Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 176 (2025)

Separating QMA from QCMA with a classical oracle

John Bostanci[1], Jonas Haferkamp[2,3], Chinmay Nirkhe[4], and Mark Zhandry[5,6]

[1]Columbia University, *New York, NY, USA*
[2]Saarland University, *Saarbrücken, Germany*
[3]Harvard University, *Cambridge, Mass., USA*
[4]University of Washington, *Seattle, Wash., USA*
[5]Stanford University, *Stanford, Calif., USA*
[6]NTT Research, Inc, *Sunnyvale, Calif., USA*

Abstract

We construct a classical oracle proving that, in a relativized setting, the set of languages decidable by an efficient quantum verifier with a quantum witness (QMA) is strictly bigger than those decidable with access only to a classical witness (QCMA). The separating classical oracle we construct is for a decision problem we coin *spectral Forrelation* – the oracle describes two subsets of the boolean hypercube, and the computational task is to decide if there exists a quantum state whose standard basis measurement distribution is well supported on one subset while its Fourier basis measurement distribution is well supported on the other subset. This is equivalent to estimating the spectral norm of a "Forrelation" matrix between two sets that are accessible through membership queries.

Our lower bound derives from a simple observation that a query algorithm with a classical witness can be run multiple times to generate many samples from a distribution, while a quantum witness is a "use once" object. This observation allows us to reduce proving a QCMA lower bound to proving a sampling hardness result which does not simultaneously prove a QMA lower bound. To prove said sampling hardness result for QCMA, we observe that quantum access to the oracle can be compressed by expressing the problem in terms of bosons – a novel "second quantization" perspective on compressed oracle techniques, which may be of independent interest. Using this compressed perspective on the sampling problem, we prove the sampling hardness result, completing the proof.

ISSN 1433-8092

Contents

Part I
Introduction

The problem of finding a standard oracle separation between QMA (the class of problems that can be verified with a quantum computer and quantum witness) and QCMA (the class of problems that can be verified with a quantum computer and *classical* witness) is a central open problem in the field of quantum query complexity and is the first question mentioned in Aaronson's list of open query complexity problems [Aar21]. The goal of separating QMA from QCMA is, in some sense, to understand the following question:

Do quantum witnesses offer more power than classical witnesses?

This question was first posed by Aharonov and Naveh [AN02], and partially answered by the work of Aaronson and Kuperberg [AK07], which provided a *quantum* oracle separation between the two classes. 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 significantly stronger new tools.

However, the quantum oracle separation of [AK07] could be considered an unsatisfying oracle separation between QMA and QCMA, as it avoids answering deeper questions about the power of quantum witnesses over classical ones. The separation constructs a unitary property testing problem for which a verifier must exactly know a Haar random state to solve the problem, essentially forcing that any classical witness for the problem must provide a full classical description of the Haar random state[1]. The question of whether QMA equals QCMA necessitates more than a lower bound on the classical description complexity of a random quantum state; it requires proving that even the single bit identified by the QMA decision problem cannot be verified by a classical witness. It could be that if QMA does equal QCMA, the QCMA verifier for a QMA-complete problem would not receive a "verbatim" description of the quantum witness (such as a circuit preparing the quantum witness), but rather some other classical information derived (potentially inefficiently) from the instance or the original quantum witness, which can be used to answer the decision problem but not necessarily to reproduce the quantum witness.

By forcing the oracle to be classical, one hopes to identify more meaningful reasons why useful properties – beyond the full description of a quantum witness – should be hard to write down in a classical witness. A final reason for studying the problem of a classical oracle separation is that the question lies in the rich field of quantum query complexity and has been linked to open questions in quantum cryptography, such as the existence of quantum money [Lut11, NZ24] and pseudorandomness against quantum adversaries [LMY25]. One hopes that finding a classical oracle separation between QMA and QCMA will provide an improved understanding of the complexity theory underlying these cryptographic questions.

[1] The notion of "knowing" the state is meant to be more intuitive rather than formal due to the result of [INN+21], where it is shown that having an oracle that solves the separating problem of [AK07] does not allow one to synthesize the state specified by the oracle.

1 Proof overview

Theorem 1.1 (Classical oracle separation between QMA and QCMA)**.** *There exists a pair of oracles* $S, U : \{0,1\}^* \to \{0,1\}$ *and a language* $\mathcal{L}^{S,U}$ *such that* $\mathcal{L}^{S,U} \in \mathsf{QMA}^{S,U} \setminus \mathsf{QCMA}^{S,U}$ *– i.e., there exists a polynomial-time quantum verifier taking as input quantum witnesses as input for deciding membership in* $\mathcal{L}^{S,U}$*, while no polynomial-time quantum verifier taking as input classical witnesses can decide membership in* $\mathcal{L}^{S,U}$*.*

This is the main result of this work. While we express the statement in terms of two oracles, as our construction is most natural to describe in terms of a quantum algorithm comparing two oracles, it is easy enough to convert the statement into that of a single oracle.

1.1 The challenges of an oracle separation

To prove this classical oracle separation between QMA and QCMA, it suffices to construct a property testing problem about classical oracles that can be decided with quantum witnesses but cannot be decided with classical witnesses. The transformation from the property testing problem to an oracle separation between complexity classes is a diagonalization argument that takes some work but is fairly standard.

The central challenge in separating QMA from QCMA is illustrated in the following thought experiment: Consider measuring a quantum witness $|\psi\rangle$ for a QMA problem to generate a classical witness. If the resulting classical object is accepted by the QMA verifier, then the problem must have been in QCMA. Therefore, for a problem to be in $\mathsf{QMA} \setminus \mathsf{QCMA}$, it follows that for any quantum witness $|\psi\rangle$ and any computational basis $|x\rangle$, $|\langle x|\psi\rangle| \leq \mathrm{negl}(n)$, as otherwise x would serve as a good classical witness for the problem[2]. Here, the phrase $\mathrm{negl}(n)$ is used to refer to any function that is smaller than every function $1/\mathrm{poly}(n)$. Therefore, assuming $\mathsf{QMA} \neq \mathsf{QCMA}$, the QMA verification procedure must certify (at least implicitly) that the quantum witness is a superposition of a super-polynomial number of basis vectors.

Most techniques for verifying such a complex superposition require highly structured oracles, which presents an inherent challenge for proving a QCMA lower bound. This is because most techniques we currently know of for proving lower bounds against quantum query algorithms are not amenable to highly structured oracles. Additionally, a classical witness would naturally be treated as a form of advice about the oracle, and most quantum query complexity techniques are not very good at distinguishing between quantum advice and classical advice. If a typical technique succeeded in proving that a language is outside of QCMA, but it can not distinguish between quantum and classical advice, it would likely show that the language is outside of QMA as well, failing to give a separation.

Our starting point is the following simple observation first made by [NZ24] and later by Zhandry [Zha25], which separates the functionality of quantum and classical witnesses: if a highly-entangled superposition of computational basis states $|\phi\rangle$ is efficiently generated from a classical witness w, then we can generate a polynomial number of measurement samples from $|\phi\rangle$, whereas if $|\phi\rangle$ is generated from a quantum witness $|\psi\rangle$, then we expect that only one measurement sample can be extracted. This is because the classical witness can be efficiently copied while the quantum witness cannot be. Somewhat paradoxically, this demonstrates a particular way in which a classical witness is *more* powerful than a quantum witness. It

[2]This can be extended to show that the measurement distribution for any quantum witness for the problem in *every computationally efficient basis* cannot have more than $\mathrm{negl}(n)$ support on any basis vector.

is precisely this boost in power that we show is too good to be true, thereby proving the impossibility of an efficient verification of a classical witness. This observation, however, does not yet indicate how to actually design the oracles used in the separation.

1.2 Spectral Forrelation

Zhandry [Zha25] gives a candidate oracle for a separation and an initial analysis, but ultimately was unable to prove the separation. Our oracles are inspired by Zhandry's, though our approach to analyzing them is very different. Zhandry [Zha25] defines a variant of the problem that we will call *spectral Forrelation*[3], which is plausibly in $\mathsf{QMA} \setminus \mathsf{QCMA}$: we say a pair of subsets $S, U \subseteq \{0,1\}^n$ are α-spectrally Forrelated if

$$\alpha = \left\| \Pi_U \cdot H^{\otimes n} \cdot \Pi_S \right\|_{\text{op}}^2 = \max_{\||\psi\rangle\|=1} \left\| \Pi_U \cdot H^{\otimes n} \cdot \Pi_S |\psi\rangle \right\|^2 \tag{1a}$$

$$\text{where } \Pi_U = \sum_{x \in U} |x\rangle\langle x|, \ \Pi_S = \sum_{x \in S} |x\rangle\langle x| . \tag{1b}$$

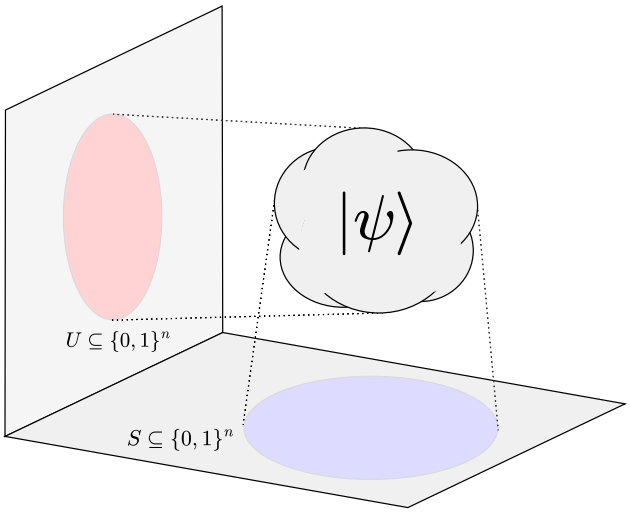


Figure 1: *A cartoon of spectral Forrelation. The subsets S and U occupy support in the standard/position and Hadamard/momentum bases, respectively. The sets S and U are $\geq \alpha$-spectrally Forrelated if there exists a state $|\psi\rangle$ such that $\|\Pi_U \cdot H^{\otimes n} \cdot \Pi_S |\psi\rangle\|^2 \geq \alpha$. Equivalently, there exists a state $|\psi\rangle$ for which the induced classical distributions by measuring in the standard and Hadamard bases are well supported on S and U, respectively.*

Spectral Forrelation can also be expressed as a property of functions $S, U : \{0,1\}^n \to \{0,1\}$, where the set identified with a function is the pre-image of 1. A decision version of the spectral Forrelation problem

[3]The name comes from the notion of Forrelation, defined by Aaronson [Aar10]. He defined the "Forrelation" between two functions $f, g : \{0,1\}^n \to \{\pm 1\}$ as the quantity $\langle +|^{\otimes n} \operatorname{diag}(f) \cdot H^{\otimes n} \cdot \operatorname{diag}(g) |+\rangle^{\otimes n}$. When f, g are described by oracles, Aaronson [Aar10, AA15] gives a simple BQP algorithm for deciding if the absolute value of the Forrelations is $\geq 3/5$ or $\leq 1/100$: simply prepare $|+\rangle^{\otimes n}$, apply the reflection $\operatorname{diag}(g)$, perform the Hadamard transform $H^{\otimes n}$, and apply the reflection $\operatorname{diag}(f)$, and measure to see if the output is $|+\rangle^{\otimes n}$. On the other hand, Aaronson shows that any algorithm making randomized classical queries to S and U needs an exponential number of queries to decide Forrelation.

[3]Note that Zhandry [Zha25] actually used different notation and used the QFT mod $N = 2^n$ in place of the Hadamard transform, but the underlying idea of two projectors in anti-commuting bases remains the same.

can be defined for two parameters $\alpha > \beta$. The task is to decide if the pair (S, U) are at least $\geq \alpha$-spectrally Forrelated (yes instance) or at most $\leq \beta$-spectrally Forrelated (no instance), promised that one of the cases holds. For $\alpha - \beta \geq 1/\text{poly}(n)$, yes instances of spectral Forrelation can be verified with a n-qubit quantum witness, namely the top singular vector $|\psi\rangle$ of $\Pi_U \cdot H^{\otimes n} \cdot \Pi_S$ (similar to plain Forrelation [Aar10], but using $|\psi\rangle$ instead of $|+\rangle^{\otimes n}$). On the other hand, it can be shown that without $|\psi\rangle$, Spectral Forrelation is hard even for quantum query algorithms. This puts Spectral Forrelation in $\mathsf{QMA} \setminus \mathsf{BQP}$. The question remains: is spectral Forrelation outside QCMA?

1.3 QCMA algorithms imply good one oracle samplers

Let us assume, for contradiction, that there exists a QCMA query algorithm $\mathcal{A}$ for spectral Forrelation requiring a $q = q(n)$ sized classical witness and making $t = t(n)$ oracle queries, where $q(n), t(n)$ are polynomials. A simple transformation can convert $\mathcal{A}$ into an algorithm which separately makes $\leq t$ queries to S and $\leq t$ queries to U.

To prove the impossibility of a QCMA algorithm, we restrict our attention to a subset of all spectrally Forrelated pairs (i.e., yes instances): First, we require pairs (S, U) with the property that the oracle S is *sparse*, consisting of approximately ℓ non-zero entries where $\ell = 2^{cn}$ for some small constant c. Secondly, we require that (Δ, U) is a no instance for all subsets $\Delta \subseteq S$ with $|\Delta| \ll \ell$. We describe pairs (S, U) satisfying this second property as being *strong*, and we will sketch in Subsection §1.4 a technique for generating a distribution over sparse and strong yes instances.

Let $w = w(S, U)$, be a witness certifying that (S, U) is a yes instance. In other words, $\mathcal{A}^{(S,U)}(w)$ accepts with high probability[4]. Since we take (S, U) to be a strong yes instance and thus have that for any small subset Δ, (Δ, U) is a no instance of spectral Forrelation, it follows that $\mathcal{A}^{(\Delta,U)}(w)$ accepts with low probability.

Let us express the algorithm $\mathcal{A}^{(S,U)}(w)$ as a sequence of unitaries interspersed with queries to the S oracle. Assuming that the queries to the U oracle are included in the unitaries $\{V_j\}$, the state of the algorithm $\mathcal{A}$ immediately before its final measurement is given by:

$$\text{pre-measurement state of } \mathcal{A}^{(S,U)}(w) = V_t \mathcal{O}_S V_{t-1} \mathcal{O}_S \dots \mathcal{O}_S V_0 \, |w, 0\rangle \, . \tag{2}$$

Where we take $\mathcal{O}_S$ to be the phase oracle for the function that checks membership in S. Then the state $V_j V_{j-1} \dots V_0 \, |w, 0\rangle$ is equivalent to running the algorithm $\mathcal{A}^{(\emptyset,U)}(w)$ until immediately prior to the j-th oracle query. Since (S, U) is a yes instance, $(\emptyset, U)$ is a no instance, and $\mathcal{A}$ must distinguish the two cases, it must be that $\mathcal{A}$ is actually putting significant query weight on points where S and $\emptyset$ differ, namely points in S. As a consequence, it follows from a hybrid argument similar to that of Bennett *et. al.* [BBBV97], that picking a uniformly random index $j \in [t]$ and measuring the query register of $V_j V_{j-1} \dots V_0 \, |w, 0\rangle$ will yield a sample x_1 from S with probability $\Omega(t^{-2})$. In other words, there exists an algorithm which only queries U and takes a witness w, but produces a sample x_1 from S with noticeable probability. This is not surprising, as information about x_1 could be encoded in w.

[4]Note that the perfect completeness of QCMA algorithms with access to classical oracles was proven by Jordan *et. al.* [JKNN12] but we do not need that result here.

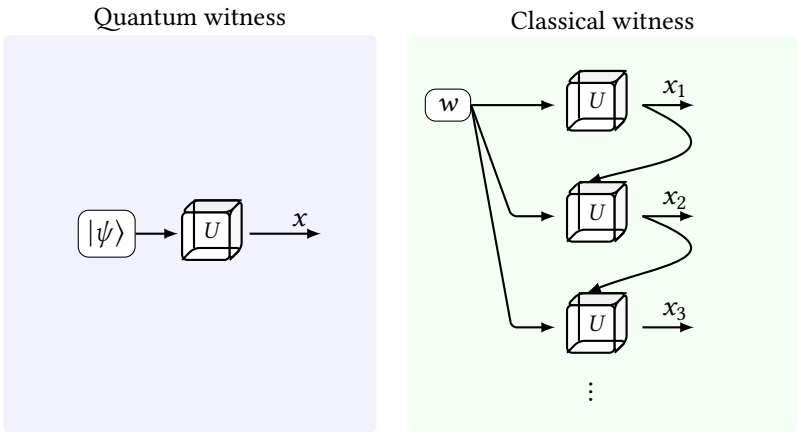


Figure 2: *Adaptive sampling from a classical witness. A quantum witness $|\psi\rangle$ yields a single sample x upon measurement with an algorithm accessing only U. Whereas, a classical witness w can be reused across successive sampling rounds with an adaptive sampler accessing only U and prior samples to generate multiple* distinct *samples.*

A more surprising fact is that a small modification of this sampler can be used to generate many unique samples[5]. Let us condition on the event that $x_1 \in S$. Now consider picking a uniformly random index $j \in [t]$ but this time measuring the query register of $V_j \mathcal{O}_\Delta V_{j-1} \mathcal{O}_\Delta \dots \mathcal{O}_\Delta V_0 |w, 0\rangle$ where $\mathcal{O}_\Delta = \mathrm{id} - 2|x_1\rangle\langle x_1|$. This is equivalent to running the algorithm $\mathcal{A}^{\{x_1\},U}(w)$ until immediately prior to the j-th oracle query. Since $(\{x_1\}, U)$ is also a no instance, the hybrid argument between (S, U) and $(\{x_1\}, U)$ yields that the generated sample x_2 satisfies

$$\mathbb{P}\left[x_2 \in S \setminus \{x_1\} \,|\, x_1 \in S\right] \geq \Omega(t^{-2}). \tag{3}$$

In other words, conditioned on $x_1 \in S$, this procedure generates a unique second point $x_2 \in S$. In general, we can repeat this process v times for $v \ll \ell$, each time generating a novel sample from S conditioned on the previous samples being from S. Therefore, there exists a sampler that only queries U and produces v unique samples from S with probability $\geq (\Omega(t^{-2}))^v$ when provided the correct witness w. See Figure 2 for a cartoon illustration of this iterated sampler. For $v \geq \Omega(q/n)$, where q is the length of the witness, this yields unexpected consequences, as it means that the sampling algorithm cannot simply read the names of samples from the witness; it must also continue to find more samples from its access to U. We can make this explicit by constructing a new sampler which guesses the witness w initially at a 2^{-q} cost in success probability, yielding the following theorem.

