by Fefferman and Kimmel [13] to the expander distinguishing problem originally studied by Ambainis, Childs, and Liu [6].

At a high level, in the *in-place permutation oracle* QMA and QCMA separation of [13], the goal was to distinguish between permutations $\pi : [N] \to [N]$ such that $\pi^{-1}([\sqrt{N}])$ is mostly $(2/3)$ supported on odd numbers from permutations mostly supported on even numbers. The original idea in Fefferman and Kimmel was that if the oracle π was provided as a classical oracle (an Nn-bit list $[\pi(1), \pi(2), \ldots, \pi(N)]$) then the subset state $|\xi_{\text{ideal}}\rangle = |\pi^{-1}([\sqrt{N}])\rangle$ would be a good quantum witness. By measuring the last qubit of a witness $|\xi\rangle$, the verifier can decide if the set $\pi^{-1}([\sqrt{N}])$ is supported mostly on either odd numbers or even numbers. What remains to verify is that the witness $|\xi\rangle$ provided is indeed $|\xi_{\text{ideal}}\rangle$. The hope would be to use the oracle for π to verify the statement as the sate $|[\sqrt{N}]\rangle$ can be easily verified by measuring in the Hadamard basis.

However, due to the index-erasure problem, a classical oracle for verifying that $|\xi\rangle = |\xi_{\text{ideal}}\rangle$ would need to allow implementation of both π and π^{-1}. However, if the oracle π^{-1} is provided, then there is a BQP algorithm for this problem. Simply, pick a random $j \in [\sqrt{N}]$ and then check if $\pi^{-1}(j)$ is odd or even. The solution in [13] was to define the oracle instead as an "in-place oracle" for π, meaning a unitary defined as $\sum_j |\pi(j)\rangle\langle j|$. Then the verifier can verify that $|\xi\rangle = |\xi_{\text{ideal}}\rangle$ and yet the BQP algorithm no longer holds.

Fefferman and Kimmel had to make one more modification to prove a QMA and QCMA oracle separation: they considered distributions over in-place oracles which mapped to the same ideal quantum witness $|\xi_{\text{ideal}}\rangle$. This was because it seems to be beyond current techniques to prove classical lower bounds without forcing a large structured set of permutations to all share the same witness — otherwise, for all we know, there might be a mathematical fact about permutations which yields a short classical certificate for any individual permutation. So the oracle is defined as a distribution over unitaries — i.e. a completely positive trace preserving (CPTP) map.

Notice that this work takes much inspiration form [13]; the quantum witnesses for both our work and [13] are subset states and we also consider distributions over oracles with the same (or similar) ideal quantum witness. This is because we are unsure how to prove that there is no property of a specific regular graph which yields a short classical witness. We elaborate on why such an impossibility result is hard to prove in the next subsection. What our result principally improves on is that the underlying oracle can be a classical string instead of a unitary.

9.2 Difficulties in proving stronger statements

Recall that our QMA upper bound does not require the setup of distributions over oracles — it was only included to prove the QCMA lower bound. How much harder is it (or is it even possible) to prove a QCMA lower bound without considering distributions?

As pointed out to us by William Kretschmer [21], if one considers *average-case* algorithms instead of *worst-case* algorithms, then this problem is $\in \mathsf{RNP}^G$, the average-case analog of NP^G. This is because the average-case version of the expander distinguishing problem is to distinguish the distributions $P_{M,\ell}$ and $P_{M,1}$. And there is a simple randomized algorithm for this problem with a classical witness. Let us recall that for a d-regular graph, the expected number of triangles in a connected component is $\Theta(d^3)$ independent of the number of vertices in the component. A similar analysis can be done for $P_{M,\ell}$ and $P_{M,1}$, to show that a random graph from $P_{M,\ell}$ has $\Theta(\ell d^3)$ triangles whereas $P_{M,1}$ has $\Theta(d^3)$ triangles. Therefore, a *classical* witness for the statement that the graph (with high probability) is drawn from $P_{M,\ell}$ (instead of $P_{M,1}$) is a list of $100 \cdot \Theta(d^3)$ triangles from the graph. This witness is easily verifiable and correctly distinguishes with high probability.

Notice that this RNP^G algorithm does not solve the expander distinguishing problem in the worst-case; since graphs exist in both distributions which are triangle-free (with

constant probability). Furthermore, it cannot distinguish the distributions considered in Theorem 1 because the proof relies on finding triangles which is property of the graph not deducible from only knowing the connected components.

But it does highlight a principal roadblock in extending Theorem 1 to distinguishing oracles that are not distributions. It is entirely possible that there exists a property of graphs revealed by looking at the edges that distinguishes graphs with many connected components from graphs with a single expanding connected component. To the best of our knowledge, we do not know of any such property but proving that none exist is beyond the techniques shown here.

Lastly, if we consider the expander distinguishing problem when in the YES case we are promised that every connected component has size at most $0.99N$, then this problem is in coAM^G. When the graph is a NO instance, the verifier can select two random points and the prover can always find a path of length $O(\log N) = O(n)$ between the two. However, when the graph is disconnected and no component is too big, with probability $\geq 1/50$, no path exists.

Therefore, our constructed oracle very finely separates the classes QMA and QCMA in the sense that small perturbations of the problem might be very easy.

References

[1] Scott Aaronson. On perfect completeness for QMA. *Quantum Info. Comput.*, 9(1): 81–89, January 2009. ISSN 1533-7146. DOI: 10.26421/qic9.1-2-5.

[2] Scott Aaronson. Open problems related to quantum query complexity. *ACM Transactions on Quantum Computing*, 2(4), December 2021. ISSN 2643-6809. DOI: 10.1145/3488559.

[3] Scott Aaronson and Greg Kuperberg. Quantum versus classical proofs and advice. In *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07)*, pages 115–128, 2007. DOI: 10.1109/CCC.2007.27.

[4] Dorit Aharonov and Tomer Naveh. Quantum NP - A survey, 2002.

[5] Andris Ambainis. Quantum lower bounds by quantum arguments. *J. Comput. Syst. Sci.*, 64(4):750–767, jun 2002. ISSN 0022-0000. DOI: 10.1006/jcss.2002.1826. URL https://doi.org/10.1006/jcss.2002.1826.