Theorem 1.2 (Good samplers from QCMA algorithms (informal))**.** *Assume there exists a classical witness query algorithm $\mathcal{A}$ for spectral Forrelation requiring a $q = q(n)$ sized classical witness and making $t = t(n)$ oracle queries. Let (S, U) be a* strong *yes instance of spectral Forrelation. For all $v = v(n)$ polynomial in n, there exists a query algorithm* CumulativeSampler *such that* $\mathsf{CumulativeSampler}^U$ *makes no queries to S, vt queries to U, and produces v unique samples from S with probability at least $\geq 2^{-q} \cdot \Omega(t)^{-2v}$.*

[5]Note that, as written, there is no guarantee that repeating this experiment with the empty oracle would generate a different sample.

Our primary goal then is to show that the sampler promised by Theorem 1.2 is too good to be true, implying that the assumed classical witness query algorithm cannot exist. A first intuition as to why this sampler should not exist is to consider a query algorithm $\widetilde{\mathsf{CumulativeSampler}}$ that only makes vt queries to S and produces v unique samples from S. It would be natural to expect that access directly to S (as opposed to some spectrally Forrelated U) should only be more useful for outputting samples from S. Yet, by the result of Hamoudi and Magniez [HM23], we know an upper bound on the success probability of $\widetilde{\mathsf{CumulativeSampler}}$ is at most $O(t^2\ell/2^n)^v$. If we compare this upper bound to the supposed CumulativeSampler guaranteed by Theorem 1.2: for t, ℓ small relative to 2^n and $v \geq \Omega(q/n)$, we have $O(t^2\ell/2^n)^v \ll 2^{-q} \cdot \Omega(t)^{-2v}$. Therefore, if the sampler CumulativeSampler from Theorem 1.2 were to actually exist, it would mean that quantum access to any set U spectrally Forrelated with S yields a significantly better sampler than quantum access to the set S itself! This is the first indication that we should be able to prove a QMA versus QCMA oracle separation using this insight about samplers. Of course, this is just intuition; we will actually need to prove that query access to U does not help too much in producing points in S.

Let us emphasize that we cannot derive an analogous theorem from a quantum witness query algorithm. While we can run the query algorithm with a random witness by using $\mathrm{id}/2^q$ as a proxy for the quantum witness, we do not know how to construct a sampling algorithm that continuously produces samples. The CumulativeSampler in Theorem 1.2 relies on being able to copy the *same* witness from one iteration to the next, which is impossible for quantum witnesses. Therefore, proving that the result of Theorem 1.2 is too good to be true does not yield a lower bound for quantum witness query algorithms. This is the fundamental reason that our QCMA-lower bounding technique can separate it from QMA.

Remark 1.3. *Note that Zhandry [Zha25] utilized a similar approach of turning a* QCMA *verifier into a sampler, except instead of removing oracle queries to S, his approach is to remove oracle queries to U while keeping queries to S. He removes queries to U under a general conjecture about the quantum indistinguishability of two oracle distributions whose k-wise marginals are close in relative error. Unfortunately, this general conjecture turns out to be false: (plain) Forrelation gives oracles whose k-wise marginals are close in relative error to random oracles (as proved by Aaronson [AA15]), but the Forrelation algorithm gives a distinguisher[6]. This refutes the general conjecture of Zhandry, but not necessarily the concrete application to removing the oracle U. Nevertheless, we opt for a different approach which sidesteps this issue by removing queries to S before analyzing U.*

1.4 A family of strong yes instances

One method of proving that Theorem 1.2 is too good to be true is to construct a distribution over strong yes instances and to argue that no sampling algorithm could succeed with the probability guaranteed by Theorem 1.2 with respect to this distribution. The introduction of strong yes instances is fundamental for two reasons. The first is type-checking; in order to argue about the behavior of the sampler, we need both yes and no instances of the original problem. By considering strong yes instances, we are implicitly considering no instances. Second, being a strong yes instance is a bare-minimum requirement for separating QMA from QCMA, as yes instances that are not strong have short classical witnesses: suppose

[6]We thank Uma Girish for pointing this out to us in the initial phases of this work.

(Δ, U) is a yes-instance for some $\Delta \subset S$ such that $|\Delta| = \text{poly}(n)$, and consider the top eigenvector $|\Psi_\Delta\rangle$ of $\Pi_\Delta \cdot H^{\otimes n} \cdot \Pi_U \cdot H^{\otimes n} \cdot \Pi_\Delta$. Since $\Delta \subset S$, we also have that $\langle\Psi_\Delta| \Pi_S \cdot H^{\otimes n} \cdot \Pi_U \cdot H^{\otimes n} \cdot \Pi_S |\Psi_\Delta\rangle$ is large. In particular, this means that $|\Psi_\Delta\rangle$ is also a good witness for the instance (S, U). However, $|\Psi_\Delta\rangle$ can be classically described using $O(|\Delta| \cdot \text{poly}(n))$ bits (since its support is limited to Δ, and therefore there are only $|\Delta|$ many amplitudes required to describe it). Therefore, any hope of proving a separation between QMA and QCMA relies on studying the behavior of strong yes instances.

We now describe how to sample from a distribution over strong yes instances. Recall that we pick $\ell = 2^{cn}$ for some small constant c. We first construct a (multi)set $S = \{s_1, \ldots, s_\ell\}$ by uniformly randomly sampling ℓ elements of $\{0,1\}^n$ with replacement. Observe that with all but $1 - \Omega(\ell^2/2^n)$ probability, the elements of S will be distinct. We then construct a distribution over sets U which, with high probability, is spectrally Forrelated with S, with $|S\rangle$ – the uniform superposition over S – serving as the witness state. Concretely, we first compute the terms for $y \in \{0,1\}^n$,

$$\gamma_y^{(S)} \stackrel{\text{def}}{=} \frac{1}{\ell} \sum_{i,j \in [\ell]} (-1)^{y \cdot (s_i \oplus s_j)}. \tag{4}$$

Observe that $\gamma_y^{(S)}$ equals 2^n times the square of the amplitude $H^{\otimes n} |S\rangle$ places on y. We construct the set U by adding each $y \in \{0,1\}^n$ to U with independent probability $1 - \frac{1}{2}e^{-\kappa \gamma_y^{(S)}}$. Here κ is a small constant, say $1/10$.

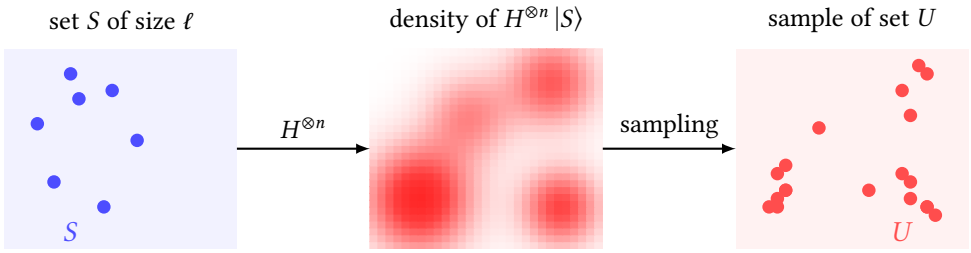


Figure 3: *A depiction of how the pair (S, U) is sampled: (a) First, the multiset S is first sampled uniformly randomly; (b) second, for each $y \in \{0,1\}^n$, a parameter $\gamma_y^{(S)}$ is calculated; (c) third, a set U is sampled by including y with independent probability $1 - \frac{1}{2}e^{-\kappa \gamma_y^{(S)}}$.*

Remark 1.4. *Zhandry [Zha25] constructs (S, U) in a similar manner, but used a Haar-random state on the support of $|S\rangle$ as the witness state, and used a slightly different probability distribution for the sampling of U.*

We prove that a random pair (S, U) sampled via this distribution, which we call Strong, is a strong yes instance with overwhelming probability. The proof of this construction is a somewhat involved medley of concentration inequalities and polynomial approximations. An intuition for why this procedure yields strong yes instances with high probability is that we expect the matrix $M = \Pi_S \cdot H^{\otimes n} \cdot \Pi_U \cdot H^{\otimes n} \cdot \Pi_S$ to concentrate tightly around $a |S\rangle\langle S| + b \cdot \text{id}_S$ for constants a and b, where id_S is the identity matrix restricted to the coordinates of S. The concentration in the last statement holds even only with respect to the randomness involved in sampling U from S. We expect such a concentration since $H^{\otimes n}|S\rangle$ has high amplitude exactly on elements that are more likely to be included in U, while for states $|\psi\rangle$ supported on

S but orthogonal to $|S\rangle$, we expect the amplitude of $H^{\otimes n}|\psi\rangle$ to be distributed largely independently of the amplitudes in U.

Due to this concentration, with high probability $|S\rangle$ is a quantum witness that (S, U) are $\approx (a+b)$-spectrally Forrelated. Moreover, for any subset $\Delta \subseteq S$, the restriction of $a\,|S\rangle\langle S| + b \cdot \mathrm{id}_S$ to Δ is just $\frac{a|\Delta|}{\ell}\,|\Delta\rangle\langle\Delta| + b \cdot \mathrm{id}_\Delta$, whose top eigenvalue is just $(b + a|\Delta|/\ell) \ll (a+b)$. By choosing our thresholds for yes and no instances appropriately, we have a large family of strong yes instances.

In our actual proof, we do not require proving tight concentration bounds for the entries of the matrix M, though we expect that such concentration bounds do hold. Nevertheless, guided by this intuition, we are able to bound the top eigenvalues of M and its restrictions to small subsets Δ.

More specifically, we prove that with probability $1 - 2^{-\Omega(n)}$, a pair (S, U) sampled according to said procedure is a strong yes instance. Combined with the CumulativeSampler of Theorem 1.2, we derive that the CumulativeSampler must produce samples from S with probability at least $2^{-q} \cdot \Omega(T)^{-2v}$ when run on U for a pair (S, U) sampled according to said procedure.

Remark 1.5. *Looking ahead to our* QCMA *lower-bound, choosing $\gamma_z^{(S)}$ proportional to the* squared *amplitude is critical. It turns out that, if we instead choose $\gamma_y^{(S)}$ to be proportional to the (non-squared) amplitude, then the instance (S, U) we obtain is actually in* BQP*, and so also in* QCMA*. This is because we can approximately synthesize $H^{\otimes n}|S\rangle$ given access to just the sign of $\langle y|H^{\otimes n}|S\rangle$ by generating the state $\frac{1}{\sqrt{2^n}}\sum_y \mathrm{sgn}(\langle y|H^{\otimes n}|S\rangle)\,|y\rangle$ (as first observed by Irani* et. al. *[INN+21]). Therefore, the critical information is embedded in the signs of $\langle y|H^{\otimes n}|S\rangle$. However, by choosing $\gamma_y^{(S)}$ to be proportional to the squared amplitude, this information is not accessible to the verification algorithm. A second reason for the importance of squaring is that $\gamma_y^{(S)}$ is invariant under shifts of S, which plays a crucial role in our sampling upper bound.*

1.5 Sampling success probability upper bounds

The remainder of the proof is a sampling probability upper bound, which we can express as the following statement. We use the notation $\mathrm{poly}(\cdot)$ to vastly simplify the statement; the technical statement is Theorem 9.1.

Theorem 1.6 (Sampling probability upper bound (informal))**.** *Let* Strong *be the distribution over pairs (S, U) discussed previously. Then any T-query algorithm accessing only U produces v distinct samples from S with probability at most*

$$\left(\frac{\mathrm{poly}(n,T)}{\mathrm{poly}(\ell)}\right)^v + \left(\frac{\mathrm{poly}(n,T)\sqrt{\ell}}{2^{\Omega(n)}}\right)^v. \tag{5}$$

This theorem will contradict the conclusion of Theorem 1.2 for a choice of $v = \Omega(q)$.

To give intuition for the form of this bound, we can consider two edge cases. First, when ℓ is large, it becomes easy to sample points from S, as random points are likely to be in S. This gives us the second term in the sum, which becomes large when ℓ approaches $2^{O(n)}$. At the same time, when ℓ becomes too small, U becomes more concentrated, potentially revealing information about the structure of S. Our proof implicitly balances these two effects to achieve our sampling upper bound.

An insightful technique for proving upper bounds on the success probability of low-query quantum algorithms is to consider the behavior of the query algorithm on a superposition of possible oracles. This technique was first used in the adversary method of Ambainis [Amb02] (to generalize the Bennett *et. al.* [BBBV97] lower bound for unconstrained search). Instead of viewing the oracles as getting access to unitaries that modify the state of the algorithm, the adversary method treats the oracle as a long vector and each oracle query as a fixed phase kickback unitary on the joint state of the algorithm and oracle. For example, a query to the oracle U can be described by the controlled phase unitary,

$$|b, x, y\rangle \otimes |\mathrm{tt}_U\rangle \mapsto (-1)^{b \cdot U(y)} |b, x, y\rangle \otimes |\mathrm{tt}_U\rangle \tag{6}$$

where $|b, x, y\rangle$ is the state of the algorithm and $|\mathrm{tt}_U\rangle$ is the long vector description of the oracle (here, tt stands for "truth table"). We note that it is known that this kind of "phase" oracle is equivalent (up to a Hadamard transform) to the standard oracle. In this manner, it is natural to consider the behavior of an algorithm when run on the superposition over oracles. Studying the behavior when run on the superposition is a useful way of arguing query lower bounds and sampling probability upper bounds over the randomness in the oracle distribution and the randomness of the algorithm. A central challenge in proving quantum query lower bounds is designing a suitable perspective on the superposition over oracles that is clean enough to prove lower bound statements. Zhandry's compressed oracle technique [Zha19] is one method for effectively describing the superposition over oracles. Hamoudi and Magniez [HM23] were some of the first to use the compressed oracle technique to prove sampling probability upper bounds for problems such as unconstrained search and collision finding.

1.6 A bosonic perspective

To prove the desired sampling upper bound in this particular scenario, we introduce a new compression technique. We construct a compression of the superposition over oracles (S, U) by expressing the oracle in terms of bosons. This perspective will naturally clean up much of the technical calculations as well as provide a physical perspective on the query algorithm. This view of the superposition over oracle pairs (S, U) can be interpreted as a "first quantization" of compressed oracle techniques, which may be of independent interest. To understand this bosonic perspective in detail, it is helpful to initially ignore the oracle U and instead concentrate on only constructing a purification of uniformly sampling a multiset S of size ℓ.

There are two natural ways to express the multiset S: we can express the set as a vector in $(\{0,1\}^n)^\ell$, each coordinate corresponding to one point in S, and the value at that coordinate telling us the value of that point. This representation is, however, not unique, as permuting the vector does not change the multiset. Alternatively, we can express the multiset as a vector in $\mathbb{Z}_{\geq 0}^{2^n}$ of 1-norm ℓ with the x'th entry representing how many times x appears in S. The second perspective can be seen as tossing ℓ indistinguishable balls into 2^n bins with each toss being uniformly random.

The quantum mechanical analog of a multiset is a collection of *bosons*. Consider a system of ℓ bosons, each of which occupies a "state" from $\{0,1\}^n$. A characteristic of bosons is that any number of bosons may

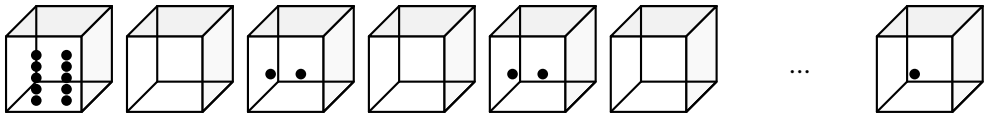


Figure 4: *A depiction of the bosonic Fock state* $|\psi\rangle = |10, 0, 2, 0, 2, 0, \ldots, 1\rangle$.

occupy the same state. Thus, the state of the ℓ bosons[7] is exactly described by a multiset S over $\{0,1\}^n$. Then, the two perspectives above give two different ways to represent a bosonic system.

We now give a brief primer on bosonic systems. Bosonic systems are described in terms of *modes*, each of which corresponds to an independent quantum degree of freedom. In physics, a mode may be defined in position space, where it is associated with a localized site or spatial region, or in momentum space, where it is associated with a plane-wave excitation. These two descriptions are related by a Fourier transform, so that switching between position and momentum modes is analogous to changing bases in a Hilbert space. In the position basis, one specifies a set of operators $\{\widehat{a}_x, \widehat{a}_x^\dagger\}$ that annihilate or create a boson at position x, respectively. The number of bosons in a particular (position) mode is unrestricted, with the number operator $\widehat{n}_x = \widehat{a}_x^\dagger \widehat{a}_x$ measuring the occupation of (position) mode x. This representation is natural when considering local interactions or spatially constrained dynamics. In the momentum basis, one works instead with operators $\{\widetilde{a}_x, \widetilde{a}_x^\dagger\}$ that annihilate or create excitations of definite momentum x, respectively. Analogously, the number of bosons in a momentum mode is unrestricted, with the number operator $\widetilde{n}_x = \widetilde{a}_x^\dagger \widetilde{a}_x$ measuring the occupation of momentum mode x. As one might suspect, the total number of bosons in the position basis $\widehat{N} = \sum_x \widehat{n}_x$ equals the total number of bosons in the momentum basis $\widetilde{N} = \sum_x \widetilde{n}_x$. The momentum description is particularly convenient for systems with translation invariance, where the state of the system takes a simple diagonal form in momentum space.

In this work, we consider a simplified "computer science" perspective on bosons. Our system consists of 2^n modes indexed by elements of $\{0,1\}^n$. Sometimes it will be convenient to index them using the isometry $[2^n] \equiv \{0,1\}^n$; so the 0-momentum mode is equivalent to the 0^n-momentum mode. Instead of relating the position and momentum creation/annihilation operators by the Fourier transform, we define our momentum creation/annihilation operators in terms of the Hadamard transform:

$$\widetilde{a}_x = \frac{1}{\sqrt{2^n}} \sum_{z \in \{0,1\}^n} (-1)^{x \cdot z} \widehat{a}_z, \quad \widetilde{a}_x^\dagger = \frac{1}{\sqrt{2^n}} \sum_{z \in \{0,1\}^n} (-1)^{x \cdot z} \widehat{a}_z^\dagger. \tag{7}$$

In this equation, the indexing variables x, z are elements of $\{0,1\}^n$ and $x \cdot z$ equals the inner product of the vectors over $\mathbb{F}_2$. In Section §8, we elaborate on the mathematics of bosons. For now, however, simply observe that the 0-momentum mode creation operator, $\widetilde{a}_0^\dagger$, is the uniform superposition over all position creation operators as it equals $\frac{1}{\sqrt{2^n}} \sum_z \widehat{a}_z^\dagger$. Therefore, the uniform superposition of a single boson in a random position mode is a single boson in the 0-momentum mode. This generalizes to: the uniform superposition of ℓ bosons each in an independently random position mode is ℓ bosons in the 0-momentum mode. We can interpret this as a computer science version of the *Heisenberg uncertainty principle* [Hei27]: certainty about the momentum of a set of bosons implies maximal uncertainty about the positions of the bosons.

[7] Whereas, for a fermionic system, by the exclusion principle, two fermions cannot occupy the same state, and so a fermionic system corresponds to an ordinary set.

Lastly, a convenient basis for studying bosonic states is the *Fock* basis, which records the number of bosons in each mode. One can write down Fock basis states in either the position or momentum basis. The bosonic (position) Fock basis is a collection of *orthonormal* states of the form $|\ell_0, \ell_1, \dots, \ell_{2^n-1}\rangle$ with each $\ell_x \in \mathbb{Z}_{\geq 0}$ where the total number of bosons is $\sum_x \ell_x$. In this setting, the annihilation and creation operators can be defined as the operators such that:

$$|\ell_0, \dots, \ell_x - 1, \dots, \ell_{2^n-1}\rangle = \frac{\widehat{a}_x}{\sqrt{\ell_x}} |\ell_0, \dots, \ell_x, \dots, \ell_{2^n-1}\rangle \tag{8a}$$

$$|\ell_0, \dots, \ell_x + 1, \dots, \ell_{2^n-1}\rangle = \frac{\widehat{a}_x^\dagger}{\sqrt{\ell_x + 1}} |\ell_0, \dots, \ell_x, \dots, \ell_{2^n-1}\rangle . \tag{8b}$$

1.7 The hardness of sampling without queries

Having set up this machinery, we observe that a purification of sampling a uniformly random multiset of size ℓ is exactly the state of ℓ bosons in the 0-momentum mode. And having defined a suitable purification, we can progress to sampling probability bounds. In the bosonic perspective, guessing elements of S is equivalent to identifying position modes where the algorithm believes bosons reside. A first step to proving that Theorem 1.2 is too good to be true, is showing that an algorithm that makes 0 queries to the oracle cannot identify with good probability position modes which contain bosons.

This "baby" problem is easy enough to prove via classical combinatorics, but we will purposefully resolve it by appealing to more quantum mechanical techniques. Observe that, by the relationship of position and momentum operators (eq. (7)), the state of a single boson in a fixed momentum mode $\widetilde{a}_x^\dagger |\text{vac}\rangle$ is equal to $\sum_y (-1)^{x \cdot y} \widehat{a_y}^\dagger |\text{vac}\rangle$, a superposition of being in every position mode, with phases corresponding to x. Therefore, the probability that any position guess y is correct is 2^{-n}. This is one example of a Heisenberg uncertainty principle [Hei27] with a more general form being that if the state of a single boson is supported on at most r momentum modes $x_1, \dots, x_r$, i.e., the state equals $\sum_{i=1}^r \alpha_i \widetilde{a}_{x_i}^\dagger |\text{vac}\rangle$, then the probability that any position guess y is correct is $\leq r/2^n$. In greater detail, what we can actually prove is for any $z \in \{0,1\}^n$, $\langle \psi | \widehat{n}_z | \psi \rangle = r\ell/2^n$ for any ℓ-boson state supported on at most r modes. Recall that $\widehat{n}_z$ is the operator *counting* the number of bosons at location z. We can define $\Pi_{\widehat{n}_z > 0}$ as the projector onto having a non-zero number of bosons in location z. Observe that $\widehat{n}_z \geq \Pi_{\widehat{n}_z > 0}$. Then, the probability that a guess of location z is correct is equal to $\langle \psi | \Pi_{\widehat{n}_z > 0} | \psi \rangle \leq \langle \psi | \widehat{n}_z | \psi \rangle = r\ell/2^n$. From here, we can conclude that the first guess has probability at most $r\ell/2^n$ of being correct, irrespective of which position mode z is guessed.

In this argument, we made a conceptual shift that turns out to be quite useful mathematically. Instead of studying the expectation of the state with respect to $\Pi_{\widehat{n}_z > 0}$, we chose to study the expectation with respect to $\widehat{n}_z$. This is mathematically equivalent to applying Markov's inequality ($\mathbb{E}[\Pi_{X>0}] = \mathbb{P}[X > 0] \leq \mathbb{E}[X]$ for any random variable $X \geq 0$). This weaker bound will be sufficient for our argument. This argument naturally generalizes: to upper bound the probability of finding bosons at each of $z_1, \dots, z_v$ for distinct indices $z_1, \dots, z_v$ (which is equal to the expectation with respect to $\Pi_{\widehat{n}_{z_1} > 0} \dots \Pi_{\widehat{n}_{z_v} > 0}$), the Markov inequality upper bound lets us instead bound the expectation with respect to $\widehat{n}_{z_1} \dots \widehat{n}_{z_v}$. Next, observe that $\widehat{n}_{z_1} \dots \widehat{n}_{z_v}$ equals $\widehat{a}_{z_1}^\dagger \dots \widehat{a}_{z_v}^\dagger \widehat{a}_{z_v} \dots \widehat{a}_{z_1}$ when the indices are distinct. Therefore, the expectation of $\widehat{n}_{z_1} \dots \widehat{n}_{z_v}$ is equal to $\left\| \widehat{a}_{z_v} \dots \widehat{a}_{z_1} |\psi\rangle \right\|^2$, the norm of the state after applying v annihilation operators.

To analyze the norm of applying two annihilation operators $\widehat{a}_{z_2}\widehat{a}_{z_1}$ when the state of the oracle is ℓ-bosons in the 0-momentum mode, we can study how the multiplicative norm decreases with each sequential operator application. We observe that the normalized state after annihilating a boson at location z_1 will again only be supported on bosons in the 0-momentum mode: annihilating this boson did not affect the remaining bosons, and so the resulting state will be $\ell - 1$ bosons in the 0-momentum mode. It follows that the following annihilation decreases the norm multiplicatively by at least $(\ell-1)/2^n$. We can continue this observation proving an upper bound of $\leq (\ell/2^n)^v$ for the norm of the vector after v annihilations. This proves a bound on the success probability of guessing v locations without making any queries at all. If our initial state was supported on r distinct momentum modes, the same argument would have produced an upper bound of $\leq (r\ell/2^n)^v$.

1.8 Understanding the action of queries to the U oracle

To prove an upper bound on the sampling probability after some queries to U, we need to extend the prior sampling probability upper bound to a broader set of initial states. The previous argument was particular to the state of ℓ bosons in the 0-momentum mode. To extend the argument, it is first illustrative to understand what the superposition of the algorithm and oracle registers looks like after making $T = \text{poly}(n)$ queries to the U oracle. What we discover is that the state of the oracle after T queries can be described as what we coin a *quasi-even condensate* and unpack below.

First, the phrase *condensate*. We borrow the term condensate from many-body physics, where it denotes a regime in which a macroscopic fraction of bosons occupy a single-particle mode. In our case, we use it to refer to a system where almost all the bosons are in the 0-momentum mode. Concretely, for us, a momentum Fock state is an r-condensate if at least $\ell - r$ of the bosons are in the 0-momentum mode. Equivalently, at most r bosons are not in the 0-momentum mode. An r-condensate may be a superposition of momentum Fock states that are r-condensates. This work considers states where $\ell = 2^{cn}$ and we will consider $r = \text{poly}(n)$, so only a negligible fraction of the bosons are not in the 0-momentum mode.

Second, the phrase *quasi-even*. For a momentum Fock state, we say that the state is o-quasi-even if at most o of the momentum modes (except the 0-mode) have an odd number of bosons in them. A general state is o-quasi-even if it is entirely supported on momentum Fock states which are o-quasi-even. A priori, the definition of quasi-even isn't nearly as motivated as the definition of a condensate (which has a natural physical interpretation). However, upon analysis of queries to the U oracle, studying how quasi-even a state is will become apparent.

Both being a condensate and being quasi-even are properties in the momentum basis. Therefore, we can define projectors onto the momentum Fock states that satisfy them. Since the projectors are both diagonal in the momentum Fock basis, they commute. So, the definition of a state as a (r, o)-quasi-even condensate is well-defined. Roughly speaking, our characterization theorem shows that the post-query state $|\psi_{\text{PQ}}\rangle$ satisfies the following: for all $\iota > 0$, there exists another state $|\psi'\rangle$ such that (a) $\| |\psi_{\text{PQ}}\rangle - |\psi'\rangle \| \leq \iota$ and (b) $|\psi'\rangle$ is a (r, o)-quasi-even condensate for $r = \text{poly}(n, T, \log(1/\iota))$ and $o \ll v/4$ with overwhelming probability (i.e., $(1 - \text{poly}(T)/\sqrt{\ell})^{O(v)}$) as long as $\text{poly}(T) \ll \ell$.

This particular characterization of the post-query state is important because we are additionally able to show that the sampling success probability of any algorithm for which the post-query state is such a quasi-even condensate decays exponentially fast with v. These two ingredients combine to prove our

sampling probability upper bound. We have not yet answered two fundamental questions: (1) why are post-query states effectively quasi-even condensates and how do we prove it, and (2) why is there a sampling probability upper bound for quasi-even condensates.

The answer to the first question comes from understanding the behavior of a single query to the oracle U at y. What we can formally prove is that a query at y applies a polynomial in the exponential function of the "double y-momentum hopping operator", which is defined as:

$$\widetilde{\mathbf{H}}_y \stackrel{\text{def}}{=} \frac{1}{\ell} \sum_{x,x' \in \{0,1\}^n} \widetilde{a}^\dagger_{x \oplus y} \widetilde{a}^\dagger_{x' \oplus y} \widetilde{a}_x \widetilde{a}_{x'}. \tag{9}$$

As the name may suggest, the action of $\widetilde{\mathbf{H}}_y$ on a momentum Fock state is to pick two modes x, x' which contain bosons and "hop" each of the bosons by y. The action of the $\widetilde{\mathbf{H}}_y$ moves around momentum Fock states, but perhaps surprisingly, it is diagonal in the position basis and its action can be described in terms of the coefficients $\gamma_z^{(S)}$ which were defined in eq. (4).

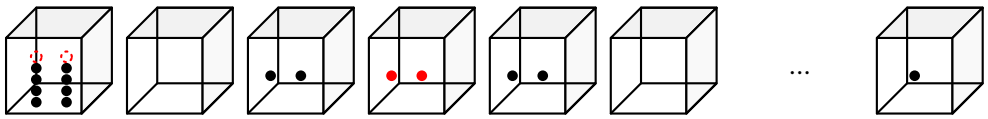


Figure 5: *A depiction of* $\widetilde{a}^\dagger_{0 \oplus 3} \widetilde{a}^\dagger_{0 \oplus 3} \widetilde{a}_0 \widetilde{a}_0 |\psi\rangle$ *which is one component in the sum* $\widetilde{\mathbf{H}}_3 |\psi\rangle$.

In slightly more detail, recall that we choose to include $y \in U$ with (independent) probability $1 - \mathrm{e}^{-\kappa \gamma_y^{(S)}}/2$ where $\gamma_y^{(S)}$ was defined in eq. (4) and κ is a small constant (think $1/10$). The use of the exponential functions will have further technical implications for the rest of the proof, but for intuition, observe that the first-order Taylor approximation of this probability equals[8] $1/2 + \kappa \gamma_y^{(S)}/2$. On the other hand, recall that shifts, or "hops", in the momentum basis are diagonalized in the position basis. For analogous reasons, $\widetilde{\mathbf{H}}_y$ is a diagonal matrix in the position Fock basis. Furthermore, for position Fock basis state $|\mathrm{tt}_S\rangle$ described by the multiset S of size ℓ, the corresponding diagonal entry is $\langle \mathrm{tt}_S | \widetilde{\mathbf{H}}_y | \mathrm{tt}_S \rangle = \gamma_y^{(S)} - 1$. See eq. (104) for details. Thus, we see that the diagonal entries of $\widetilde{\mathbf{H}}_y$ in the position Fock basis are closely related to the probability that z is included in U. This relationship allows us to express the effect of querying U at z in terms of $\widetilde{\mathbf{H}}_y$.

What we will show is that a query to a random U according to the prescribed description can be expressed as an exponential operator in the terms $-\gamma_y^{(S)}$, which by the prior relation is an exponential operator in the terms $-\widetilde{\mathbf{H}}_y$. Proving this to be true is significantly more challenging, and we save most of the details for the technical content of this paper. Of the numerous challenges required to prove this statement, the first is that for a fixed S oracle, we need to analyze the result of querying a purification of the distribution of U oracles. This is because evaluating whether the samples guessed by the algorithm are correct does not require U; therefore, the quantity we are actually interested in studying is the reduced density matrix on the algorithm's register and the register of the oracle encoding the S oracle – i.e., we can take the partial trace over the register encoding the U oracle. To make calculating this partial trace easier, we encode queries to the U oracle using another compressed oracle. By doing so, we can express

[8] An astute reader might question why the constants $1/2$ and κ appear. Both of these constants are necessary for our proof techniques, but are probably not necessary.

the action of queries to the U oracle on the algorithm and S oracle registers in terms of *Kraus operators*. These Kraus operators can be defined as a function of $\widetilde{\mathbf{H}}_y$ for the locations y queried by the algorithm.

This in turn leads us to the second challenge: the Kraus operators (formally defined in eq. (91)) will be exponential functions in $-\widetilde{\mathbf{H}}_y$. The exponential function appears here because we include $y \in U$ with probability $1 - \mathrm{e}^{-\kappa\gamma_y^{(S)}}/2$. Recall that our intended goal was to prove that the action of a query can be well approximated by a polynomial in $\widetilde{\mathbf{H}}_y$. However, we can observe from the definition that the values $\gamma_y^{(S)}$ can lie in the range $[0, \ell]$. But, any polynomial approximation of $\mathrm{e}^{-\kappa\gamma_y^{(S)}}$ that is ε-accurate on the entire region $[0, \ell]$ will require a very large degree — for example, a Taylor series approximation will require degree $\Omega(\ell \log(1/\varepsilon))$. Furthermore, a polynomial of such degree in $\widetilde{\mathbf{H}}_y$ applied to the initial state will no longer be a condensate, and we don't know of techniques for proving sampling upper bounds for states which are not condensates. The point being that an approximation on the entire range $[0, \ell]$ is untenable. However, it is also unnecessary. Instead, what we observe is that each $\gamma_y^{(S)}$ for $y \neq 0$ is approximately distributed as the square of a normal Gaussian. Furthermore, since our analysis is, in some sense, computing the average success probability over all possible (S, U) pairs, what is actually required is that the polynomial approximation of the exponential function is good *proportional* to the distribution over $\gamma_y^{(S)}$. In other words, since most of the mass is around 1, the polynomial approximation needs to be very strong in this region, but the approximation can be exponentially bad at large values, as the probability mass of $\gamma_y^{(S)}$ is likewise exponentially small. We show that applying "flat" polynomial approximations [BLMT24] to the exponential function[9] developed by Narayanan [Nar24] suffices to give a $\mathrm{poly}(n, t, \log(1/\iota))$-degree polynomial that well-approximates the action of a query for our means. See Subsection §12.1 for a technical overview of the polynomial approximation we take.

1.9 A characterization of post-query states

We promote this observation to prove that the action of the queries can be approximated to ι-precision by a polynomial of degree $\mathrm{poly}(n, T, \log(1/\iota))$ in the operators $\widetilde{\mathbf{H}}_y$. This will prove that the state is close to the subspace of r-condensates for $r = \mathrm{poly}(n, T, \log(1/\iota))$ as each action of the double-momentum hopping operator $\widetilde{\mathbf{H}}_y$ will move at most 2 bosons from the 0-momentum subspace.

However, recall that we start with $\ell = 2^{cn}$ bosons in the 0-momentum mode. Therefore, any $\mathrm{poly}(n)$ query algorithm, in expectation, moves an imperceptible number of the bosons from the 0-momentum mode. Again, this is the motivation behind the nomenclature choice of "condensate". Had we been able to prove that the state was an r-condensate for $r \ll v$, the proof would be over as we could apply an uncertainty principle. Unfortunately, this is not the case, as the sampler we construct makes $T = vt$ queries, if we started from a QCMA algorithm that made t queries.

What we will be able to prove is that with each query, the probability that the state becomes less "quasi-

[9] An astute reader might question why we choose to use an exponential function to define the probability of including $y \in U$. Indeed, we too spent much time attempting to prove the result without the use of an exponential function. If we had been able to use a polynomial function, then there would be no need to consider a suitable approximation. However, we were unable to find a low-degree polynomial that is within $[0, 1]$ on the range $[0, \ell]$ and has sufficient signal to provide a QMA algorithm for this problem. Therefore, we required using the exponential function to guarantee that the probability of including $y \in U$ was well defined. In turn, this necessitated the use of a flat approximation. We suspect that a truncation function (for example, $\max\{1, \varepsilon\gamma_y^{(S)}\}$ would have also sufficed, provided there exists a good flat approximation of this function.

even" is negligible in n. In total, it is very unlikely that the post-query state is not a quasi-even condensate. To understand why this is the case, let us imagine a simplified setting where the algorithm makes a query, the action on the oracle register is equal to applying the double-hopping operator $\widetilde{\mathbf{H}}_y$. This is an obvious simplification as the operator is not unitary, but ignore this for now. We think of this operator, in words, as the following operation: pick two random bosons at modes x and x', shift them to be at $x \oplus y$ and $x' \oplus y$, and multiply the norm of the state by the number of bosons already at modes $x \oplus y$ and $x' \oplus y$. Then it is clear to see that on a condensate (i.e., a state such that the vast majority of bosons are in the 0-momentum mode), the double-hopping operator is almost entirely characterized by its restriction onto the 0- and y-momentum modes (i.e., shifting bosons into or out of the condensate). If this really was an exact characterization, then the state of a quantum algorithm querying this non-physical operation would have exactly 2 bosons in every momentum mode that was queried by the algorithm, and every non-zero momentum mode would be perfectly paired up. The real double-hopping operator has some probability (we bound this by $\mathrm{poly}(r)/\sqrt{\ell}$) on each query to affect a boson that is not in the condensate, but since this number is negligible compared to the number of queries the algorithm makes, we can prove a strong bound on the probability that we observe $\ll v/4$ many momentum modes occupied by an odd number of bosons.

The analysis is not as simple, though, as the exact action of a query is not the double-hopping operator, but consists of terms looking like the exponential of the double-hopping operator: $\mathrm{e}^{-\kappa\widetilde{\mathbf{H}}_y}$. This presents new challenges to analyzing the effects of queries, which we resolve by using techniques from perturbation theory. Having shown that the double-hopping operator mostly does not change the number of odd indices in the condensate, we interpret it as a small perturbation (i.e., the part of the double-hopping operator that would change the number of odd indices) added to a large operator (i.e., the part of the double-hopping operator that preserves the evenness of all entries). Applying the Dyson series for the exponential function allows us to bound the effect of $\mathrm{e}^{-\kappa\widetilde{\mathbf{H}}_y}$ on the number of even entries in a similar manner as before.