[6] Andris Ambainis, Andrew M. Childs, and Yi-Kai Liu. Quantum property testing for bounded-degree graphs. In Leslie Ann Goldberg, Klaus Jansen, R. Ravi, and José D. P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 365–376, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg. ISBN 978-3-642-22935-0. DOI: 10.1007/978-3-642-22935-0_31.

[7] Anurag Anshu, Nikolas P. Breuckmann, and Chinmay Nirkhe. NLTS hamiltonians from good quantum codes. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023, page 1090–1096, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450399135. DOI: 10.1145/3564246.3585114.

[8] Atul Singh Arora, Alexandru Gheorghiu, and Uttam Singh. Oracle separations of hybrid quantum-classical circuits. 2022. DOI: 10.48550/arXiv.2201.01904.

[9] Charles H. Bennett, Ethan Bernstein, Gilles Brassard, and Umesh V. Vazirani. Strengths and weaknesses of quantum computing. *SIAM J. Comput.*, 26(5):1510–1523, 1997. DOI: 10.1137/S0097539796300933.

[10] Adam D. Bookatz. QMA-complete problems. 2012. DOI: 10.48550/ARXIV.1212.6312.

[11] Sergey Bravyi and Barbara Terhal. Complexity of stoquastic frustration-free hamiltonians. *SIAM J. Comput.*, 39(4):1462–1485, nov 2009. ISSN 0097-5397. DOI: 10.1137/08072689x.

[12] Sergey Bravyi, David P. Divincenzo, Roberto Oliveira, and Barbara M. Terhal. The complexity of stoquastic local hamiltonian problems. *Quantum Info. Comput.*, 8(5): 361–385, may 2008. ISSN 1533-7146. DOI: 10.26421/qic8.5-1.

[13] Bill Fefferman and Shelby Kimmel. Quantum vs. classical proofs and subset verification. In Igor Potapov, Paul G. Spirakis, and James Worrell, editors, *43rd International Symposium on Mathematical Foundations of Computer Science, MFCS 2018, August 27-31, 2018, Liverpool, UK*, volume 117 of *LIPIcs*, pages 22:1–22:23. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018. DOI: 10.4230/LIPIcs.MFCS.2018.22.

[14] Honghao Fu. Personal Communication, Oct 2022.

[15] Alex Bredariol Grilo, Iordanis Kerenidis, and Jamie Sikora. QMA with subset state witnesses. In Giuseppe F. Italiano, Giovanni Pighizzini, and Donald T. Sannella, editors, *Mathematical Foundations of Computer Science 2015*, pages 163–174, Berlin, Heidelberg, 2015. Springer Berlin Heidelberg. ISBN 978-3-662-48054-0. DOI: 10.1007/978-3-662-48054-0_14.

[16] Sandy Irani, Anand Natarajan, Chinmay Nirkhe, Sujit Rao, and Henry Yuen. Quantum Search-To-Decision Reductions and the State Synthesis Problem. In Shachar Lovett, editor, *37th Computational Complexity Conference (CCC 2022)*, volume 234 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 5:1–5:19, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. ISBN 978-3-95977-241-9. DOI: 10.4230/LIPIcs.CCC.2022.5.

[17] Stephen P Jordan, David Gosset, and Peter J Love. Quantum-merlin-arthur-complete problems for stoquastic hamiltonians and markov matrices. 81(3), 3 2010. ISSN 1050-2947. DOI: 10.1103/PHYSREVA.81.032331.

[18] A.Yu. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303 (1):2 – 30, 2003. ISSN 0003-4916. DOI: https://doi.org/10.1016/S0003-4916(02)00018-0.

[19] Hartmut Klauck and Supartha Podder. Two results about quantum messages. In *International Symposium on Mathematical Foundations of Computer Science*, pages 445–456. Springer, 2014. DOI: 10.1007/978-3-662-44465-8_38.

[20] Dexter Kozen. Lower bounds for natural proof systems. In *Proceedings of the 18th Annual Symposium on Foundations of Computer Science*, SFCS '77, page 254–266, USA, 1977. IEEE Computer Society. DOI: 10.1109/SFCS.1977.16.

[21] William Kretschmer. Personal Communication, Jul 2022.

[22] Andrew Lutomirski. Component mixers and a hardness result for counterfeiting quantum money, 2011.

[23] Michael A Nielsen and Isaac L Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511976667.

[24] Chinmay Nirkhe. *Lower bounds on the complexity of quantum proofs*. PhD thesis, EECS Department, University of California, Berkeley, Nov 2022. URL `http://www2.eecs.berkeley.edu/Pubs/TechRpts/2022/EECS-2022-236.html`.

[25] Chinmay Nirkhe. NLTS Hamiltonians from codes, 2022. URL `https://simons.berkeley.edu/events/quantum-colloquium-nlts-hamiltonians-codes`. Simons Institute for the Theory of Computing Quantum Colloquium. Panel Umesh Vazirani, Dorit Aharanov, Matthew Hastings, Anand Natarajan, and Chinmay Nirkhe.

A Omitted concentration inequalities for random graphs

Lemma 10 (Adaptation of Lemma 16 of [6])**.** *Assume $|F| \leq N^{1/100}$. Then with probability at least $\geq 1 - O(N^{-3})$, a graph drawn from distribution $P_{M,\ell}(F)$ consists of exactly ℓ connected components each α-expanding and consisting of between $(1-\gamma)z$ and $(1+\gamma)z$ vertices.*

Likewise, the probability that a graph drawn from the distribution $P_{M,1}$ is α-expanding is $\geq 1 - O(N^{-3})$.