1.10 Momentum conservation and boson pairs

One might look at the property of being a quasi-even condensate outlined in the previous subsection and wonder why it *should* say anything about the probability that a sampler succeeds at sampling many points from the oracle. To see this, it will be instructive to think about an analog of Noether's theorem [Noe18] to our oracle. In classical mechanics, Noether's theorem roughly states that conservation laws correspond to symmetries in a system. In our case, one such conservation law is the total momentum of the system. In particular, for all y, the double-hopping operator $\widetilde{\mathbf{H}}_y$ maps a system with total momentum 0 (where we define total momentum to use addition over $\mathbb{F}_2$) to a system that also has total momentum 0. Analogous with classical mechanics, this corresponds exactly to translational invariance. In other words, the state of the system does not change if we apply the shift operator, $\mathsf{Shift}_x \widehat{a}_z \mathsf{Shift}_x = \widehat{a}_{z\oplus x}$. Thus, the conservation of momentum trivially implies a bound on the probability that any algorithm outputs a single point of S.

However, making this observation about the system as a whole is too brittle on its own to bound the probability of sampling more than one point. At a proof level, an algorithm that successfully guesses one point from S will change the total momentum of the system by effectively annihilating one of the bosons. Even worse, an algorithm that queries the entire truth table of U would expect to be able to output many points of S given a single point in S. Thus, we must use a more fine-grained feature of the states of an

algorithm to provide an upper bound on the sampling probability.

One way to reconcile this is to make the following observation: translational invariance applies to *any* collection of bosons that have 0-total momentum, simply because applying the shift operator corresponds to applying a phase in the momentum basis, and having 0-total momentum is equivalent to the phases canceling each other out. Thus, whenever bosons are paired up (i.e., there is another boson in the same momentum mode), the pair of bosons will obey translational invariance, and (in a loose sense, since all bosons are identical) the sampler is very unlikely to sample one of the bosons from this pair. This property is far more robust than the original observation about 0-total momentum, as successfully sampling a boson affects at most one of the paired up bosons, although our sampler upper bound proof does not use an inductive argument. This critical observation is the reason why having almost all of the momentum modes occupied by an even number of bosons allows us to prove a sampler upper bound.

While the intuition that paired up bosons should be hard to find, our proof that algorithms whose purified states are supported on quasi-even condensates are hard to sample from does not use an inductive argument. Instead, we directly upper bound the spectral norm of any operator of the form $\widehat{n}_{z_1} \dots \widehat{n}_{z_v}$ on the subspace of quasi-even condensates using the max row 1-norm, and use the quasi-even and condensate properties to carefully bound the constructive interference that quasi-even condensates can generate.

1.11 Putting it all together

All together, we have argued that the existence of a QCMA algorithm for spectral Forrelation implies a sampler which morally guesses the positions of bosons given only query access to their momentum information. We then prove, by considering the query action on the purification of all oracles, that such a sampler cannot perform well unless it makes a superpolynomial number of queries.

The proof technique described in this paper is the technique we arrived at after considerable research and challenges. We highlight the challenges at the end of the result in our concluding remarks (Section §15).

2 History of the QMA versus QCMA problem

The question of whether QMA equals QCMA first appeared in the survey of Aharanov and Naveh [AN02] with the first indication of a separation given by Aaronson and Kuperberg [AK07]. The following is a brief review of the progress made on the problem since then: An early candidate classical oracle separation was given by Lutomirski [Lut11], but the candidate lacked a proof. More recently, a number of results have made progress towards the goal of a classical oracle separation by proving separations under different restrictions on how the oracle is accessed. Fefferman and Kimmel [FK15] showed a separation assuming that the oracle is an "in-place permutation oracle", a non-standard model where the oracle irreversibly permutes the input state; the resultant object was "less quantum" than a reflection about a Haar-random state but still inherently quantum. [FK15] also presents the first techniques for proving query lower bounds for QCMA; they lift an AM lower bound for set size estimation of Goldwasser and Sipser [GS86] to a QCMA lower bound for set size. Unfortunately, as is the case for many QCMA lower bounds, the same lower bound will apply for QMA, thereby not providing a separation. The corresponding QMA lower bound for set size estimation was proven later using Laurent polynomial techniques by Aaronson *et. al.* [AKKT20]. Future results starting from Natarajan and Nirkhe [NN24] showed QMA versus QCMA separations for weakened

notions of a prover or verifier. Natarajan and Nirkhe build on the prior works for set size estimation to construct a set size estimation problem with an efficient QMA algorithm by adding graph structure. For example, [NN24] proved the separation assuming the witness was only a function of some portion of the oracle – this can be equivalently expressed as a distribution testing problem. Recently, Agarwal and Kundu [AK25] have shown that one should be cautious when dealing with distributional oracles or unitary oracles as separations with respect to such oracles may not imply classical oracle separations. Secondly, a line of work has used quantum advantage relative to unstructured oracles [YZ24] to separate QMA from QCMA in settings where the quantum verifier was limited: [Liu22, LLPY23] gave a separation assuming the verifier can only make classical oracle queries, and more recently [BDK24] gave a separation which allows the verifier to make quantum queries, but assumes the adaptivity of the queries is sub-logarithmic.

Motivated by this seemingly dual requirement of having to apply quantum query complexity tools to highly structured oracles, a pair of works, by Zhandry [Zha25] (which this work takes some inspiration from) and Liu, Mutreja, and Yuen [LMY25], showed connections between the QMA versus QCMA problem and *pseudorandomness* against quantum adversaries. [LMY25] improved the lower bound analysis of [NN24] and proposed a conjecture about the pseudorandomness of δ-dense permutations, which, if proven, demonstrates another classical oracle separation between QMA and QCMA. Previously, [GLLZ21] proved that the δ-dense conjecture would imply that any quantum algorithm making queries to a random oracle can be simulated by an efficient classical one, a major open question in quantum query complexity [AA09]. A similar open problem would need to be resolved in order to make the separation in [LMY25] unconditional. We nevertheless believe that the results and the statistical conjecture of [LMY25] remain interesting, even in light of this result, since proving that the oracle of [LMY25] works to separate QCMA and QMA (either through their conjecture about δ-dense permutations or by different means) would provide an alternate oracle separation that sheds light on the Aaronson-Ambainis conjecture.

3 Observations and open questions

1. **Black-box separations and the two basis thesis** Recently, Ma and Natarajan [MN25] identified a QMA_1-complete family of local Hamiltonians where every term is 6-local and either diagonal in the standard or Hadamard bases. The collection of local Hamiltonian terms diagonal in the standard basis forms a constraint satisfaction problem (CSP). Likewise, the local Hamiltonian terms diagonal in the Hadamard basis form a second CSP. If we let S be the solutions to the first CSP and U the solutions to the second CSP, Ma and Natarajan's result can be expressed as the statement: "Deciding 1 vs $1 - 1/\mathrm{poly}(n)$ spectral Forrelation is QMA_1-complete even when S and U are the solution sets to CSPs". Previously, Cubitt and Montanaro [CM16] had proven that the local Hamiltonian problem where all terms are of the type $\alpha_{ij} X_i \otimes X_j$ or $\beta_{ij} Z_i \otimes Z_j$ is QMA-complete for a completeness-soundness gap of $1/\mathrm{poly}(n)$. However, Cubitt and Montanaro's Hamiltonians were necessarily *frustrated* as they were built from perturbation theory gadgets.

   The result presented here is an oracular variant of these two results. Ma and Natarajan's result also shows that it is QMA-complete to decide the $1 - 1/\exp(n)$ vs $1 - 1/\mathrm{poly}(n)$ spectral Forrelation problem even when S and U are the solution sets to CSPs. Whereas this result shows that the problem of deciding $59/100$ vs $57/100$ spectral Forrelation for general sets S and U is not in QCMA. Therefore, if QCMA were

to equal QMA, our black-box separation concretely says that the QCMA algorithm *must depend on the structure* of the two CSPs. This is the analog of how the unconstrained search problem lower bound for BQP [BBBV97] proves that if a BQP algorithm exists for NP, it must depend on the structure of the CSPs.

Furthermore, our oracle separation does not have perfect completeness, which, to the best of our knowledge, all previous candidate oracle separations did. And we do not know a technique for adapting this protocol to have perfect completeness, as the state $H^{\otimes n}|S\rangle$ has some support on every basis vector.

More broadly, our results fall into a larger family of quantum computation results that satisfy the "two-basis thesis"—that, computationally speaking, it suffices to consider computation or Hamiltonian terms that are either in the standard or Hadamard basis. Other results that satisfy the two-basis thesis include the BB84 protocol [BB14], Weisner's quantum money scheme [Wie83], Aaronson and Christiano's money scheme [AC13], the Mermin-Peres magic square game [Mer90, Per90], Mahadev's measurement protocols [Mah18], and quantum codes such as the NLTS Hamiltonian construction of Anshu, Breuckmann, and Nirkhe [ABN23].

2. **Upgrading quantum oracle separations** Variations of the Aaronson and Kuperberg [AK07] oracle are incredibly prevalent in quantum complexity theory and quantum cryptography. Having demonstrated that we can suitably replace the [AK07] quantum oracle with a classical oracle in this particular setting, we posit that more results are due for upgrades to classical oracles. In particular, we identify the question of separating QMA-search from QMA-decision with respect to a classical oracle as incredibly pertinent; the quantum oracle result was proven by Irani *et. al.* [INN$^+$21].

3. **Cryptographic primitives.** More broadly, quantum oracles are equally prevalent in quantum complexity theory and quantum cryptography. For example, many complexity-theoretic and cryptography *separations* are proven by means of quantum oracles. One notable example is Kretschmer's separation [Kre21] between quantum pseudorandomness and (for example) one-way functions. Very recently, there has been an explosion of unitary oracle separations separating various notions of quantum cryptography [CCS25, BCN25, BMM$^+$25, GZ25, Bar25, GLMY25, AGL25]. Often, for the cryptographic use cases, there hasn't been an easy classical oracle replacement. Can we use the techniques introduced in this work to make these improvements?

4. **The computational complexity of clonability** Nehoran and Zhandry [NZ24] introduce the concept of ClonableQMA, which is the class of decision problems decidable with a quantum witness that is also efficiently clonable. *It is not difficult to see that our proof also extends to separate* ClonableQMA *from* QMA *with respect to a classical oracle, as the sampler generated in Theorem 1.2 can be constructed given the clonability.* The complexity class ClonableQMA was identified as the complexity-theoretic generalization of many cryptographic tasks that build on the idea that some states are hard to clone while still easy to verify. Using the constructed oracle separation, can we work backwards and identify new cryptographic protocols that can be proven secure with respect to a classical oracle?

5. **Boolean function analysis** Liu, Mutreja, and Yuen [LMY25] identify an inherent connection between QMA vs QCMA and the Aaronson-Ambainis conjecture [AA09], a major open question in boolean function analysis. There are two versions of the Aaronson-Ambainis conjecture: one in terms of the influ-

ence of boolean functions and the other in terms of quantum query algorithms. Our result does not directly address either version of the Aaronson-Ambainis conjecture, but we hope that it might offer a new interesting perspective with which the problem might be tackled.

6. **Matching upper bounds.** We've shown that spectral Forrelation is hard for QCMA. However, we do not know what the optimal BQP or QCMA algorithms for this oracle are. Similarly, the problem of sampling points from S given oracle access to S and U (or just U) seems like an interesting quantum query complexity problem in its own right. What is the best sampling algorithm for producing v problems in S from access to both S and U?

7. **Proof improvements** Lastly, we admit that many of the components of our proof are likely suboptimal. We made many decisions in our proof that could be modified to make the proof simpler. One particular decision that stands out is to include an element y in the set U with probability $1 - e^{-\kappa\gamma_y^{(S)}}/2$. Is there a better choice of function that makes the proof simpler? We chose the exponential function since we could find a family of good polynomial approximations to the functions, and previously Zhandry [Zha25] used a similar function since it was amenable to integration. However, we do not know if this was optimal, as it caused other challenges. Other decisions, such as analyzing the bosonic system in terms of quasi-even condensates, could also be optimized. Is there a better generalization (instead of quasi-even condensates) for which we can prove sampling probability upper bounds? We suspect and hope that a simpler reformulation of this proof will be found.

4 Outline of the paper

This paper is broken up into multiple parts. In the remainder of Part I, we introduce relevant notations, definitions, and formalize the definitions of the complexity classes QMA, QCMA, and oracle separations. We, however, defer introducing the quantum mechanics of bosons to Section §8 of part III. Next, Part II and Parts III and IV play dual roles to each other, together proving the impossibility of an efficient query algorithm with a polynomial-sized witness for spectral Forrelation in the property-testing regime.

Part II proves that an efficient QCMA algorithm for the spectral Forrelation problem implies a sampling probability *lower bound* for a particular sampling task, namely sampling points from S given oracle access to U, the heavy points of the Hadamard transform of $|S\rangle$. The proof highly resembles the hybrid method originally used by [BBBV97], with the key idea being to use the fact that a classical witness can be reused without degrading its quality to get arbitrarily many samples from S.

Then, Parts III and IV prove a strictly smaller sampling *upper bound* for that same task. Part III introduces a broad family of states, which we call quasi-even condensates, that (approximately) satisfy a large number of translational symmetries. Using a combinatorial argument, we bound the maximum success probability of any sampler whose purifying register is a quasi-even condensate. Finally, Part IV shows that after only making a few queries to U, the purifying register of every query algorithm will be supported on quasi-even condensates. Part IV introduces a new compressed oracle technique for sparse oracles, and combines it with flat polynomial approximations to the exponential and tools from perturbation theory.

Finally, Part V brings together all the important results from the previous parts to prove a property testing oracle separation which can be lifted to a complexity class separation. We end Part V with concluding

remarks about how we came to this proof.

5 Preliminaries

Organizational notes Some of the proofs in this paper are naturally lengthy and require smaller technical/mathematical lemmas to prove. Notationally, we write these required lemmas in inline boxed environments. The intention is that a first pass reading of the paper can effectively skip over these lemmas by only reading the main listed lemmas/theorems in the boxed environments.

5.1 Mathematical notation

The following notations are used in this work. Most are standard, but we reiterate them for the sake of clarity.

- $\|\cdot\|$ or $\|\cdot\|_{\mathrm{op}}$ for a matrix will refer to the operator norm of a matrix, unless specified otherwise, and $\|\cdot\|_1$ denotes the matrix 1-norm, $\|X\|_1 = \mathrm{Tr}[|X|]$. $\|\cdot\|$ for a vector will refer to the 2-norm unless specified otherwise.
- When using the notation $\prod_{i=a}^{b} X_i$ for X_i that do not commute, the product expands from left to right starting from a. For example, when $a < b$, it expands as $X_a X_{a+1} \dots X_b$. When $b < a$, it expands as $X_a X_{a-1} \dots X_b$.
- Given a $x \in \{0,1\}^n$, we define 1_x to be the length 2^n tuple with entries indexed by n-bit strings, with 1 in the index corresponding to x, and 0 elsewhere.
- For two functions $f, g : \mathbb{N} \to \mathbb{R}$, we say that $f = O(g)$ and $g = \Omega(f)$ if $\exists\ n_0 \geq 0$ and $c > 0$ such that $f(n) \leq c g(n)$ for all $n \geq n_0$.
- We employ the notation $\delta_{x,y}$ for the indicator function of $x = y$ and we also use $\delta(b)$ for the indicator function when b is a boolean predicate such as $b = (x \in A)$.
- The binomial function $\binom{a}{b}$ can be extended naturally to non-integer valued a. In this paper, we will only need to use $a = 1/2$. For $b > 0$,

$$\binom{1/2}{b} \stackrel{\text{def}}{=} \frac{(1/2)(1/2-1)(1/2-2)\cdots(1/2-b+1)}{b!}. \tag{10}$$

  We also define $\binom{1/2}{0} \stackrel{\text{def}}{=} 1$. An important fact that we use is that the absolute value of this is always < 1 for $b > 0$.

Indexing Throughout most of this paper, we will be indexing over the set $\{0,1\}^n$; this set is isomorphic to the set $[N]$ where $N = 2^n$ under the lexicographic ordering. It is sometimes convenient to switch indexing sets between $\{0,1\}^n$ and $[N] \stackrel{\text{def}}{=} \{0, \dots, 2^n - 1\}$ for notational simplicity.

5.2 Quantum query complexity

Quantum query algorithms For this paper, we employ the following definitions of quantum query circuits/algorithms. These are standard definitions in the literature and are restated here for convenience. We use the term "an oracle of size n" to refer to a function $\mathcal{O} : \{0,1\}^n \to \{0,1\}$. This term matches notions from query complexity. We also use the phrase "an oracle" to refer to a function $\mathcal{O} : \{0,1\}^* \to \{0,1\}$ for a function mapping arbitrary strings to bits.

Definition 5.1 (Quantum query circuit/algorithm)**.** *For this paper, we define a* quantum query algorithm *as an inputless quantum circuit that interacts coherently with a boolean oracle function $\mathcal{O} : \{0,1\}^n \to \{0,1\}$. The oracle is accessed via a* phase query gate $|b, x\rangle \longmapsto (-1)^{b \cdot \mathcal{O}(x)} |b, x\rangle$*, which acts on an $n+1$-qubit query register. The algorithm is described by an alternating sequence of unitaries (drawn from any fixed universal gate set) and oracle gates. Here t denotes the* number of oracle queries *made by the algorithm and n denotes the size of the boolean oracle function. After all gates are applied, a designated output qubit is measured in the computational basis to determine acceptance. The initial state is assumed to be all qubits as $|0\rangle$ and all intermediate unitaries may act on an arbitrary (but finite) number of qubits. This definition naturally extends to multiple oracles (as will be the use case in this paper).*

This model captures the standard query-complexity viewpoint: there is no explicit input string—only access to the oracle $\mathcal{O}$ of a particular input size n. Furthermore, we don't place any restrictions on the complexity of the interleaved unitaries, and the overall circuit size is not of importance beyond the number of queries to $\mathcal{O}$. We note that the phase query gate described here is equivalent to other standard notions of oracle access, including the bit-flip oracle $|b, x\rangle \longmapsto |b \oplus \mathcal{O}(x), x\rangle$.

We note that a pair of functions $(S, U) : \{0,1\}^{n-1} \to \{0,1\}$ can be represented as a single function on n bits by

$$\mathcal{O}(b, x) = \begin{cases} S(x) & \text{if} \quad b = 0\,, \\ U(x) & \text{if} \quad b = 1\,. \end{cases} \tag{11}$$

Furthermore, controlled-access to the "combined unitary" can be constructed from controlled-access to both S and U. Therefore, for the remainder of this paper (and the statement of Theorem 1.1), we use the notation of accessing two oracles as it is equivalent and notationally easier to digest.

Definition 5.2 (Quantum query algorithm with witness)**.** *A quantum query algorithm may also receive an auxiliary* witness*. There are two main types of witnesses we consider.*

- *A* quantum witness *is a state $|\psi\rangle$ on q qubits.*
- *A* classical witness *is a bit string $w \in \{0,1\}^q$, treated as a computational-basis state.*

The definition is identical to the previous except the input state is now $|\psi\rangle \otimes |0 \dots 0\rangle$ or $|w\rangle \otimes |0 \dots 0\rangle$. The algorithm's acceptance probability may depend on both the oracle and the witness.

Quantum complexity theory The previous definitions of quantum query circuits/algorithms are defined in terms of a parameter n, which specifies the input length for oracle queries. Most of the paper will involve proving property-testing lower bounds for such algorithms. Only in the end will we need to reconcile these definitions with quantum complexity theory.

In complexity theory, we are interested in families of quantum algorithms that run on all possible input sizes. Notationally, n is often used as the length of the input, which we want to classify as a yes or no instance of a decision problem. An oracular complexity class is one defined in terms of a function $\mathcal{O} : \{0,1\}^* \to \{0,1\}$ (or equivalently a family of functions $\mathcal{O} : \{0,1\}^n \to \{0,1\}$ for each integer length n).

To formally define the quantum complexity classes $\mathsf{BQP}^{\mathcal{O}}$, $\mathsf{QCMA}^{\mathcal{O}}$, and $\mathsf{QMA}^{\mathcal{O}}$, we will need to define the notion of a uniform quantum oracle algorithm. For the purposes of this result, we will only need to define P-uniform quantum oracle algorithms. This is due to the result of Yao [Yao93], which proved that quantum complexity classes can be defined in terms of P-uniform families of quantum circuits.

Definition 5.3 (Quantum oracle circuits)**.** *We define a family of quantum oracle circuits/algorithms $\{\mathcal{A}_n\}_{n\geq 1}$, where the index n corresponds to the length of the explicit input to the computational problem. Each $\mathcal{A}_n$ can be represented concretely as a quantum circuit consisting of:*

- *elementary quantum gates drawn from a fixed, complete gate set (for example, Hadamard, phase, and controlled-NOT),*
- oracle phase gates *providing coherent access to the Boolean oracles*

$$\mathcal{O}_k : \{0,1\}^k \to \{0,1\} \text{ by } |b,x\rangle \mapsto (-1)^{b\cdot\mathcal{O}_k(x)}|b,x\rangle \text{ for } x \in \{0,1\}^k, b \in \{0,1\}. \tag{12}$$

This can be viewed as the circuit model definition of accessing $\mathcal{O}$ at various lengths.

The family $\{\mathcal{A}_n\}$ is P-uniform *if there exists a deterministic polynomial-time Turing machine M that, on input 1^n, outputs a full classical description of the circuit $\mathcal{A}_n$. Because M runs in time polynomial in n, the resulting circuit $\mathcal{A}_n$ must satisfy the following polynomial bounds for some polynomial functions:*

- *it takes as input a classical input of size n and consists of unitary gates followed by the measurement of a single qubit for a binary output,*
- *it contains at most* $\mathrm{poly}(n)$ *gates (either oracle or elementary), and*
- *the largest length k which it can query $\mathcal{O}_k$ is at most* $\mathrm{poly}(n)$.

This ensures that the circuit family represents an efficiently describable quantum algorithm operating within polynomial resources.

Using the definition of P-uniform quantum oracle algorithms, we define the standard oracle quantum complexity classes. The natural extension of BQP is the following.

Definition 5.4 (Oracle BQP). *A promise language $\mathcal{L}^{\mathcal{O}} = (\mathcal{L}_{\text{yes}}, \mathcal{L}_{\text{no}})^{\mathcal{O}} \subseteq \{0,1\}^*$ is in $\mathsf{BQP}^{\mathcal{O}}$ if there exists a P-uniform family of quantum oracle circuits $\mathcal{A}_n$ such that for every input x of length $n = |x|$,*

$$\textit{(Completeness)} \quad x \in \mathcal{L}_{\text{yes}} \implies \mathbb{P}[\mathcal{A}_n^{\mathcal{O}}(x) \textit{ accepts}] \geq \tfrac{2}{3}, \tag{13a}$$

$$\textit{(Soundness)} \quad x \in \mathcal{L}_{\text{no}} \implies \mathbb{P}[\mathcal{A}_n^{\mathcal{O}}(x) \textit{ accepts}] \leq \tfrac{1}{3}. \tag{13b}$$

The pair of constants $2/3$ and $1/3$ is not important. Standard parallel repetition techniques can be used to choose different constants or even functions as close as $(1/2+1/\text{poly}(n), 1/2-1/\text{poly}(n))$ or as far apart as $(1-2^{-\text{poly}(n)}, 2^{-\text{poly}(n)})$. Next, we define the generalizations of QCMA and QMA.

Definition 5.5 (Oracle QCMA). *A promise language $\mathcal{L}^{\mathcal{O}} = (\mathcal{L}_{\text{yes}}, \mathcal{L}_{\text{no}})^{\mathcal{O}} \subseteq \{0,1\}^*$ is in $\mathsf{QCMA}^{\mathcal{O}}$ if there exists a P-uniform family of quantum oracle circuits $\mathcal{A}_n$ with $\mathcal{A}_n$ accepting a witness of length $q(n)$, such that for every input x of length $n = |x|$,*

$$\textit{(Completeness)} \quad x \in \mathcal{L}_{\text{yes}} \implies \exists\ w \in \{0,1\}^{q(n)} \textit{ s.t. } \mathbb{P}[\mathcal{A}_n^{\mathcal{O}}(x, w) \textit{ accepts}] \geq \tfrac{2}{3}, \tag{14a}$$

$$\textit{(Soundness)} \quad x \in \mathcal{L}_{\text{no}} \implies \forall\ \widetilde{w} \in \{0,1\}^{q(n)},\ \mathbb{P}[\mathcal{A}_n^{\mathcal{O}}(x, \widetilde{w}) \textit{ accepts}] \leq \tfrac{1}{3}. \tag{14b}$$

Definition 5.6 (Oracle QMA). *A promise language $\mathcal{L}^{\mathcal{O}} = (\mathcal{L}_{\text{yes}}, \mathcal{L}_{\text{no}})^{\mathcal{O}} \subseteq \{0,1\}^*$ is in $\mathsf{QMA}^{\mathcal{O}}$ if there exists a P-uniform family of quantum oracle circuits $\mathcal{A}_n$ with $\mathcal{A}_n$ accepting a witness of length $q(n)$, such that for every input x of length $n = |x|$,*

$$\textit{(Completeness)} \quad x \in \mathcal{L}_{\text{yes}} \implies \exists\ |\psi\rangle \in (\mathbb{C}^2)^{\otimes q(n)} \textit{ s.t. } \mathbb{P}[\mathcal{A}_n^{\mathcal{O}}(x, |\psi\rangle) \textit{ accepts}] \geq \tfrac{2}{3}, \tag{15a}$$

$$\textit{(Soundness)} \quad x \in \mathcal{L}_{\text{no}} \implies \forall\ |\widetilde{\psi}\rangle \in (\mathbb{C}^2)^{\otimes q(n)},\ \mathbb{P}[\mathcal{A}_n^{\mathcal{O}}(x, |\widetilde{\psi}\rangle) \textit{ accepts}] \leq \tfrac{1}{3}. \tag{15b}$$

5.3 Sets, multisets, and functions

Throughout the paper, we use the following notation to refer to sets, multisets, and the functions associated with them.

- Given a multiset S, we use the notation $\{x_1, \ldots, x_\ell\}$ to refer to the *unordered* collection of elements in S, where the elements $x_1, \ldots, x_\ell$ may be non-distinct.
- We abuse notation and take $S(x)$ to be the indicator function $\delta(x \in S)$, which is 1 if x is in S with any multiplicity, and 0 otherwise.
- We will use the notation $\mathcal{O}_S$ to refer to the phase oracle for the function $S(\cdot)$, i.e., the unitary that acts as $\mathcal{O}_S |b, x\rangle = (-1)^{b \cdot S(x)} |b, x\rangle$.
- When applying operations between sets, like $S \setminus T$, we refer to the operation of first mapping S and T to the set of distinct elements in S and T respectively, and then outputting the operation applied to the resulting sets. We also use the notation $T \subseteq S$ to mean that T is a subset of the set of distinct elements of S.
- We also define a projector Π_S for a multiset S to be the projection onto basis states x in S, treating S as a set, i.e., $\sum_{x \in S} |x\rangle\langle x|$.

- Throughout the paper, we use the terminology "oracle access to a multiset S" to mean query access to $\mathcal{O}_S$, and may write an algorithm querying $\mathcal{O}_S$ as $\mathcal{A}^S$ for brevity. Algorithms may receive access to multiple multisets, in which case we write $\mathcal{A}^{(S_1,S_2)}$.
- We will also write, for a multiset $S = \{s_1, \ldots, s_\ell\} \subseteq \{0,1\}^n$, the state $|S\rangle = \frac{1}{\sqrt{\ell}} \sum_{i=1}^{\ell} |s_i\rangle$ in the Hilbert space $(\mathbb{C}^2)^{\otimes n}$. Note that when S has no multiplicities greater than 1, this is a normalized state, but otherwise it may be unnormalized. This is as opposed to the state $|\mathrm{tt}_S\rangle$ in the Hilbert space $\mathbb{C}^{\mathbb{Z}_{\geq 0}^{2^n}}$, the classical description of the multiset.

5.4 Spectral Forrelation

Definition 5.7 (Spectral Forrelation)**.** *We say two subsets $S, U \subset \{0,1\}^n$ are α-spectrally Forrelated if*

$$\alpha = \left\| \Pi_U \cdot H^{\otimes n} \cdot \Pi_S \right\|_{\mathrm{op}}^2 = \max_{\||\psi\rangle\|=1} \left\| \Pi_U \cdot H^{\otimes n} \cdot \Pi_S |\psi\rangle \right\|^2. \tag{16}$$

Here, the projectors Π_U and Π_S are the projectors onto the subspaces spanned by the basis vectors $|x\rangle$ for $x \in U$ and $x \in S$, respectively.

Remark 5.8. *First note that this definition can be extended to multisets S and U by taking the set of elements included in S and U. In this way, spectral Forrelation is defined to be a property of pairs of multisets as well.*

This definition can also be extended to functions/oracles $S, U : \{0,1\}^n \to \{0,1\}$ by taking the sets S and U to be the pre-images of 1 for the functions S and U respectively. In this way, spectral Forrelation is also defined as a property of a pair of oracles.

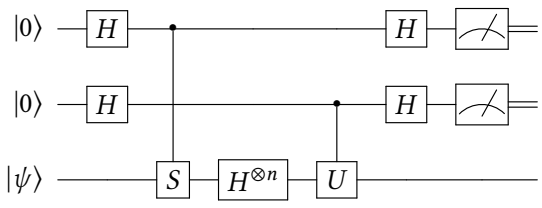
Remark 5.9. *The spectral Forrelation of any two oracles is a number $\in [0,1]$. By definition, if either oracle is the constant function 0 (i.e., corresponds to the empty set), then the spectral Forrelation is 0.*

In this paper, we will use the notation "at least α-spectrally Forrelated" to mean that a pair of subsets are β-spectrally Forrelated for some $\beta \geq \alpha$, and similarly for "at most α-spectrally Forrelated". We will typically refer to a pair (S, U) that is at least $59/100$-spectrally Forrelated as a yes instance of spectral Forrelation, and a pair (S, U) that is at most $57/100$-spectrally Forrelated as a no instance.

The following is implicit in [Zha25], but we re-prove it here.

Theorem 5.10 (QMA containment)**.** *For any $\alpha > \beta$, there is a $O(1/(\alpha-\beta)^2)$ quantum query algorithm with a n-qubit quantum witness that, given oracle access to oracles $S, U : \{0,1\}^n \to \{0,1\}$, accepts with probability at least $2/3$ if they are at least α-spectrally Forrelated (yes instances), and probability at most $1/3$ if they are at most β-spectrally Forrelated (no instances).*

Proof. There is a simple verifier that accepts yes instances with probability $\geq \alpha$ and accepts no instances with probability $\leq \beta$: Let $|\psi\rangle$ be the n-qubit quantum witness. The following quantum circuit describes the verifier.



The verifier accepts if both measurements equal 1. It is easy to check that the probability that both measurements output 1 is exactly $\|\Pi_U \cdot H^{\otimes n} \cdot \Pi_S |\psi\rangle\|^2$. By definition, there exists a $|\psi\rangle$ such that this probability is $\geq \alpha$ for yes instances and for all $|\psi\rangle$, this probability is $\leq \beta$ for no instances. Using the Marriott-Watrous amplification [MW05], using $O(1/(\alpha - \beta)^2)$ queries, we can convert the verifier into one that decides spectral Forrelation with completeness-soundness of $(2/3, 1/3)$. □

Note that the previous circuit is effectively measuring the input witness with projectors Π_S and $H^{\otimes n} \cdot \Pi_U \cdot H^{\otimes n}$. The Marriott-Watrous amplification protocol [MW05] will consist of alternating these two measurements and using the classical outputs from the measurements to decide the problem.

Part II
From QCMA algorithms to samplers

The goal of Part II is to show that there is a reduction from a QCMA algorithm that decides spectral Forrelation (i.e., whether two oracles of length n are at least 59/100-spectrally Forrelated or at most 57/100-spectrally Forrelated, given a classical witness), to a sampler such that for all (S, U) that are at least 59/100-spectrally Forrelated, the sampler outputs many points from S, only querying U.

Section §6 will describe a general reduction for taking QCMA algorithms that query two oracles and decide between two families of (pairs of) oracles, and producing samplers that query one of the two oracles. The reduction will work for families of oracles satisfying a special property. Given a family of yes and no instances defined on pairs of oracles (S, U), we say that (S, U) is a *strong yes instance* if (S, U) is a yes instance, and for any small $\Delta \subseteq S$, (Δ, U) is a no instance. This section will show that any QCMA algorithm that distinguishes between families of pairs of yes and no oracle instances of this problem can be used to sample points from strong yes instances. Finally, Section §7 shows that if we define yes instances to be sets (S, U) that are at least 59/100-spectrally Forrelated, and no instances to be sets (S, U) that are at most 57/100-spectrally Forrelated, there is a procedure for sampling from a large family of strong yes instances. These constants (59/100 and 57/100) are artifacts of the proof, and not necessarily fundamental to spectral Forrelation.

6 Constructing samplers from strong yes instances

Assume the existence of a t-query quantum query algorithm $\mathcal{A}^{(S,U)}$ with q-bit classical witness which solves the spectral Forrelation problem. Formally, assume the quantum query algorithm $\mathcal{A}^{(S,U)}$ has the following properties:

1. (Completeness) If (S, U) is at least 59/100-spectrally Forrelated, there exists a witness $w \in \{0,1\}^q$ such that $\mathcal{A}^{(S,U)}(w)$ accepts with probability at least 2/3.

2. (Soundness) If (S, U) is at most 57/100-spectrally Forrelated, then for all witnesses $\widetilde{w} \in \{0,1\}^q$, $\mathcal{A}^{(S,U)}(\widetilde{w})$ accepts with probability at most 1/3.

In this section, we describe pairs (S, U) that are at least 59/100-spectrally Forrelated as yes instances of spectral Forrelation, and pairs (S, U) that are at most 57/100-spectrally Forrelated as no instances of spectral Forrelation.

As described in the preliminaries, oracle access to the multisets S and U can be described as the linear extensions of the following maps

$$|b, x\rangle |z\rangle \overset{\mathcal{O}_S}{\mapsto} (-1)^{b \cdot S(x)} |b, x\rangle |z\rangle \ , \tag{17a}$$

$$|b, x\rangle |z\rangle \overset{\mathcal{O}_U}{\mapsto} (-1)^{b \cdot U(x)} |b, x\rangle |z\rangle \ . \tag{17b}$$

where b is a control bit, x describes the input to the oracle, and z describes the state of the remainder of the system. The algorithm $\mathcal{A}^{(S,U)}$ can be thought of starting from a state $|w\rangle |0 \dots 0\rangle$ and applying a sequence

of general unitaries interlaced with t queries to S and t queries to U. The algorithm $\mathcal{A}^{(S,U)}$ concludes by measuring the first qubit in the standard basis.

6.1 Sampling from S

We can treat the algorithm $\mathcal{A}^{(S,U)}$ as $(\mathcal{A}^U)^S$ — i.e., an algorithm $\mathcal{A}^U$, consisting of standard unitary gates, as well as U gates, that makes queries the S oracle. When expressed this way, the state of the algorithm $(\mathcal{A}^U)^S$ immediately before its final measurement is given by

$$V_t \mathcal{O}_S V_{t-1} \mathcal{O}_S \dots \mathcal{O}_S V_0 \left|w, 0\right\rangle \tag{18}$$

with the unitaries $\{V_j\}$ including queries to U. Consider the following algorithm, which takes as input a witness w and a set $\Delta \subseteq \{0,1\}^n$. The algorithm $\mathsf{Sampler}^U(w, \Delta)$ is a sampler for generating an additional sample from S given a prior subset Δ of found points. Recall that for a multiset S and set Δ, we are taking $S \setminus \Delta$ to be the set of elements that appear with multiplicity at least 1 in S that do not appear in Δ.

Query algorithm $\mathsf{Sampler}^U(w, \Delta)$:

1. Sample $j \leftarrow \{0, \dots, t-1\}$ uniformly randomly.
2. Compute the state $V_j \mathcal{O}_\Delta V_{j-1} \mathcal{O}_\Delta \dots \mathcal{O}_\Delta V_0 \left|w, 0\right\rangle$, where $\mathcal{O}_\Delta$ is the unitary defined by
$$\left|b, x\right\rangle \left|z\right\rangle \overset{\mathcal{O}_\Delta}{\mapsto} (-1)^{b \cdot \delta(x \in \Delta)} \left|b, x\right\rangle \left|z\right\rangle . \tag{19}$$
3. Measure the state in the standard basis for output (b, x, z).
4. If $x \in \Delta$, output the alphabetically first symbol not in Δ. Else, output x.

A minor adjustment can be made to ensure that $\mathsf{Sampler}$ makes exactly t queries to U. At a high level, the idea behind the sampler is that for $\mathcal{A}^U$ to distinguish between S and Δ, it must be querying points in S that are not in Δ, since the oracles f_Δ and S are identical outside of those points. We prove the following claim about $\mathsf{Sampler}$. The proof closely resembles the hybrid method of [BBBV97].

Lemma 6.1. *Let (S, U) be a yes instance of the oracle problem (i.e., at least $59/100$-spectrally Forrelated). Let $\Delta \subseteq S$ be a subset of the set of elements in S such that (Δ, U) is a no instance (i.e., at most $57/100$-spectrally Forrelated). Let w be a witness that causes the query algorithm $\mathcal{A}$ to accept (S, U) with probability at least $2/3$. Then, the algorithm $\mathsf{Sampler}^U(w, \Delta)$ makes t queries to the oracle U, no queries to the oracle S, and produces a sample from $S \setminus \Delta$ with probability at least $\left(\frac{1}{36t^2}\right)$.*

Proof. We begin by defining a sequence of hybrids. Recall that the verification algorithm's state immediately before its final measurement is given by eq. (18). Let the j-th hybrid state $\left|h_j(w)\right\rangle$ be define as

$$\left|h_j(w)\right\rangle \overset{\text{def}}{=} V_t \mathcal{O}_S \dots \mathcal{O}_S \underbrace{V_j \mathcal{O}_\Delta V_{j-1} \mathcal{O}_\Delta \dots V_1 \mathcal{O}_\Delta V_0 \left|w, 0\right\rangle}_{\overset{\text{def}}{=} \left|\psi_j(w)\right\rangle}, \tag{20}$$

where we call the state $\left|\psi_j(w)\right\rangle$ the j-th prefix state. Intuitively, the prefix state $\left|\psi_j(w)\right\rangle$ corresponds to running the algorithm $\mathcal{A}$ with oracles (Δ, U) until the $(j+1)$-th query, and the hybrid state $\left|h_j(w)\right\rangle$ corresponds to replacing the first j queries to the S oracle with queries to the Δ oracle.