Proof. The proof for $P_{M,1}$ follows as a special case for $\ell = 1$ and $F = \emptyset$. Consider a graph drawn from the distribution $P_{M,\ell}(F)$ as described in Section 4. Let $X_{k,j}$ for $k \in [\ell]$ and $j \in V$ be the indicator random variable that $k_j = i$. Let $X_k = \sum_{j \in V} X_{k,j}$, the size of $\iota^{-1}(V_k)$. Since the image of a connected component under ι must lie in some V_k, then the size of any connected component is upper-bounded by $(1+\gamma)z$. For any $k > 1$, $\mathbf{E}(X_k) = (N - |F|)/\ell = z - |F|/\ell$. Since $|F| \ll \ell$, then $|F|/\ell < \gamma z/2$. Therefore, by a Chernoff bound,

$$\begin{aligned}
\mathbf{Pr}\left[|X_k - z| \geq \gamma z\right] &\leq \mathbf{Pr}\left[|X_k - \mathbf{E}(X_k)| \geq \gamma\, \mathbf{E}(X_k)\right] && \text{(80a)}\\
&\leq 2\exp\left(-\frac{(\gamma/2)^2\, \mathbf{E}(X_k)}{3}\right) && \text{(80b)}\\
&\leq 2\exp\left(-\frac{\gamma^2 N}{12\ell}\right) && \text{(80c)}\\
&\leq 2\exp\left(-\frac{N^{7/10}}{12}\right). && \text{(80d)}
\end{aligned}$$

For $k = 1$, the situation is only slightly different since already $|F|$ terms are guaranteed to be in X_1.

A union bound of all these probabilities bounds shows that with all but

$$2N^{1/10} \cdot \exp\left(-\frac{N^{7/10}}{12}\right) \text{ probability,} \tag{81}$$

the preimage under ι of every subset V_k has size $\in [(1-\gamma)z, (1+\gamma)z]$. It remains now to show that, conditioned on this holding, the preimages of every subset V_k are good expanders. Recall the definition of a β-edge expander is any graph such that for any subset $U \subset V$ of at most half the vertices, $|\partial U| \geq \beta |U|$, where ∂U is the set of neighbors of U excluding U itself.

For any $k \in [\ell]$, let $B_k = \iota^{-1}(V_k)$ be the preimages under ι of the defined regions of G'. If the induced graph on B_k is not a β-edge expander, there exists some subset U_1 of at most half the vertices of B_k and a subset $U_2 \subseteq B_k \setminus U_1$ such that every non-self neighbor of U_1 is contained in U_2 and, in particular, $|U_2| < \beta |U_1|$. Let E_{U_1,U_2} be the event that this occurs. Note that for a fixed B_k, the probability of the event $\mathbf{Pr}[E_{U_1,U_2}] = g(|U_1|, |U_2|)$ for some function g — i.e. the probability only depends on the sizes of the two sets. This is because the graph G' and the injective function ι are picked independently. Therefore, a union bound on the probability that the connected component is not an edge expander

is

$$\mathbf{Pr}\left[B_k \text{is } \textit{not} \text{ a } \beta\text{-edge expander}\right] \leq \sum_{U_1:|U_1|\leq|B_k|/2} \sum_{U_2\subseteq V\setminus U_1,|U_2|<\beta|U_1|} \mathbf{Pr}[E_{U_1,U_2}] \tag{82a}$$

$$\leq \sum_{i=1}^{\zeta/2} \binom{\zeta}{i}\binom{\zeta}{\beta i} g(i,\beta i). \tag{82b}$$

We will soon prove that

$$g(i,\beta i) \leq \left(2\gamma + \frac{(1+\beta)i}{z}\right)^{\frac{di}{2}}. \tag{83}$$

Assuming eq. (83), we can bound the probability by using two bounds, one for small i and one for large i. For small i — i.e. whenever $(1+\beta)i \leq 6\gamma z$, then $g(i,\beta i) \leq (8\gamma)^{\frac{di}{2}}$ and so

$$\binom{\zeta}{i}\binom{\zeta}{\beta i} g(i,\beta i) \leq \zeta^{(1+\beta)i} \cdot \frac{8^{\frac{di}{2}}}{N^{\frac{di}{20}}} \tag{84a}$$

$$\leq z^{(1+\beta)i} \cdot 2^{(1+\beta)i} \cdot \frac{8^{\frac{di}{2}}}{N^{\frac{di}{20}}} \tag{84b}$$

$$\leq \left(\frac{2^{1+\beta+3d/2}}{N^{\frac{d}{20}-\frac{9}{10}(1+\beta)}}\right)^{i} \stackrel{\text{def}}{=} r^i \tag{84c}$$

Assuming $\beta < 4$, then $\frac{d}{20} > \frac{9}{10}(1+\beta)$, and therefore, for sufficiently large N, $r \ll \frac{1}{2}$. So this geometric series is bounded by

$$\sum_{(1+\beta)i\leq 6\gamma z} \binom{\zeta}{i}\binom{\zeta}{\beta i} g(i,\beta i) \leq \frac{r}{1-r} \leq r(1+2r) \leq 3r \leq N^{-(4-\beta)} \cdot 2^{154}. \tag{85}$$

For large i — i.e. whenever $(1+\beta)i > 6\gamma z$ and $i \leq \zeta/2$, note that

$$g(i,\beta i) \leq \left(\frac{4}{3}\frac{(1+\beta)i}{z}\right)^{\frac{di}{2}}. \tag{86}$$

Therefore,

$$\binom{\zeta}{i}\binom{\zeta}{\beta i} g(i,\beta i) \leq \left(\frac{e\zeta}{i}\right)^{i}\left(\frac{e\zeta}{\beta i}\right)^{\beta i}\left(\frac{4}{3}\frac{(1+\beta)i}{z}\right)^{\frac{di}{2}} \tag{87a}$$

$$\leq \left(\frac{(2e)^{1+\beta}}{\beta^\beta}\left(\frac{4(1+\beta)}{3}\right)^{\frac{d}{2}} \cdot \left(\frac{i}{z}\right)^{\frac{d}{2}-(1+\beta)}\right)^{i} \tag{87b}$$

$$\leq \left(\frac{(2e)^{1+\beta}}{\beta^\beta}\left(\frac{4(1+\beta)}{3}\right)^{\frac{d}{2}} \cdot \left(\frac{51}{100}\right)^{\frac{d}{2}-(1+\beta)}\right)^{i} \tag{87c}$$

where in the last line we used that $i < \zeta/2$ and $\zeta = (1+\gamma)z$ so for sufficiently large N, $i/z \leq 51/100$. For choice of $\beta = 1/100$ and $d = 100$, we get that

$$\binom{\zeta}{i}\binom{\zeta}{\beta i} g(i,\beta i) \leq 2^{-23i}. \tag{88}$$

Therefore,

$$\sum_{\frac{6\gamma z}{1+\beta}<i<\frac{\zeta}{2}} \binom{\zeta}{i}\binom{\zeta}{\beta i} g(i,\beta i) \leq z\cdot 2^{-23\gamma z} \leq N^{9/10}\cdot 2^{-23N^{-8/10}}. \tag{89}$$