Then, $|h_0(w)\rangle$ corresponds to running $\mathcal{A}$ on (Δ, U) up until the final measurement and $|h_t(w)\rangle$ corresponds to running $\mathcal{A}$ on (S, U) up until the final measurement. Since (Δ, U) is a no instance of spectral Forrelation and (S, U) is a yes instance of spectral Forrelation, there is a measurement that accepts $|h_0(w)\rangle$ with probability at most $1/3$ and accepts $|h_t(w)\rangle$ with probability at least $2/3$, namely to measure the first qubit in the computational basis. Therefore, we have that

$$\frac{1}{3} \leq \mathrm{Tr}\left[|0\rangle\langle 0| \left(|h_t(w)\rangle\langle h_t(w)| - |h_0(w)\rangle\langle h_0(w)|\right)\right] \tag{21a}$$

$$\leq \frac{1}{2}\left\| |h_t(w)\rangle\langle h_t(w)| - |h_0(w)\rangle\langle h_0(w)| \right\|_1 \tag{21b}$$

$$\leq \left\| \, |h_t(w)\rangle - \, |h_0(w)\rangle \right\|. \tag{21c}$$

A proof of these inequalities can be found in e.g. [Wil17]. By the triangle inequality, we have that

$$\frac{1}{3} \leq \left\| |h_t(w)\rangle - |h_0(w)\rangle \right\| \tag{22a}$$

$$\leq \sum_{j=1}^{t} \left\| \, |h_j(w)\rangle - \, |h_{j-1}(w)\rangle \right\| \tag{22b}$$

$$\leq \sum_{j=0}^{t-1} \left\| (\mathcal{O}_\Delta - \mathcal{O}_S) \, |\psi_j(w)\rangle \right\| \tag{22c}$$

$$= 2 \sum_{j=0}^{t-1} \left\| \Gamma \, |\psi_j(w)\rangle \right\| \quad \text{where } \Gamma \stackrel{\text{def}}{=} \sum_{x \in S \setminus \Delta} |1, x\rangle\langle 1, x| \otimes \mathrm{id}. \tag{22d}$$

Here, we use the fact that the j and $(j+1)$'st hybrids differ only by the oracle query immediately after the j-th prefix state, and we recall that S is the controlled phase flip oracle that applies a sign of -1 when $b = 1$ and $x \in S$, so $\mathcal{O}_\Delta - \mathcal{O}_S$ is exactly twice the projector onto strings of the form $|1, x\rangle$ for $x \in S \triangle \Delta = S \setminus \Delta$ as $\Delta \subseteq S$. Here, recall that operations like set minus and symmetric difference act on S and Δ as if they were sets. For each $j \in \{0, \ldots, t-1\}$, express

$$|\psi_j(w)\rangle = \sum_x \beta_x^{(j)} \, |x\rangle \otimes \, |\psi_j(x, w)\rangle \;\text{ for } \beta_x \in \mathbb{R}^+. \tag{23}$$

Here, the control bit b is captured in $\big|\psi_j(x, w)\big\rangle$. Then, using the Cauchy-Schwarz inequality, we have that,

$$\frac{1}{6} \leq \sum_{j=0}^{t-1} \sqrt{\sum_{x \in S \setminus \Delta} (\beta_x^{(j)})^2} \leq \sqrt{t} \sqrt{\sum_{j=0}^{t-1} \sum_{x \in S \setminus \Delta} (\beta_x^{(j)})^2}. \tag{24}$$

This implies that,

$$\frac{1}{36t^2} \leq \frac{1}{t} \sum_{j=0}^{t-1} \sum_{x \in S \setminus \Delta} (\beta_x^{(j)})^2. \tag{25}$$

The right hand side is exactly the probability that $\mathsf{Sampler}^U(w, \Delta)$ outputs a x from $S \setminus \Delta$. □

6.2 Witness-free sampler

The algorithm $\mathsf{Sampler}^U(w, \Delta)$ requires a witness w to run. Furthermore, conditioned on $\mathsf{Sampler}^U(w, \Delta)$ outputting a novel point from S, we can iterate the sampler to generate multiple points. This generates a notion of a Cumulative Sampler which also requires a witness. However, we can simply guess the witness w at a cost to the success probability of the algorithm:

Query algorithm $\mathsf{CumulativeSampler}^U$:

1. Sample a random $w \in \{0,1\}^q$.
2. Initialize $\Delta \leftarrow \emptyset$.
3. For v rounds, run $\mathsf{Sampler}(w, \Delta)$ with fresh independent randomness. Append output x: $\Delta \leftarrow \Delta \cup \{x\}$.
4. Output the resulting Δ.

Notice that the algorithm $\mathsf{CumulativeSampler}^U$ takes no witness as an input. Meaning the same sampler is producing samples from S with the aforementioned probability for *every* yes instance (S, U) such that (Δ, U) are no instances for small subsets Δ. We formalize this observation in the following theorem.

Theorem 6.2 (Good samplers from QCMA algorithms)**.** *Assume there exists a quantum query algorithm $\mathcal{A}$ with classical witness for spectral Forrelation instances such that for instances of size n, $\mathcal{A}$ takes a q-sized classical witness and makes t oracle queries. Let (S, U) be a yes instance of spectral Forrelation and $v \in \mathbb{N}$ be such that (Δ, U) is a no instance of spectral Forrelation for all subsets $\Delta \subset S$ with $|\Delta| \leq v$. There exists a query algorithm* CumulativeSampler *(with implicit dependence on v) such that for* $\mathsf{CumulativeSampler}^U$ *makes 0 queries to S, vt queries to U and produces v unique samples from S with probability at least*

$$\geq 2^{-q} \cdot \left(\frac{1}{36t^2}\right)^v . \tag{26}$$

Proof. Define the following events with respect to the running of CumulativeSampler. Let G be the event that the witness w sampled is a good witness for the pair (S, U). Second, let $E_1 \ldots, E_v$ be the events that the corresponding guesses are in S. By construction, the samples are distinct. The probability they are all correct is

$$\begin{aligned}
\mathbb{P}[E_v, \ldots, E_1] &\geq \mathbb{P}[G]\,\mathbb{P}[E_v, \ldots, E_1 | G] && \text{(27a)}\\
&\geq \mathbb{P}[G] \cdot \prod_{j=1}^{v} \mathbb{P}[E_j | E_{j-1}, \ldots, E_1, G] && \text{(27b)}\\
&\geq 2^{-q} \cdot \left(\frac{1}{36t^2}\right)^v . && \text{(27c)}
\end{aligned}$$

Here, we apply Theorem 6.1, which tells us that the probability that the j-th sample is in S, conditioned on the first $j-1$ samples being in S is at least $\frac{1}{36t^2}$. □

7 Strong yes instances for spectral Forrelation

Theorem 6.2 demonstrates that there exists a sampler algorithm with good success probability that, when given oracle access to the U oracle in a yes instance (S, U) of spectral Forrelation, with the additional property that (Δ, U) is a no instance of spectral Forrelation for all subsets $\Delta \subset S$ such that $|\Delta| \leq v$, outputs v points from S. We define such yes instances as *strong* yes instances.

Definition 7.1 (Strong yes instance)**.** *For any t_1, t_2, v with $t_1 < t_2$, we will say that a pair (S, U) is a (t_1, t_2, v)-strong yes instance if*

1. *(completeness): (S, U) are at least t_2-spectrally Forrelated.*
2. *(soundness): For any subset $\Delta \subset S$ such that $|\Delta| \leq v$, (Δ, U) are at most t_1-spectrally Forrelated.*

Intuitively, disproving the consequence derived in Theorem 6.2 will require constructing many strong yes instances and proving that no small query sampler can be successful for all these yes instances. To create yes instances, we want to sample a random multiset S of size ℓ, and then construct (with high probability) a set U that is Forrelated with S. For this, define

$$\gamma_y^{(S)} = \left(\frac{1}{\sqrt{\ell}} \sum_{i \in [\ell]} (-1)^{y \cdot s_i} \right)^2 = \frac{1}{\ell} \sum_{i,j} (-1)^{y \cdot (s_i + s_j)} = 1 + \frac{1}{\ell} \sum_{i \neq j} (-1)^{y \cdot (s_i + s_j)} . \tag{28}$$

Observe that $\gamma_y^{(S)}$ equals $2^n \cdot |\langle y | H^{\otimes n} | S \rangle|^2$, where $|S\rangle$ is the superposition over S, weighted by the multiplicities of the elements, divided by $\sqrt{\ell}$. When S is a set (i.e., has no multiplicities that are not 0 or 1), this is a normalized state, but if it is a multiset it may not be normalized. The following lemma proves that there exists a distribution that is overwhelmingly supported on strong yes instances. In conjunction with Theorem 6.2, it gives that a QCMA algorithm implies a sampler for a particular distribution. In the next section (Section §11), we will prove that such a sampler cannot exist.

Definition 7.2 (The Strong distribution)**.** *Let the distribution Strong_κ over pairs (S, U) be defined by*

(a) sampling the multiset $S = \{s_1, \ldots, s_\ell\}$ of size ℓ by sampling each s_i uniformly randomly from $\{0,1\}^n$,

(b) followed by, sampling a set U by adding each point $y \in \{0,1\}^n \setminus \{0^n\}$ to U with probability $1 - \mathrm{e}^{-\kappa \gamma_y^{(S)}}/2$ where $\gamma_y^{(S)}$ is defined in eq. (28).

The distribution $\mathsf{Strong} = \mathsf{Strong}_\kappa$ can also be seen as a distribution over oracle pairs (S, U) by considering the indicator oracles for the multisets.

Note that in the distribution Strong_κ, U is always a set, even though S may be have multiplicities. Second, observe that the point $y = 0^n$ has a corresponding probability of $1 - \mathrm{e}^{\kappa \ell}/2$ which is an invariant irrespective of the instance. Therefore, querying y provides no information for a verification algorithm about the instance that it couldn't have achieved by flipping a coin.

Lemma 7.3 (The strong yes property). *For all $\kappa \in [0,1]$, $\rho \geq 0$ such that $t_1 < t_2$ in eq. (30), a pair (S,U) sampled $\sim \mathsf{Strong}_\kappa$ is a (t_1,t_2,v)-strong yes instance, except with probability at most*

$$\ell^6 2^{-n} + 2\ell^2 \exp\left(-\frac{\rho^2 2^n}{2\ell^2}\right), \tag{29}$$

where

$$t_1 = \frac{1+\kappa}{2} + \frac{v}{\ell} + \rho\,, \tag{30a}$$

$$t_2 = \frac{1+3\kappa}{2} - \frac{15\kappa^2}{4} - \frac{5\kappa}{\ell} - \rho\,. \tag{30b}$$

In particular, if $500 \leq \ell \ll 2^{n/6}$, by setting $\rho = \frac{2\ell^2}{2^n} \ln\left(\frac{2^n}{2\ell^4}\right)$ such that $2\ell^2 \exp\left(-\frac{\rho^2 2^n}{2\ell^2}\right) = \ell^6 2^{-n}$ and choosing $\kappa = 1/10$, and we have that except with probability at most $2\ell^6 2^{-n}$, (S,U) is a $(57/100, 59/100, \ell/100)$-strong yes instance[10]. Further note that for some choices of κ, ρ, for example, when $t_1 \geq 1$ and $t_2 \leq 0$, the lemma trivially holds, and when $t_1 \geq t_2$, it is not very useful. Before we give the proof of Lemma 7.3, we remark on the choice of adding y with probability a function of $\gamma_y^{(S)}$.

Remark 7.4. *It is important that the oracle U is constructed to have y added with probabilities that are functions of $\gamma_y^{(S)} = 2^n |\langle y|H^{\otimes n}|S\rangle|^2$. For certain choices of U where the probability of including y in U is a function of $\langle y|H^{\otimes n}|S\rangle$, or in particular its sign, there are ways to synthesize the state $|S\rangle$ using oracle access to U, see e.g. [INN+21].*

Proof of Lemma 7.3. For a list $S = \{s_1, \ldots, s_\ell\}$ and a positive integer k, we say an equation $s_{i_1} \oplus \ldots \oplus s_{i_{2k}} = 0^n$ is "trivial" if the index multiset $\{i_1, \ldots, i_{2k}\}$ contains each index with even multiplicity. Notice that such trivial equations hold independent of the values of s_i. We say that S is "good" if for $k = 1, 2$ or 3, the only identities $s_{i_1} \oplus \ldots \oplus s_{i_{2k}} = 0^n$ that hold are the trivial identities.

Claim 7.5. *Except with probability $\ell^6/2^n$ over the choice of random S of size ℓ, S is good.*

Proof. There are at most $\ell^6 + \ell^4 + \ell^2$ equations we need to consider. Since we only need to handle non-trivial equations, it is actually straightforward to bound the number of non-trivial equations of $2, 4, 6$ terms by the number of ways to choose 2, 4 or 6 elements from S, which we can then upper bound as $\binom{\ell}{6} + \binom{\ell}{4} + \binom{\ell}{2} \leq \ell^6$. For each non-trivial equation, over the choice of random S, the value of $s_{i_1} \oplus \cdots \oplus s_{i_{2k}}$ is just a random element in $\{0,1\}^n$. As such, the probability that the non-trivial equation holds is exactly 2^{-n}. Union-bounding over all non-trivial equations gives the claim. □

Since the probability that S is not good is very small, we will only calculate the probability that (S,U) is a strong yes instance conditioned on S being good. For a matrix M and a subset Δ, let $M_{[\Delta]}$ denote the principal submatrix of M obtained by discarding all rows and columns outside of Δ. Let $M^{S,U} = \Pi_S \cdot H^{\otimes n} \cdot \Pi_U \cdot H^{\otimes n} \cdot \Pi_S$. Notice that all the rows and columns of $M^{S,U}$ indexed by $x \notin S$ will be 0, so with a slight abuse notation, we will think of $M^{S,U}$ as the the sub-matrix of $H^{\otimes n} \cdot \Pi_U \cdot H^{\otimes n}$.

[10] We emphasize that the constants here are not particularly important; any $t_2 > t_1 + 1/\mathrm{poly}(n)$ would have sufficed.

Completeness then corresponds to showing that the largest eigenvalue of $M^{S,U}$ is $\geq t_2$. To prove this, it suffices to construct a state $|\psi\rangle$ such that $\langle\psi|M^{S,U}|\psi\rangle \geq t_2$. We use $|\psi\rangle$ equaling the uniform superposition over S. Soundness, on the other hand, corresponds to proving that $\|M^{S,U}_{[\Delta]}\| \leq t_1$ for any subset $\Delta \subseteq S$ s.t. $|\Delta| \leq v$.

Let $M^S = \mathbb{E}_U[M^{S,U}]$ where the expectation is over U sampled based on S. We will use concentration statements to relate the maximal eigenvalues of $M^{S,U}$ and its principal submatrices to those of M^S, which allows us to focus on M^S.

Claim 7.6. *For all S, $\mathbb{P}_U\left[\max_{x,x'}\left|M^S_{x,x'} - M^{S,U}_{x,x'}\right| > \rho\right] \leq 2\ell^2 \mathrm{e}^{-\rho^2 2^n/2}$.*

Proof. Fix S. For $y \in \{0,1\}^n$, let U_y denote the random variable which is 1 if $y \in U$ and 0 otherwise. For $x, x' \in S$, write

$$M^{S,U}_{x,x'} = \frac{1}{2^n}\sum_{y\in U}(-1)^{(x\oplus x')\cdot y} = \frac{1}{2^n}\sum_{y\in\{0,1\}^n}(-1)^{(x\oplus x')\cdot y}U_y\,, \tag{31a}$$

$$M^S_{x,x'} = \mathop{\mathbb{E}}_U[M^{S,U}_{x,x'}] = \frac{1}{2^n}\sum_{y\in\{0,1\}^n}(-1)^{(x\oplus x')\cdot y}\mathop{\mathbb{E}}_U[U_y]\,. \tag{31b}$$

Observe that, since S is fixed, $M^{S,U}_{x,x'}$ is the sum of 2^n independent (but not identical) random variables with each $\in \{\pm 2^{-n}\}$. The independence lets us apply Hoeffding's inequality, showing that

$$\mathop{\mathbb{P}}_u\left[\left|M^{S,U}_{x,x'} - M^S_{x,x'}\right| \geq \rho\right] \leq 2\mathrm{e}^{-\rho^2 2^n/2}\,. \tag{32}$$

Claim 7.6 follows by union-bounding over all ℓ^2 pairs $(x, x') \in S^2$. □

Claim 7.7. *For any S, except with probability at most $2\ell^2\mathrm{e}^{-\rho^2 2^n/2\ell^2}$ over the choice of U, we have that:*

$$\|M^{S,U}\| \geq \|M^S\| - \rho\,, \tag{33a}$$

$$\text{and for any subset } \Delta \subseteq S,\ \|M^{S,U}_{[\Delta]}\| \leq \|M^S_{[\Delta]}\| + \rho\,. \tag{33b}$$

Proof. We invoke Claim 7.6 with $\rho' = \rho/\ell$ to bound $|M^{S,U}_{x,x'} - M^S_{x,x'}| \leq \rho' = \rho/\ell$ for all x, x', except with probability $2\ell^2\mathrm{e}^{-\rho^2 2^n/2\ell^2}$. Under this condition, by the Gershgorin circle theorem[a], $\|M^{S,U} - M^S\|_{\mathrm{op}} \leq \ell\rho' = \rho$ and $\|M^{S,U}_{[\Delta]} - M^S_{[\Delta]}\|_{\mathrm{op}} \leq |\Delta|\rho' \leq \ell\rho' = \rho$ for any subset $\Delta \subseteq S$. Claim 7.7 then follows by the triangle inequality. □

[a]The Gershgorin circle theorem states that for a complex $W = (W_{xy})$ matrix, every eigenvalue of W lies in the union of the discs $D(W_{xx}, R_x)$ where $R_x = \sum_{y:y\neq x}|W_{xy}|$.