Then by adding eq. (85) and eq. (89), and seeing that clearly eq. (85) dominates for sufficiently large N, for choice of $\beta = 1/100$,

$$\mathbf{Pr}[B_k \text{ is } \textit{not} \text{ a } \beta\text{-edge expander}] \leq 2^{155} N^{-3.99}. \tag{90}$$

Consider the graph on B_k induced plus the self-loops introduced. This graph is d-regular (including self-loops), then by Cheeger's inequality if it is a β-edge expander then for $\beta = 1/100$ and $d = 100$, it is a α-spectral expander with a second normalized eigenvalue of at most

$$\alpha \leq 1 - \frac{1}{2\cdot 10^8}. \tag{91}$$

We can now perform one more union bound over both the probability that every V_k is of near-optimal size (eq. (81)) and over all $k \in [\ell]$ that B_k is *not* an α-expander. Since $\ell = N^{1/10}$, this overall probability is $O(N^{-3})$ as stated in the statement.

It remains to prove eq. (83). Consider disjoint sets U_1, U_2 contained in B_k where $|U_1| = i$ and $|U_2| = \beta i$. Then, the event E_{U_1,U_2} is equivalent to the event that every edge emanating from U_1 is contained in $U_1 \cup U_2$. Since the graph G from $P_{M,\ell}(F)$ is built by considering a graph G' on M vertices built of d perfect matchings and then considering the induced graph on N vertices under the injective map $\iota : V \to V'$, then this implies that every edge emanating from $\iota(U_1)$ is contained in

$$A \stackrel{\text{def}}{=} (V_k \setminus \iota(B_k)) \cup \iota(U_1) \cup \iota(U_2). \tag{92}$$

the size of A is easily bounded as $|A| \leq [z(1+\gamma) - z(1-\gamma)] + i + \beta i = 2\gamma z + (1+\beta)i$. For each of the d matchings (corresponding to a different color), the probability that all κ-th colored edges emanating $\iota(U_1)$ lie in A is

$$\frac{|A|-1}{\zeta-1}\cdot\frac{|A|-3}{\zeta-3}\cdot\ldots\cdot\frac{|A|-i+1}{\zeta-i+1} \leq \left(\frac{|A|}{\zeta}\right)^{i/2} \leq \left(\frac{|A|}{z}\right)^{i/2} \leq \left(2\gamma + \frac{(1+\beta)i}{z}\right)^{i/2}. \tag{93}$$

Since the d edge colorings are independently sampled, then the net probability is bounded by eq. (83).

□

Lemma 11. *Let m be $\leq N^{1/100}$. Let $\mathcal{D}_1$ be the distribution on pairs (G, F) obtained by sampling $G \sim P_{M,\ell}$, choosing a uniformly random vertex $v \in G$, and then choosing F to be a uniformly random subset of the connected component of G containing v of size m. Let $\mathcal{D}_2$ be the distribution on pairs (G, F) obtained by first choosing F to be a uniformly random subset of V with size m, and then sampling $G \sim P_{M,\ell}(F)$. Then these distribution are close in statistical distance:*

$$\|\mathcal{D}_1 - \mathcal{D}_2\| \leq 3N^{-9/200}. \tag{14}$$

Proof. In order to prove this lemma, it is illustrative to decompose the procedure for sampling the distributions $P_{M,\ell}$ and $P_{M,\ell}(F)$. Let us note that while the procedure is stated sequentially, there are three different independent components. First, the sampling of the graph G' on M vertices is independent of the construction of the injective map ι. Furthermore sampling ι is constructed by independently sampling $k : V \to [\ell]$ and then

defining the injective map $\iota : V \hookrightarrow V'$ by assigning each vertex j to a vertex in $V_{k(j)}$ without replacement. An equivalent sampling algorithm is to sample uniformly random independent injective maps $\pi_1, \ldots, \pi_\ell$ where each $\pi_i : [M/\ell] \hookrightarrow V_i$ is an enumeration of the vertices of V_i. Then, define ι by using the random enumerations $\{\pi_i\}$ to sequentially assign each vertex j a vertex in $V_{k(j)}$. Formally, if j is the s-th vertex such that $k(j) = i$; then $\iota(j) = \pi_i(s)$.

Therefore, the sampling procedure can be rewritten as a deterministic process based on the three samples: the graph G', the map k, and enumerations $\{\pi_i\}$. Notice that the only difference between $P_{M,\ell}$ and $P_{M,\ell}(F)$ is the map k. Notice, we can further simplify by thinking only of the map $k' : V \to \{0, 1\}$ where $k' = 0$ if $k \neq 1$. This is because we can sample a uniformly random function $k'' : V \to [\ell] \setminus \{1\}$ and define

$$k(j) = \begin{cases} 1 & \text{if } k'(j) = 1 \\ k''(j) & \text{if } k'(j) \neq 1. \end{cases} \tag{94}$$

Let us say that $k' : V \to \{0, 1\}$ is a *uniformly random map* if each $k(j)$ for $j \in V$ is an iid random variable with $\mathbf{Pr}[k'(j) = 1] = 1/\ell$. Now, let $\mathcal{D}'_1$ be the distribution on (k', F) defined by sampling a uniformly random map k' and then uniformly randomly a set F of size m from $(k')^{-1}(1)$. Let $\mathcal{D}'_2$ be the distribution on (k, F) formed by sampling a uniformly random set F of m vertices and then a uniformly random map k' such that $k'(F) = 1$. By the previously stated decomposition of the sampling procedures for $\mathcal{D}_1$ and $\mathcal{D}_2$, if we show that $\mathcal{D}'_1$ and $\mathcal{D}'_2$ are statistically indistinguishable, then $\mathcal{D}_1$ and $\mathcal{D}_2$ are at least as indistinguishable.