It remains to lower-bound $\left\|M^S\right\|_{\mathrm{op}}$ and upper-bound $\left\|M^S_{[\Delta]}\right\|_{\mathrm{op}}$. To do this, we compute approximate matrices in PSD ordering. Define

$$A^S \stackrel{\text{def}}{=} H^{\otimes n}\cdot\mathsf{Diag}\left[\left(\frac{1}{2} + \frac{\kappa}{2}\gamma^{(S)}_y - \frac{\kappa^2}{4}(\gamma^{(S)}_y)^2\right)_y\right]\cdot H^{\otimes n}\,, \tag{34a}$$

$$B^S \stackrel{\text{def}}{=} H^{\otimes n}\cdot\mathsf{Diag}\left[\left(\frac{1}{2} + \frac{\kappa}{2}\gamma^{(S)}_y\right)_y\right]\cdot H^{\otimes n}\,. \tag{34b}$$

Here, the notation $\mathsf{Diag}[(f(y))_y]$ denotes the diagonal matrix where for all $y \in \{0,1\}^n$, the (y,y)'th entry of the matrix is $f(y)$.

Claim 7.8. *For any S and any subset $\Delta \subseteq \{0,1\}^n$ (including $\Delta = \{0,1\}^n$), we have that $A^S_{[\Delta]} \preccurlyeq M^S_{[\Delta]} \preccurlyeq B^S_{[\Delta]}$, and in particular $\|A^S_{[\Delta]}\| \le \|M^S_{[\Delta]}\| \le \|B^S_{[\Delta]}\|$.*

Proof. Recall that $\mathbb{E}_U[\Pi_U] = \mathsf{Diag}[(1 - \frac{1}{2}e^{-\kappa\gamma^{(S)}_y})_y]$. We can use the small Taylor series expansions of e^{-x} to bound, for all non-negative real x,

$$\frac{1}{2} + \frac{\kappa}{2}x - \frac{\kappa^2}{4}x^2 \le 1 - \frac{1}{2}e^{-\kappa x} \le \frac{1}{2} + \frac{\kappa}{2}x \tag{35}$$

for all non-negative real x. For diagonal matrices, as PSD ordering is equivalent to the ordering of the diagonal entries:

$$\mathsf{Diag}\left[\left(\frac{1}{2} + \frac{\kappa}{2}\gamma^{(S)}_y - \frac{\kappa^2}{4}(\gamma^{(S)}_y)^2\right)_y\right] \preccurlyeq \mathsf{Diag}\left[\left(1 - \frac{1}{2}e^{-\kappa\gamma^{(S)}_y}\right)_y\right] \preccurlyeq \mathsf{Diag}\left[\left(\frac{1}{2} + \frac{\kappa}{2}\gamma^{(S)}_y\right)_y\right]. \tag{36}$$

As PSD ordering is preserved under transformations $M \mapsto C^\dagger MC$, it follows that $A^S \preccurlyeq M^S \preccurlyeq B^S$. As PSD ordering is also preserved for all principal submatrices, this proves Claim 7.8. □

Now we are ready to bound the maximal eigenvalues of (principal submatrices of) A^S and B^S and, in turn, the maximal eigenvalues of M^S. Recall that for all good S, all ℓ elements in the multiset S are distinct and so are the elements in the sumset $S \oplus S = \{x \oplus y : x, y \in S.x \ne y\}$.

Bounding the top eigenvalue of B^S. For $x \in \{0,1\}^n$, let δ_x be the indicator function for $x = 0^n$. Observe that:

$$B^S_{x,x'} = \frac{1}{2^n}\sum_y (-1)^{(x\oplus x')\cdot y}\left(\frac{1}{2} + \frac{\kappa}{2}\gamma^{(S)}_y\right) \tag{37a}$$

$$= \frac{1}{2}\left(\frac{1}{2^n}\sum_y (-1)^{(x\oplus x')\cdot y}\right) + \frac{\kappa}{2\ell}\left(\frac{1}{2^n}\sum_y (-1)^{(x\oplus x')\cdot y}\sum_{x_0,x_0'\in S}(-1)^{(x_0\oplus x_0')\cdot y}\right) \tag{37b}$$

$$= \frac{1}{2}\delta_{x\oplus x'} + \frac{\kappa}{2\ell}\sum_{x_0,x_0'\in S}\delta_{x\oplus x'\oplus x_0\oplus x_0'}. \tag{37c}$$

For diagonal entries $x = x'$, $\delta_{x\oplus x'} = 1$. Moreover, for diagonal entries $\delta_{x\oplus x'\oplus x_0\oplus x_0'} = 0$ unless $x_0 = x_0'$; there are exactly ℓ such cases since S is good. Thus, the diagonal entries all equal $B^S_{x,x} = \frac{1}{2} + \frac{\kappa}{2}$. Meanwhile, for off-diagonal entries $x \ne x'$, $\delta_{x\oplus x'} = 0$. Moreover, because S is good, $\delta_{x\oplus x'\oplus x_0\oplus x_0'} = 0$ except for the two cases $(x_0, x_0') = (x, x')$ or $(x_0', x_0) = (x, x')$. Thus, the off diagonal entries all equal $B^S_{x,x'} = \frac{\kappa}{\ell}$. Observe, then, that the matrix $B^S_{[\Delta]}$ can be more succinctly expressed as

$$B^S_{[\Delta]} = \left(\frac{1}{2} + \frac{\kappa}{2} - \frac{\kappa}{\ell}\right)\mathrm{id}_\Delta + \frac{\kappa|\Delta|}{\ell}|\Delta\rangle\langle\Delta| \tag{38}$$

where $|\Delta\rangle = \frac{1}{\sqrt{\Delta}} \sum_{x\in\Delta} |x\rangle$, the uniform superposition over Δ. $|\Delta\rangle$ is thus the top eigenvector. Therefore,

$$\left\| B^S_{[\Delta]} \right\|_{\text{op}} = \langle \Delta | B^S_{[\Delta]} | \Delta \rangle \leq \frac{1+\kappa}{2} + \frac{|\Delta|\kappa}{\ell} . \tag{39}$$

Bounding the top eigenvalue of A^S. This term is slightly more complicated due to the quadratic term in $\gamma^{(S)}_y$. We have:

$$A^S_{x,x'} = \frac{1}{2^n} \sum_y (-1)^{(x\oplus x')\cdot y} \left(\frac{1}{2} + \frac{\kappa}{2} \gamma^{(S)}_y - \frac{\kappa^2}{4} (\gamma^{(S)}_y)^2 \right) \tag{40a}$$

$$= B^S_{x,x'} - \frac{\kappa^2}{4\ell^2} \left(\frac{1}{2^n} \sum_y \sum_{x_0,x_1,x'_0,x'_1} (-1)^{(x\oplus x')\cdot y} (-1)^{(x_0\oplus x'_0)\cdot y} (-1)^{(x_1\oplus x'_1)\cdot y} \right) \tag{40b}$$

$$= B^S_{x,x'} - \frac{\kappa^2}{4\ell^2} \sum_{x_0,x_1,x'_0,x'_1} \delta_{x\oplus x'\oplus x_0\oplus x'_0\oplus x_1\oplus x'_1} . \tag{40c}$$

We have already evaluated $B^S_{x,x'}$, and instead focus on evaluating the final term. For diagonal entries $x = x'$, since S is good, there are only two ways for $\delta_{x\oplus x'\oplus x_0\oplus x'_0\oplus x_1\oplus x'_1} = \delta_{x_0\oplus x'_0\oplus x_1\oplus x'_1}$ to be non-zero:

- $x_0 = x'_0$ and $x_1 = x'_1$. There are ℓ^2 such terms.
- $x_0 \neq x'_0$, and either $(x_1, x'_1) = (x_0, x'_0)$ or $(x'_1, x_1) = (x_0, x'_0)$. There are $\ell \times (\ell - 1) \times 2$ such terms.

This gives a total of $3\ell^2 - 2\ell$ such terms. As a consequence, we have

$$A^S_{x,x} = B^S_{x,x} - \frac{\kappa^2}{4\ell^2}(3\ell^2 - 2\ell) = \frac{1}{2} + \frac{\kappa}{2} - \frac{3\kappa^2}{4} + \frac{\kappa^2}{2\ell} . \tag{41}$$

For off-diagonal entries $x \neq x'$, the only way for $x \oplus x' \oplus x_0 \oplus x'_0 \oplus x_1 \oplus x'_1$ to be 0 is for one of x_0, x'_0, x_1, x'_1 to be x, another to be x', and the remaining two must be equal. This again follows from the goodness of S. There are 4×3 ways of choosing which of the four elements are equal to x and x', and ℓ ways to choose the remaining pair. This slightly over-counts, since for example $(x, x', x_0, x'_0, x_1, x'_1) = (x, x', x, x, x, x')$ would be counted 3 times, for $x_0 = x$, $x'_0 = x$ and $x'_1 = x$. There are 8 terms of this form (4 where there is one x' and 3 x among (x_0, x'_0, x_1, x'_1), and 4 where there is one x and 3 x'). Each term of this form is over-counted 2 extra times (for a total of 3 times). The correct number of terms with $x \oplus x' \oplus x_0 \oplus x'_0 \oplus x_1 \oplus x'_1 = 0$ is therefore: $12\ell - 16$. This means the off-diagonal entries are equal to:

$$A^S_{x,x'} = B^S_{x,x'} - \frac{\kappa^2}{4\ell^2}(12\ell - 16) = \frac{\kappa}{\ell} - \frac{3\kappa^2}{\ell} + \frac{4\kappa^2}{\ell^2} . \tag{42}$$

Similar to the calculation for B^S_Δ,

$$A^S = \left[\frac{1}{2} + \kappa\left(\frac{1}{2} - \frac{1}{\ell} \right) + \kappa^2 \left(-\frac{3}{4} + \frac{7}{2\ell} - \frac{4}{\ell^2} \right) \right] \text{id}_S + \left[\kappa - 3\kappa^2 + \frac{4\kappa^2}{\ell} \right] |S\rangle\langle S| . \tag{43}$$

Therefore, the top eigenvector is indeed $|S\rangle$ and

$$\left\| A^S \right\|_{\text{op}} = \langle S | A^S | S \rangle \tag{44a}$$

$$= \frac{1}{2} + \kappa\left(\frac{3}{2} - \frac{1}{\ell}\right) + \kappa^2\left(-\frac{15}{4} + \frac{15}{2\ell} - \frac{4}{\ell^2}\right) \tag{44b}$$

$$\geq \frac{1}{2} + \frac{3\kappa}{2} - \frac{15\kappa^2}{4} - \frac{5\kappa}{\ell}\,. \tag{44c}$$

Combining the previous Hoeffding's inequality with these bounds on the spectral norm completes the proof. □

Part III
Sampling probability upper bound for quasi-even condensates

The proof components in Part II combine to prove the existence of a QCMA algorithm for spectral Forrelation implies sampling probability *lower bound*. More specifically, Section §6 showed how to transform a QCMA algorithm for deciding between yes (at least 59/100-spectrally Forrelated) and no (at most 57/100-spectrally Forrelated) instances into a successful sampler that samples points from strong yes instances, and Lemma 7.3 in Section §7 constructed a distribution Strong_κ which is, with overwhelming probability, supported on strong yes instances. For the rest of the paper, we will choose $\kappa = 1/10$ and use $\mathsf{Strong} = \mathsf{Strong}_{1/10}$ for brevity.

Together, we have proven that if there exists a QCMA decision algorithm, then there exists a lower bound on the success probability of a sampler at guessing points from S given query access to U when (S, U) is sampled according to the definition of Lemma 7.3. The remainder of the proof is to marry this probability lower bound with a contradictory probability upper bound. In this section, we will use compressed oracle/purification techniques defined in terms of bosons. The main theorem for Part III is Theorem 9.1, and it will be stated after we introduce the bosonic framework in Section §8.

Before starting, we would like to emphasize that the proofs and bounds we generate are almost certainly far from tight. We will loosely bound terms for legibility, and we acknowledge that the resulting bounds will not be tight. Nevertheless, they will be sufficient to contradict Theorem 6.2.

8 Quantum mechanics of bosons

As stated in the Introduction, our proof will require understanding purifications of quantum oracles as bosonic systems. This section will provided the facts and definitions needed to understand the rest of the proof. For the purposes of this note, it suffices to consider a bosonic system with 2^n modes with the modes indexed by n-bit vectors. For notational convenience, it is also useful to be able to index the modes with $\{0, \dots, 2^n - 1\}$ under the standard isometry. In particular we refer to the 0^n-mode as the 0-mode.

8.1 A natural basis

The bosonic (position) Fock basis is a collection of *orthonormal* states of the form $|\ell_0, \ell_1, \dots, \ell_{2^n-1}\rangle$ with each $\ell_x \in \mathbb{N}$. This is the basis state corresponding to ℓ_0 bosons in the 0-mode, ℓ_1 bosons in the 1-mode, etc. The total number of bosons is $\sum_x \ell_x$.

We will exclusively consider states with ℓ total bosons; however, it is mathematically helpful to be able to add and remove bosons at will to describe the transformation from one state of ℓ bosons to another.

8.2 The second quantization

Let $|\text{vac}\rangle \stackrel{\text{def}}{=} |0, \dots, 0\rangle$ be the vacuum state representing 0 bosons in the system. Abiding by the traditional notation from physics, let $\widehat{a}_x$ and $\widehat{a}_x^\dagger$ be the annihilation and creation operators for a boson in the x-th mode, respectively. These are defined by their action on the position Fock basis as follows

$$\widehat{a}_x \, |\ell_0, \dots, \ell_x, \dots, \ell_{2^n-1}\rangle = \sqrt{\ell_x} \, |\ell_0, \dots, \ell_x - 1, \dots, \ell_{2^n-1}\rangle \quad \text{and,} \tag{45a}$$

$$\widehat{a}_x^\dagger \, |\ell_0, \dots, \ell_x, \dots, \ell_{2^n-1}\rangle = \sqrt{\ell_x + 1} \, |\ell_0, \dots, \ell_x + 1, \dots, \ell_{2^n-1}\rangle \, . \tag{45b}$$

We will refer to these as the annihilation and creation operators for bosons in the *position* basis. Note that these operators are *not* unitary. Very roughly, the factor $\sqrt{\ell_x}$ corresponds to the fact that the bosons are indistinguishable and so we do not know which of the ℓ_x bosons were annihilated. Likewise, for creation. In calculations about bosons, it is useful to oscillate between the Fock representation and the creation/annihilation perspective. In physics, the two representations are referred to as first and second quantizations, respectively.

It follows from the definition of the position creation operator that

$$\frac{1}{\sqrt{\prod_{x=0}^{2^n-1} \ell_x!}} \prod_{x=0}^{2^n-1} \left(\widehat{a}_x^\dagger\right)^{\ell_x} |\text{vac}\rangle = |\ell_0, \dots, \dots, \ell_{2^n-1}\rangle \, . \tag{46}$$

The commutation relations for bosonic position operators are given by

$$[\widehat{a}_x, \widehat{a}_y^\dagger] = \widehat{a}_x \widehat{a}_y^\dagger - \widehat{a}_y^\dagger \widehat{a}_x = \delta_{xy} \, , \qquad [\widehat{a}_x, \widehat{a}_y] = [\widehat{a}_x^\dagger, \widehat{a}_y^\dagger] = 0 \, . \tag{47}$$

One may expect all annihilation and creation operators to commute since, for example, annihilating a boson at x seems independent of creating a boson at y. This is true generally, except for annihilating and creating at the same mode x. Intuitively, the reason for a lack of commutation is that if there are zero bosons at mode x, annihilation actually does nothing since there is nothing to annihilate. This causes an asymmetry since annihilation-before-creation may have nothing to annihilate, but annihilation-after-creation will always have something to annihilate.

8.3 A momentum basis

We can also define the annihilation and creation operators in the *momentum* basis by the Hadamard transform. Note, that this is our *computer science* interpretation. Usually, the transform from position to momentum basis is given by the quantum Fourier transform over the group $\mathbb{Z}_{2^n}$.

$$\widetilde{a}_y \stackrel{\text{def}}{=} \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} (-1)^{x \cdot y} \widehat{a}_x \, , \tag{48a}$$

$$\widetilde{a}_y^\dagger \stackrel{\text{def}}{=} \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} (-1)^{-x \cdot y} \widehat{a}_x^\dagger \, . \tag{48b}$$

The commutation relations for momentum operators can be derived to be the analogs of eq. (47):

$$[\widetilde{a}_x, \widetilde{a}_y^\dagger] = \widetilde{a}_x \widetilde{a}_y^\dagger - \widetilde{a}_y^\dagger \widetilde{a}_x = \delta_{xy} \, , \qquad [\widetilde{a}_x, \widetilde{a}_y] = [\widetilde{a}_x^\dagger, \widetilde{a}_y^\dagger] = 0 \, . \tag{49}$$

Having defined momentum annihilation and creation operators, we can derive that there exists a second Fock basis that can be used to describe states. This is the momentum Fock basis, and it is a collection of orthonormal states $|\ell_0, \ell_1, \dots, \ell_{2^n-1}\rangle$ where the integers ℓ_x describe how many bosons are in each *momentum* mode. It is the exact analog of the position Fock basis.

8.4 Number operators

Additionally, we can define the position and momentum number operators as $\widehat{n}_x \stackrel{\text{def}}{=} \widehat{a}_x^\dagger \widehat{a}_x$ and $\widetilde{n}_x \stackrel{\text{def}}{=} \widetilde{a}_x^\dagger \widetilde{a}_x$, respectively. These are diagonal matrices in the position and momentum Fock bases that multiply a Fock basis state by the *number* of bosons in the x'th mode — i.e., $\widehat{n}_0 |\ell_0, \dots\rangle = \ell_0 |\ell_0, \dots\rangle$, hence the name. These have the following commutation relations with the creation and annihilation operators.

$$\widehat{n}_x \widehat{a}_x = \widehat{a}_x^\dagger \widehat{a}_x \widehat{a}_x = (\widehat{a}_x \widehat{a}_x^\dagger - 1)\widehat{a}_x = \widehat{a}_x(\widehat{a}_x^\dagger \widehat{a}_x - 1) = \widehat{a}_x(\widehat{n}_x - 1)\,, \tag{50a}$$

$$\widehat{n}_x \widehat{a}_x^\dagger = \widehat{a}_x^\dagger \widehat{a}_x \widehat{a}_x^\dagger = \widehat{a}_x^\dagger(1 + \widehat{a}_x^\dagger \widehat{a}_x) = \widehat{a}_x^\dagger(1 + \widehat{n}_x)\,. \tag{50b}$$

Intuitively, the number of bosons *after* annihilating is just one less than the number *before* annihilating (and vice-versa for creating). We also define the total number operator $\widehat{N} = \sum_x \widehat{n}_x$. One can verify that $\widehat{N} = \widetilde{N}$.