We first remark that these distributions are not the same as the expected Hamming weight of a vector k' from $\mathcal{D}'_1$ is N/ℓ whereas the expected Hamming weight of the vector from $\mathcal{D}'_2$ is $m + (N - m)/\ell$. To show, however, that these distributions are statistically close, we will use some simple Chernoff bounds and Pinsker's inequality. Let E be the event that the sampled vector k' has expected Hamming weight either $< (1 - \gamma)z$ or $> (1 + \gamma)z$. Let a_1 and a_2 be the probabilities of the event E over distributions $\mathcal{D}'_1$ and $\mathcal{D}'_2$. Since $m \ll z$, a simple Chernoff bound shows that both a_1 and a_2 are at most

$$a_1, a_2 \leq 2\exp\left(-\frac{\gamma^2 N}{12\ell}\right) \leq 2\exp\left(-\frac{N^{7/10}}{12}\right). \tag{95}$$

Let $\mathcal{D}''_1$ and $\mathcal{D}''_2$ be the distributions conditioned on the event $\neg E$, respectively. Therefore, the statistical distances are at most

$$\left\|\mathcal{D}'_1 - \mathcal{D}''_1\right\| \leq 2a_1, \quad \left\|\mathcal{D}'_2 - \mathcal{D}''_2\right\| \leq 2a_2. \tag{96}$$

We now bound the statistical distance between $\mathcal{D}''_1$ and $\mathcal{D}''_2$ using Pinsker's inequality and a bound on the KL divergence. We calculate the probability of outputting (k', F) under both distributions. First,

$$\mathcal{D}''_1(k', F) = \frac{1}{1 - a_1} \cdot \left(\frac{1}{\ell}\right)^{|k'|} \left(1 - \frac{1}{\ell}\right)^{N - |k'|} \cdot \frac{1}{\binom{|k'|}{m}}. \tag{97}$$

The $(1 - a_1)^{-1}$ term is due to the conditioning on event $\neg E$. The next two terms give the probability of sampling k' as each index is set to 1 with iid probability $1/\ell$. The last is the probability of selecting the specific set F given k'. Second,

$$\mathcal{D}''_2(k', F) = \frac{1}{1 - a_2} \cdot \frac{1}{\binom{N}{m}} \cdot \left(\frac{1}{\ell}\right)^{|k'| - m} \left(1 - \frac{1}{\ell}\right)^{(N - m) - (|k'| - m)}. \tag{98}$$

Likewise, the $(1-a_2)^{-1}$ term is due to the conditioning on event $\neg E$. The next term is the probability of sampling F. The last two terms yield the probability of sampling the rest of k'. Then, as a sub-calculation in the KL divergence,

$$\ln\left(\frac{\mathcal{D}_1''(k',F)}{\mathcal{D}_2''(k',F)}\right) = \ln\left(\frac{1-a_2}{1-a_1}\cdot\left(\frac{1}{\ell}\right)^m\cdot\frac{\binom{N}{m}}{\binom{|k'|}{m}}\right) \tag{99a}$$

$$= \ln\left(\frac{1-a_2}{1-a_1}\cdot\frac{N}{|k'|\ell}\cdot\frac{N-1}{(|k'|-1)\ell}\cdot\ldots\cdot\frac{N-m+1}{(|k'|-m+1)\ell}\right) \tag{99b}$$

$$\leq \ln\left(\frac{1-a_2}{1-a_1}\right) + \sum_{i=0}^{m-1}\ln\left(\frac{N-i}{(|k'|-i)\ell}\right) \tag{99c}$$

$$\leq 2a_1 + \sum_{i=0}^{m-1}\ln\left(\frac{N-i}{((1-\gamma)z-i)}\cdot\frac{z}{N}\right) \tag{99d}$$

$$\leq 2a_1 + \sum_{i=0}^{m-1}\ln\left(\frac{z}{(1-\gamma)z-i}\right) \tag{99e}$$

$$\leq 2a_1 + \sum_{i=0}^{m-1}\ln\left(\frac{1}{1-2\gamma}\right) \tag{99f}$$

$$\leq 2a_1 + 4m\gamma \tag{99g}$$

$$\leq 5N^{-9/100}. \tag{99h}$$

Here we used $\ln\left(\frac{1}{1-x}\right) \leq 2x$ twice and eq. (99d) follows from $|k'| \geq (1-\gamma)z$ since we condition on the event $\neg E$ and eq. (99f) follows from $m \ll \gamma z$. To pass from eq. (99g) to eq. (99h), we used the fact that $2\alpha_1 \ll m\gamma \leq N^{-9/100}$. Therefore, the total KL divergence between $\mathcal{D}_1''$ and $\mathcal{D}_2''$ is also bounded by

$$\mathrm{KL}(\mathcal{D}_1'',\mathcal{D}_2'') = \sum_{(k',F)}\mathcal{D}_1''(k',F)\ln\left(\frac{\mathcal{D}_1''(k',F)}{\mathcal{D}_2''(k',F)}\right) \leq 5N^{-9/100}\sum_{(k',F)}\mathcal{D}_1''(k',F) = 5N^{-9/100}. \tag{100}$$

By Pinsker's inequality, then

$$\|\mathcal{D}_1''-\mathcal{D}_2''\| \leq \sqrt{\frac{1}{2}\cdot\mathrm{KL}(\mathcal{D}_1'',\mathcal{D}_2'')} \leq \sqrt{\frac{1}{2}\cdot 5N^{-9/100}} \leq 2N^{-9/200}. \tag{101}$$

As this dominates the bounds in eq. (96), we can conclude by the triangle inequality for $\|\cdot\|$ that for sufficiently large N, $\|\mathcal{D}_1'-\mathcal{D}_2'\| \leq 3N^{-9/200}$. Finally, as remarked previously, this implies that $\|\mathcal{D}_1-\mathcal{D}_2\| \leq 3N^{-9/200}$. □

B Ommited proofs for the adversary method

Every oracle query algorithm using Hilbert space $\mathcal{H}_A$ can be adapted into a "ficticious" oracle query algorithm using Hilbert space $\mathcal{H}_A\otimes\mathcal{H}_I$ where an oracle gate $|i\rangle \mapsto (-1)^{a_i}|i\rangle$ is turned into a gate $|i\rangle|a,b\rangle \mapsto (-1)^{a_i}|i\rangle|a,b\rangle$ where the second register is $\mathcal{H}_I$. this allows us to analyze how the algorithm would behave on a superposition over oracles.

Assume an algorithm $\mathcal{A}$ makes T queries to the oracle. For $1 \leq t \leq T$, let $\rho_t = \mathrm{tr}_{\mathcal{H}_A}(|\psi_t\rangle\langle\psi_t|)$ be the reduced density matrix (for corresponding pure state $|\psi_t\rangle$) of the computation on the register $\mathcal{H}_I$ immediately before the t-th oracle gate. We assume the algorithm from an initial starting state of $|0\rangle_{\mathcal{H}_A}\otimes|\chi_0\rangle_{\mathcal{H}_I}$ evolves to

$$|\psi_0\rangle = |\xi_0\rangle_{\mathcal{H}_A}\otimes|\chi_0\rangle_{\mathcal{H}_I} \tag{102}$$

immediately before the application of the first oracle query. The convenient start, as suggested by Ambainis, is to consider the algorithm run with

$$|\chi_0\rangle = \frac{1}{\sqrt{2|X|}} \sum_{x\in X} |x\rangle + \frac{1}{\sqrt{2|Y|}} \sum_{y\in Y} |y\rangle . \tag{103}$$

In this case[7], for $x \in X$ and $y \in Y$,

$$(\rho_0)_{xy} = \frac{1}{2\sqrt{|X||Y|}}. \tag{104}$$

For any algorithm $\mathcal{A}$ achieving eq. (33), let X_{good} and Y_{good} be the set of x and y, respectively, such that $\mathbf{Pr}_{\mathcal{A}}[\mathcal{A}^x = 0]$ and $\mathbf{Pr}_{\mathcal{A}}[\mathcal{A}^y = 1]$ are $\geq 1 - \epsilon$. By Markov's inequality, $|X_{\text{good}}| \geq (1-\epsilon)|X|$ and $|Y_{\text{good}}| \geq (1-\epsilon)|Y|$. For any $(x, y) \in X_{\text{good}} \times Y_{\text{good}}$, we will show that

$$\left|(\rho_T)_{xy}\right| \leq \sqrt{\frac{\epsilon(1-\epsilon)}{|X||Y|}}. \tag{105}$$

To show eq. (105), let $|\psi_x\rangle$ and $|\psi_y\rangle$ be the final states of $\mathcal{A}$ when run with inputs $|x\rangle$ and $|y\rangle$ in register $\mathcal{H}_I$, respectively. Then, the final state of the algorithm after T queries will be

$$|\psi_{\text{final}}\rangle = \frac{1}{\sqrt{2|X|}} \sum_{x\in X} |\psi_x\rangle |x\rangle + \frac{1}{\sqrt{2|Y|}} \sum_{y\in Y} |\psi_y\rangle |y\rangle . \tag{106}$$

Take a basis of $\mathcal{H}_A$ of the form $|z\rangle |v\rangle$ where $|z\rangle$ corresponds to the answer bit and $|v\rangle$ is a basis for the remainder of the work register. In this basis, let

$$|\psi_x\rangle = \sum_{z,v} \alpha_{z,v} |z\rangle |v\rangle , \qquad |\psi_y\rangle = \sum_{z,v} \beta_{z,v} |z\rangle |v\rangle . \tag{107}$$

Since the algorithm $\mathcal{A}$ cannot effect the amplitudes of the oracle string in $\mathcal{H}_A$, then

$$\begin{aligned}
(\rho_T)_{xy} &= \langle y| \rho_T |x\rangle && \text{(108a)}\\
&= \sum_{z,v} \langle z, v, y|\psi_{\text{final}}\rangle \langle \psi_{\text{final}}|z, v, x\rangle && \text{(108b)}\\
&= \sum_{z,v} \frac{\alpha^\dagger_{z,v}}{\sqrt{2|Y|}} \frac{\beta_{z,v}}{\sqrt{2|X|}} && \text{(108c)}\\
&= \frac{1}{2\sqrt{|X||Y|}} \sum_{z,v} \alpha^\dagger_{z,v}\beta_{z,v}. && \text{(108d)}
\end{aligned}$$

Like Ambainis, define

$$\epsilon_0 \stackrel{\text{def}}{=} \sum_v |\alpha_{0,v}|^2, \qquad \epsilon_1 \stackrel{\text{def}}{=} \sum_v |\beta_{1,v}|^2. \tag{109}$$

[7] Ambainis uses the notation $\rho_{xy} = \langle y| \rho |x\rangle$ to refer to indices of matrices (this matches the standard notation).

Since $x \in X_{\text{good}}$ and $y \in Y_{\text{good}}$, then $\epsilon_0 \leq \epsilon$ and $\epsilon_1 \leq \epsilon$. Then

$$\sum_{z,v} \alpha_{z,v}^\dagger \beta_{z,v} \leq \sum_{z,v} |\alpha_{z,v}||\beta_{z,v}| \tag{110a}$$

$$= \sum_{v} |\alpha_{0,v}||\beta_{0,v}| + |\alpha_{1,v}||\beta_{1,v}| \tag{110b}$$

$$\leq \sqrt{\sum_{v} |\alpha_{0,v}|^2} \sqrt{\sum_{v} |\beta_{0,v}|^2} + \sqrt{\sum_{v} |\alpha_{1,v}|^2} \sqrt{\sum_{v} |\beta_{1,v}|^2} \tag{110c}$$

$$= \sqrt{\epsilon_0(1-\epsilon_1)} + \sqrt{\epsilon_1(1-\epsilon_0)} \tag{110d}$$

$$\leq 2\sqrt{\epsilon(1-\epsilon)}. \tag{110e}$$

where eq. (110c) follows from the Cauchy-Schwarz inequality. Combining eq. (108d) and eq. (110e) gives eq. (105).

Theorem 13. *Let $f : \{0,1\}^{N+M} \to \{0,1\}$ be a function and let $X, Y \subset \{0,1\}^N \times \{0,1\}^M$ be two subsets such that $X \subset f^{-1}(0)$ and $Y \subset f^{-1}(1)$. Let $R \subset X \times Y$ be a relation such that*

1. *For every $x \in X$, let $R_x \subset Y$ equal $R_x = \{y : (x,y) \in R\}$ such that $\underline{m} \leq |R_x| \leq \overline{m}$.*
2. *For every $y \in Y$, let $R_y \subset X$ equal $R_y = \{x : (x,y) \in R\}$ such that $\underline{m'} \leq |R_x| \leq \overline{m'}$.*
3. *For every $x = (a,b) \in X$ and $i \in [N]$, let $\ell_{x,i}$ be the number of $y = (c,d) \in Y$ such that $(x,y) \in R$ and $a_i \neq c_i$. Likewise, for every $y = (c,d) \in Y$ and $i \in [N]$, let $\ell_{y,i}$ be the number of $x = (a,b) \in X$ such that $(x,y) \in R$ and $a_i \neq c_i$. Let $\ell_{\max}$ be the maximum product $\ell_{x,i}\ell_{y,i}$ over $(x,y) \in R$ and $i \in [N]$ such that $a_i \neq c_i$.*

Then any quantum algorithm $\mathcal{A}$ which only queries the first N bits of the oracle and computes f such that

$$\mathop{\mathbf{E}}_{x=(a,b)\in X} \mathop{\mathbf{Pr}}_{\mathcal{A}} [\mathcal{A}^a \neq 0] \leq \epsilon^2 \quad \text{and} \quad \mathop{\mathbf{E}}_{y=(c,d)\in Y} \mathop{\mathbf{Pr}}_{\mathcal{A}} [\mathcal{A}^c \neq 1] \leq \epsilon^2 \tag{33}$$

uses

$$\geq \left(1 - 2\sqrt{\epsilon(1-\epsilon)}\right) \sqrt{\frac{(\underline{m} - 2\epsilon\overline{m})\left(\underline{m'} - 2\epsilon\overline{m'}\right)}{\ell_{\max}}} \quad \text{queries.} \tag{34}$$

Proof. Let X_{good} and Y_{good} be the sets as previously defined and let

$$R_{\text{good}} \stackrel{\text{def}}{=} R \cap X_{\text{good}} \times Y_{\text{good}}. \tag{111}$$

We can bound the size of R_{good} as follows. First note that

$$|R| \geq \max\left(\underline{m}|X|, \underline{m'}|Y|\right) \tag{112a}$$

$$\geq \frac{\underline{m}|X| + \underline{m'}|Y|}{2}. \tag{112b}$$

Second, by the upper bound on the degree,

$$|R \setminus R_{\text{good}}| \leq \overline{m}|X \setminus X_{\text{good}}| + \overline{m'}|Y \setminus Y_{\text{good}}| \tag{113a}$$

$$\leq \epsilon\overline{m}|X| + \epsilon\overline{m'}|Y|. \tag{113b}$$

Therefore,

$$|R_{\text{good}}| \geq \frac{1}{2}\Big((\underline{m} - 2\epsilon\overline{m})|X| + \Big(\underline{m'} - 2\epsilon\overline{m'}\Big)|Y|\Big) \tag{114a}$$

$$\geq \sqrt{(\underline{m} - 2\epsilon\overline{m})\Big(\underline{m'} - 2\epsilon\overline{m'}\Big)} \cdot \sqrt{|X||Y|}. \tag{114b}$$

Now define,

$$S_t \stackrel{\text{def}}{=} \sum_{(x,y)\in R_{\text{good}}} |(\rho_t)_{xy}|. \tag{115}$$

Notice that then

$$S_0 - S_T = \sum_{(x,y)\in R_{\text{good}}} |(\rho_0)_{xy}| - |(\rho_T)_{xy}| \tag{116a}$$

$$\geq |R_{\text{good}}| \cdot \left(\frac{1 - 2\sqrt{\epsilon(1-\epsilon)}}{2\sqrt{|X||Y|}}\right) \tag{116b}$$

$$\geq \sqrt{(\underline{m} - 2\epsilon\overline{m})\Big(\underline{m'} - 2\epsilon\overline{m'}\Big)} \cdot \frac{1 - 2\sqrt{\epsilon(1-\epsilon)}}{2}. \tag{116c}$$

We will next show that the difference $S_t - S_{t+1}$ is minimal:

$$S_t - S_{t+1} \leq \frac{\sqrt{\ell_{\max}}}{2}. \tag{117}$$

Therefore,

$$S_0 - S_T \leq \frac{T\sqrt{\ell_{\max}}}{2} \tag{118}$$

which yields a lower bound of

$$T \geq \left(1 - 2\sqrt{\epsilon(1-\epsilon)}\right)\sqrt{\frac{(\underline{m} - 2\epsilon\overline{m})\Big(\underline{m'} - 2\epsilon\overline{m'}\Big)}{\ell_{\max}}}. \tag{119}$$

To show eq. (117), we first write the intermediate state $|\psi_t\rangle$ as

$$|\psi_t\rangle = \sum_{i,r,z} \alpha_{i,r,x,z} \, |i, r, z\rangle \otimes |x\rangle \tag{120}$$

where $i \in [N]$ is the index of the input variable being queried, r is the answer bit for the query, z are the bits not involved in the query or answer, and x is the oracle on $N + M$ bits being queried. Immediately after querying the oracle, the state of the algorithm is precisely

$$\sum_{i,r,z} \alpha_{i,r,x,z} \, |i, r \oplus x_i, z\rangle \otimes |x\rangle \, . \tag{121}$$

Like Ambainis, if we denote

$$|\xi_{i,r,z}\rangle = \sum_x \alpha_{i,r,x,z} \, |x\rangle \, , \tag{122a}$$

$$\left|\xi'_{i,r,z}\right\rangle = \sum_x \alpha_{i,r\oplus x_i,x,z} \, |x\rangle \, , \tag{122b}$$

$$\rho_{t,i} = \sum_{r,z} |\xi_{i,r,z}\rangle\langle\xi_{i,r,z}| \, , \tag{122c}$$

$$\text{and } \rho_{t+1,i} = \sum_{r,z} \left|\xi'_{i,r,z}\right\rangle\left\langle\xi'_{i,r,z}\right| , \tag{122d}$$

then, notice $\rho_{t,i}$ and $\rho_{t+1,i}$ are the parts of ρ_t and ρ_{t+1} corresponding to querying index i, with each being the sum over $i \in [N]$ of the corresponding parts. Notice that

$$|(\rho_t)_{xy}| - |(\rho_{t+1})_{xy}| \leq |(\rho_t)_{xy} - (\rho_{t+1})_{xy}| \tag{123a}$$

$$= \left|\sum_i (\rho_{t,i})_{xy} - (\rho_{t+1,i})_{xy}\right| \tag{123b}$$

$$\leq \sum_i |(\rho_{t,i})_{xy} - (\rho_{t+1,i})_{xy}|. \tag{123c}$$

Notice that if $x_i = y_i$, then

$$(\rho_{t,i})_{xy} = \sum_{r,z} \alpha^\dagger_{i,r,x,z}\alpha_{i,r,y,z} = \sum_{r,z} \alpha^\dagger_{i,r\oplus x_i,x,z}\alpha_{i,r\oplus y_i,y,z} = (\rho_{t+1,i})_{xy}. \tag{124}$$

And when $x_i \neq y_i$, then

$$(\rho_{t+1,i})_{xy} = -(\rho_{t,i})_{xy} \implies \left|(\rho_{t,i})_{xy} - (\rho_{t+1,i})_{xy}\right| = 2\left|(\rho_{t,i})_{xy}\right|. \tag{125}$$

Therefore, the only indices we need to consider in eq. (123c) are $i \in [N]$ for which $x_i \neq y_i$. Let $S_{t,i}$ equal

$$S_{t,i} \stackrel{\text{def}}{=} \sum_{(x,y)\in R_{\text{good}} \text{ s.t. } x_i \neq y_i} |(\rho_{t,i})_{xy}|. \tag{126}$$

Then,

$$S_t - S_{t+1} = \sum_{(x,y)\in R_{\text{good}}} |(\rho_t)_{xy}| - |(\rho_{t+1})_{xy}| \leq 2\sum_{i\in[N]} S_{t,i}. \tag{127}$$

To bound $S_{t,i}$, since $\rho_{t,i}$ is non-negative by construction,

$$0 \leq \left(\langle y| \sqrt{\ell_{x,i}} - \langle x| \sqrt{\ell_{y,i}}\right)\rho_{t,i}\left(\sqrt{\ell_{x,i}}\,|y\rangle - \sqrt{\ell_{y,i}}\,|x\rangle\right) \tag{128a}$$

$$= \ell_{x,i}(\rho_{t,i})_{yy} + \ell_{x,i}(\rho_{t,i})_{xx} - \sqrt{\ell_{x,i}\ell_{y,i}}\left((\rho_{t,i})_{xy} + (\rho_{t,i})_{yx}\right) \tag{128b}$$

and therefore

$$\left|(\rho_{t,i})_{xy}\right| \leq \frac{1}{2}\left(\sqrt{\frac{\ell_{y,i}}{\ell_{x,i}}}|(\rho_{t,i})_{xx}| + \sqrt{\frac{\ell_{x,i}}{\ell_{y,i}}}\left|(\rho_{t,i})_{yy}\right|\right). \tag{129}$$

So,

$$S_{t,i} \leq \sum_{(x,y)\in R_{\text{good}} \text{ s.t. } x_i \neq y_i} \frac{1}{2}\left(\sqrt{\frac{\ell_{y,i}}{\ell_{x,i}}}|(\rho_{t,i})_{xx}| + \sqrt{\frac{\ell_{x,i}}{\ell_{y,i}}}\left|(\rho_{t,i})_{yy}\right|\right) \tag{130a}$$

$$\leq \frac{1}{2}\left(\sum_{x\in X_{\text{good}}} \sqrt{|\ell_{x,i}\ell_{y,i}|}|(\rho_{t,i})_{xx}| + \sum_{y\in Y_{\text{good}}} \sqrt{|\ell_{x,i}\ell_{y,i}|}\left|(\rho_{t,i})_{yy}\right|\right) \tag{130b}$$

$$\leq \frac{\sqrt{\ell_{\max}}}{2}\operatorname{tr}(\rho_{t,i}). \tag{130c}$$

where eq. (130c) follows from positivity and the definition of trace. Therefore,

$$S_t - S_{t+1} = \sum_{i\in[N]} S_{t,i} \leq \frac{\sqrt{\ell_{\max}}}{2}\sum_{i\in[N]} \operatorname{tr}(\rho_{t,i}) = \frac{\sqrt{\ell_{\max}}}{2}\operatorname{tr}(\rho_t) = \frac{\sqrt{\ell_{\max}}}{2}. \tag{131}$$

□

Corollary 14. *Let X and Y be two subsets of $\{0,1\}^{N+M}$ satisfying the three conditions listed in Theorem 13. Then, any query algorithm $(1-\delta)$-distinguishing the uniform distributions over X and Y, must use eq. (34) queries for $\epsilon = 2\delta$.*

Proof. Suppose there exists an algorithm $\mathcal{A}$ such that

$$\mathop{\mathbf{E}}_{y\in Y} \mathop{\mathbf{Pr}}_{\mathcal{A}}[\mathcal{A}^y = 1] - \mathop{\mathbf{E}}_{x\in X} \mathop{\mathbf{Pr}}_{\mathcal{A}}[\mathcal{A}^x = 0] \geq 1-\delta. \tag{132}$$

Then,

$$\mathop{\mathbf{E}}_{y\in Y} \mathop{\mathbf{Pr}}_{\mathcal{A}}[\mathcal{A}^y = 1] \geq 1-2\delta \quad \text{and} \quad \mathop{\mathbf{E}}_{x\in X} \mathop{\mathbf{Pr}}_{\mathcal{A}}[\mathcal{A}^x = 0] \leq 2\delta. \tag{133}$$

Define $X' = X \times \{0\}$ and $Y' = Y \times \{1\}$. Then, these sets in $\{0,1\}^{N+M+1}$ are disjoint and satisfy eq. (33) for $\epsilon = 2\delta$ and we can apply Theorem 13. □