8.5 A random bosonic setup

A typical problem in (classical) combinatorics might start with "place ℓ indistinguishable balls uniformly randomly into N boxes". The quantum mechanical interpretation is to uniformly place ℓ bosons into $N = 2^n$ modes. One exact purification of this setup is to consider the state with ℓ bosons in the 0-momentum mode, which equals

$$\frac{1}{\sqrt{\ell!}} \left(\widetilde{a}_0^\dagger\right)^\ell |\text{vac}\rangle = \frac{1}{\sqrt{\ell! \cdot 2^{n\ell}}} \sum_{x_1,\dots,x_\ell} \left(\prod_{i=1}^{\ell} \widehat{a}_{x_i}^\dagger\right) |\text{vac}\rangle \tag{51}$$

with the right side demonstrably equivalent to creating ℓ bosons uniformly randomly into the vacuum. Observe that the $\sqrt{\ell!}$ additional normalization term might seem strange initially; however, it is indeed necessary. Indeed, this is a consequence of the bosons being indistinguishable, and therefore not "knowing" the order in which they are created. This results in overcounting, which is fixed by dividing by $\sqrt{\ell!}$. Furthermore, suppose we measure this state in the position Fock basis and interpret the measurement as a multiset. Note that the probability of measuring a particular multiset S, which contains element x, ℓ_x many times, when measuring eq. (51) can be calculated as

$$\mathbb{P}[S] = \frac{\ell!}{2^{n\ell} \prod_x \ell_x!}\,, \tag{52}$$

which is exactly the uniform distribution over multisets of size ℓ.

8.6 Bosonic Hilbert space

Bosonic systems on 2^n modes are states in an infinite-dimensional Hilbert space. As there is no bound on the number of bosons in a bosonic system, this Hilbert space is infinite-dimensional. However, the

Hilbert space can be expressed as the direct sum of finite-dimensional Hilbert spaces by restricting to a fixed number of bosons:

$$\mathcal{H}_{\text{boson}} = \bigoplus_{\ell=0}^{\infty} \mathcal{H}_{\text{boson}}^{(\ell)} \text{ where } \mathcal{H}_{\text{boson}}^{(\ell)} = \text{span of Fock states of } \ell \text{ bosons}. \tag{53}$$

Since $\widehat{N} = \widetilde{N}$, the number of bosons in the position and momentum bases is the same; therefore, the transformation mapping between the position and momentum Fock bases (i.e., the Hadamard transform) is block-diagonal with respect to this decomposition.

In this work, we will restrict ourselves to considering bosonic systems with a fixed number, ℓ, of bosons. Therefore, the states are in the space $\mathcal{H}_{\text{boson}}^{(\ell)}$. Both the position and momentum Fock bases of this Hilbert space can be indexed by non-negative integer tuples of length 2^n with total sum ℓ. Observe that this is isomorphic to the set of truth-tables of multisets $S \subseteq \{0,1\}^n$ of size ℓ. In the rest of the paper, we identify the states $|\text{tt}_S\rangle$ with the position Fock basis state with bosons in locations given by the elements of S.

9 Sampler upper bound statement and organization

This section will state the main theorem of Part III, and give some intuition for how the proof will go, and the organization of the rest of the part. The goal of Part III is to prove a probability *upper bound* on the success of any polynomial query sampler when run on the Strong distribution defined in Definition 7.2. This is the counterpart to Theorem 6.2, which proves a contradictory lower bound on the success probability, assuming a QCMA algorithm for spectral Forrelation exists.

9.1 Theorem statement

The main theorem of this part is the following.

Theorem 9.1 (Sampling probability upper bound)**.** *For all v, for all quantum algorithms $\mathcal{A}^U$ accessing an oracle U and outputting v distinct outputs, while making t queries per output, if a pair of oracles (S, U) are sampled according to distribution* Strong *(defined in Definition 7.2), then the probability that all v of the outputs of $\mathcal{A}^U$ are elements of S is at most*

$$\leq 2\left(\frac{4v((vt)^{30} + v(vt)^{20})\sqrt{\ell}}{2^{n/4}}\right)^v + \left(\left(\frac{(vt)^4}{\ell^{1/32}}\right)^v + e^{-5vt}\right)^2. \tag{54}$$

In this theorem, the exponents are not important and are the consequence of loose bounding for legibility. What matters is that for $v, t = \text{poly}(n)$ and $\ell = 2^{cn}$, the numerators are significantly smaller than the denominators. Therefore, as v grows, this quantity decreases exponentially fast. We will actually prove a slightly stronger statement, namely we will show a bound for any algorithm that first makes T queries to the oracle U and then outputs v guesses. For $T = vt$, this is a strictly larger class of algorithms than those that make t queries per guess. The rationale for studying this stronger model is that it natively handles the complexities of the memory of the algorithm between guesses.

9.2 Proof overview and intuition

As suggested in the Introduction, we will heavily use the bosonic framework to analyze the success probability of a sampler algorithm. The bosonic framework will be used to analyze a uniform superposition over the pairs (S, U) sampled from the Strong distribution defined in Definition 7.2. By considering a uniform superposition over pairs (S, U), we can conveniently analyze the average-case success probability of the sampler.

Our proof strategy for proving Theorem 9.1 is reminiscent of ideas introduced by Hamoudi and Magniez [HM23]. We will define a family of subspaces $\{\mathsf{QEC}_{(r,o)}\}$ indexed by positive integers r and o, with the property that if the state after T queries is mostly contained in $\mathsf{QEC}_{(r,o)}$ for $r \leq \mathrm{poly}(n)$ and $o \leq v/4$, then the success probability bound for the guessing algorithm is enough to prove Theorem 9.1. Once we have established that states supported on the subspace $\mathsf{QEC}_{(r,o)}$ have a low sampling success probability, we will prove that the state of all T-query algorithms querying the purified (S, U) oracles is supported almost entirely on the $\mathsf{QEC}_{(r,o)}$ subspace for some r that is polynomial in T. Putting these together we will have a sampling probability upper bound.

The notation $\mathsf{QEC}_{(r,o)}$ refers to something we call a (r, o)-*quasi-even condensate*. Informally, a (r, o)-quasi-even condensate is a state in the span of momentum Fock states where the number of *odd* number operators is $\leq o$ and the number of momentum modes that are non-zero is at most r. See the following remark for more details about this choice of naming.

Remark 9.2 (Nomenclature)**.** *We borrow the term condensate from many-body physics, where it denotes a regime in which a macroscopic fraction of bosons occupy a single-particle mode—typically the zero-momentum mode—giving rise to collective, mean-field–like behavior. In our setting, we use condensate analogously to indicate that most bosons remain in the zero-momentum mode, with only a small number of excitations occupying other modes. The modifier quasi-even reflects that the occupations across modes are almost parity-symmetric: only a few modes have odd occupation numbers.*

Hence, a quasi-even condensate refers to a subspace of Fock states that are both condensate-like (dominated by the zero-momentum mode) and nearly even under mode-parity. This terminology is not meant in a thermodynamic sense, but rather as a descriptive analogy capturing the structure of states that remain close to an even-parity condensate configuration with a few parity defects.

9.3 Organization

We begin by proving a sampling probability upper-bound for quasi-even condensates in Section §10. Then, in Section §11 we prove how to represent the state of an algorithm that has made a few queries to the oracle U when run in superposition over all pairs $(S, U) \sim$ Strong (defined in Definition 7.2). Next, in Part IV we prove that the state of an algorithm that has made polynomial in n many queries is a quasi-even condensate. This proves that we can apply the previously derived sampling probability upper-bound and completes the proof.

10 Sampler upper bounds for quasi-even condensates

This section will define states that are quasi-even condensates and prove an upper bound on the sampling success probability of all algorithms whose states are supported on quasi-even condensates. In future sections, we prove that the post-query state of all polynomial query samplers is close to a quasi-even condensate.

Recall that the larger goal is to prove an upper bound on the success probability of a query algorithm on a random instance from Strong (defined in Definition 7.2). We can imagine that the query algorithm is split into two steps: (a) the first is the T queries of interaction with the oracle U and (b) the measurement with respect to S of the guesses. Interaction with a $(S, U) \sim \mathsf{Strong}$ can be studied by first purifying the distribution over oracles and interacting coherently with each pair (S, U) in superposition. By purifying, we mean that queries to the oracle U are replaced by the linear extension of the unitary

$$|b, x, y\rangle |\mathrm{tt}_S\rangle |\mathrm{tt}_U\rangle \mapsto (-1)^{b \cdot U(y)} |b, x, y\rangle |\mathrm{tt}_S\rangle |\mathrm{tt}_U\rangle \tag{55}$$

where $\mathrm{tt}_{(\cdot)}$ is our notation for the truth table of the corresponding oracle. Checking the v guesses output by the algorithm is equivalent to measuring the final (entangled) state of the algorithm and oracle with respect to

$$\Pi_{\mathrm{succ}} \stackrel{\mathrm{def}}{=} \sum_{\substack{z_1,\ldots,z_v \in (\{0,1\}^n)^v \\ \text{distinct}}} |z_1, \ldots, z_v\rangle\langle z_1, \ldots, z_v| \otimes \left(\sum_{S\,:\, z_1,\ldots,z_v \in S} |\mathrm{tt}_S\rangle\langle \mathrm{tt}_S| \right). \tag{56}$$

The success probability of this sampler over the distribution of random oracles (S, U) is equal to the probability that measuring Π_{succ} on post-query state of the sampler, when run on the purification of the oracles, accepts. We now define (r, o)-quasi-even condensates.

Definition 10.1 (Quasi-even condensates)**.** *Let $u = (u_x)_{x \in \{0,1\}^n}$ be a tuple of non-negative integers such that $\sum_x u_x = \ell$, representing a momentum Fock state of ℓ bosons. Then we say that u describes an* (r, o)-quasi-even condensate *if*

1. *(Condensate) $u_0 \geq \ell - r$. I.e., most of the bosons are in the 0-mode.*
2. *(Quasi-even) At most o many u_x, except for u_0, are odd.*

We define (r, o)-quasi-even condensates *to be any state in the span of momentum Fock states corresponding to quasi-even condensate tuples, or* $\mathrm{span}\{\, |u\rangle \,:\, u \text{ is } (r, o)\text{-quasi-even condensate} \,\}$*. We further define $\mathsf{QEC}_{(r,o)}$ to be the projector onto this subspace. Additionally, we define projectors Con_r and QE_o as the projectors onto r-condensates and o-quasi-even states, respectively. We also define $\mathsf{QE}_{=o}$ and $\mathsf{QE}_{\geq o}$ to be the projectors onto states with exactly o odd u_x's (excluding u_0) and $\geq o$ many odd u_x's (excluding u_0). All of these projectors are diagonal in the momentum Fock basis and therefore commute. By definition,*

$$\mathsf{QEC}_{r,o} = \mathsf{Con}_r \cdot \mathsf{QE}_o = \mathsf{QE}_o \cdot \mathsf{Con}_r. \tag{57}$$

Note that all of the previously defined projectors require a state of exactly ℓ bosons. Within this subspace, Con_0 is the projector onto the state that has all ℓ bosons in the 0-momentum mode, i.e., $|\ell, \ldots\rangle$, and Con_ℓ is the projector onto all states with ℓ bosons. The following theorem is the main result of this section, which shows a sampling probability upper bound for $(r, v/4)$-quasi-even condensates when $r \ll \ell$.

Theorem 10.2 (Quasi-even condensate probability upper bound)**.** *Let* Π_{succ} *be the previously defined success operator and* $\mathsf{QEC}_{(r,v/4)}$ *be the projector onto* $(r, v/4)$*-quasi-even condensates, which only acts on the* S *register. Then,*

$$\left\|\mathsf{QEC}_{(r,v/4)} \cdot \Pi_{\text{succ}} \cdot \mathsf{QEC}_{(r,v/4)}\right\|_{\text{op}} \leq 2\left(\frac{4v(r^3 + vr^2)\sqrt{\ell}}{2^{n/4}}\right)^v . \tag{58}$$

The key lemma required to prove Theorem 10.2 is the following. It provides an upper bound on the success probability of a quasi-even condensate that makes distinct guesses $z_1, \ldots, z_v$. A priori, Lemma 10.3 may seem strange as it calculates the maximum eigenvalue of a product of number operators on the space of quasi-even condensates. However, we will prove that this is sufficient for proving an upper bound on the maximum eigenvalue of Π_{succ} within the space of quasi-even condensates. This proof will rely on the fact that the upper bound proven in Lemma 10.3 is independent of the choice of guess locations.

Lemma 10.3. *For distinct coordinates* $z_1, \ldots, z_v \in \{0,1\}^n$*,*

$$\left\|\mathsf{QEC}_{(r,v/4)} \cdot \widehat{n}_{z_1} \ldots \widehat{n}_{z_v} \cdot \mathsf{QEC}_{(r,v/4)}\right\|_{\text{op}} \leq 2\left(\frac{4v(r^3 + vr^2)\sqrt{\ell}}{2^{n/4}}\right)^v . \tag{59}$$

We first prove that Lemma 10.3 implies Theorem 10.2 and then finish this section with the proof of Lemma 10.3.

Proof of Theorem 10.2. Observe that we can reformulate Π_{succ} as the following:

$$\Pi_{\text{succ}} = \sum_{\substack{z_1,\ldots,z_v \in (\{0,1\}^n)^v \\ \text{distinct}}} |z_1, \ldots, z_v\rangle\langle z_1, \ldots, z_v| \otimes \Pi_{z_1,\ldots,z_v} . \tag{60}$$

where define $\Pi_{z_1,\ldots,z_v}$ to be the projection onto states $|\text{tt}_S\rangle$ that have at least one boson in position modes $z_1, \ldots z_v$. Next, we note that for distinct $z_1, \ldots, z_v$, the annihilation operators commute and we have that

$$\left(\widehat{a}^\dagger_{z_1} \ldots \widehat{a}^\dagger_{z_v} \widehat{a}_{z_1} \ldots \widehat{a}_{z_v}\right) = \widehat{n}_{z_1} \ldots \widehat{n}_{z_v} \succcurlyeq \Pi_{z_1,\ldots,z_v} . \tag{61}$$

This statement is conceptually equivalent to applying Markov's inequality: that $\mathbb{P}[X > 0] \leq \mathbb{E}[X]$ for non-negative random variable X. Therefore, it follows that

$$\left\|\mathsf{QEC}_{(r,v/4)} \cdot \Pi_{\text{succ}} \cdot \mathsf{QEC}_{(r,v/4)}\right\|_{\text{op}} \leq \left\|\mathsf{QEC}_{(r,v/4)} \cdot \Lambda_{\text{succ}} \cdot \mathsf{QEC}_{(r,v/4)}\right\|_{\text{op}} \tag{62a}$$

$$\text{where } \Lambda_{\text{succ}} \stackrel{\text{def}}{=} \sum_{\substack{z_1,\ldots,z_v \\ \text{distinct}}} |z_1, \ldots, z_v\rangle\langle z_1, \ldots, z_v| \otimes \left(\widehat{a}^\dagger_{z_1} \ldots \widehat{a}^\dagger_{z_v} \widehat{a}_{z_1} \ldots \widehat{a}_{z_v}\right) . \tag{62b}$$

Now we will prove the following bound, applying Lemma 10.3.

$$\left\|\mathsf{QEC}_{(r,v/4)} \cdot \Lambda_{\text{succ}} \cdot \mathsf{QEC}_{(r,v/4)}\right\|_{\text{op}} \leq 2\left(\frac{4v(r^3 + vr^2)\sqrt{\ell}}{2^{n/4}}\right)^v . \tag{63}$$

Consider any state $|\varphi\rangle$ supported on $\mathsf{QEC}_{(r,v/4)}$ and write it in its decomposition based on guesses:

$$|\varphi\rangle = \sum_{z_1,\ldots,z_v} \alpha_{z_1,\ldots,z_v} |z_1, \ldots, z_v\rangle \otimes |\varphi_{z_1,\ldots,z_v}\rangle \tag{64}$$

where $|\varphi_{z_1,\ldots,z_v}\rangle$ is the remainder of the state (normalized). Then,

$$\langle\varphi|\Lambda_{\text{succ}}|\varphi\rangle = \sum_{\substack{z_1,\ldots,z_v \\ \text{distinct}}} \left|\alpha_{z_1,\ldots,z_v}\right|^2 \cdot \left\langle \varphi_{z_1,\ldots,z_v}\middle|\widehat{n}_{z_1}\ldots\widehat{n}_{z_v}\middle|\varphi_{z_1,\ldots,z_v}\right\rangle \tag{65a}$$

$$\leq \max_{\substack{z_1,\ldots,z_v \\ \text{distinct}}} \left\langle \varphi_{z_1,\ldots,z_v}\middle|\widehat{n}_{z_1}\ldots\widehat{n}_{z_v}\middle|\varphi_{z_1,\ldots,z_v}\right\rangle \tag{65b}$$

$$\leq 2\left(\frac{4v(r^3+vr^2)\sqrt{\ell}}{2^{n/4}}\right)^v. \tag{65c}$$

Here, the final line is the application of Lemma 10.3. □

It only remains to prove Lemma 10.3, which we prove next.

Proof of Lemma 10.3. Recall that an upper bound on the spectral norm of a Hermitian matrix is the max 1-norm over all rows. Our goal is to study $\widehat{n}_{z_1}\ldots\widehat{n}_{z_v}$ within the space defined by quasi-even condensates. So, we restrict our attention to the row indexed by u in the momentum Fock basis where u is a (r,o)-quasi-even condensate tuple. Therefore, the goal is to bound

$$\sum_{(r,o)-\text{QEC tuple } w} \left|\langle w|\widehat{n}_{z_1}\ldots\widehat{n}_{z_v}|u\rangle\right| \tag{66a}$$

$$= \sum_{d\geq 0}\left(\sum_{\substack{(r,o)-\text{QEC tuple } w \\ |w-u|=2d}} \left|\langle w|\widehat{n}_{z_1}\ldots\widehat{n}_{z_v}|u\rangle\right|\right) \tag{66b}$$

$$\leq \sum_{d\geq 0}\left(\max_{\substack{(r,o)-\text{QEC tuple } w \\ |w-u|=2d}} \left|\langle w|n_{z_1}\ldots\widehat{n}_{z_v}|u\rangle\right|\right)\cdot \#\left\{w : \substack{(r,o)-\text{QEC tuple } w \\ |w-u|=2d}\right\}. \tag{66c}$$

As the previous equation suggests, we will bound the terms in the previous equation using two additional claims. The first is a bound on the entry of the matrix corresponding to a tuple w, at distance $2d$ from a fixed tuple u.

Claim 10.4. *Fix two (r,o)-quasi-even condensate tuples u, w such that $|u-w| = 2d$ with $\ell \geq 2d$. Then the following holds for all distinct $z_1,\ldots,z_v$.*

$$\left|\langle u|\widehat{n}_{z_1}\ldots\widehat{n}_{z_v}|w\rangle\right| \leq \begin{cases} v!(2r)^{v+d/2}\dfrac{\ell^{v-d/2}}{2^{nv}} & \text{if } d\leq v, \\ 0 & \text{if } d>v. \end{cases} \tag{67}$$

The proof of this claim is deferred to after the proof of Lemma 10.3. Our second claim is an upper bound on the number of quasi-even condensates w at distance $2d$ from our initial quasi-even condensate u for all values of d.

Claim 10.5. *For every (r,o)-quasi-even condensate tuple u and $d \geq 0$, the number of (r,o)-quasi-even condensate tuples w that are at exactly distance $2d$ from u is upper bounded by*

$$(2^{n+1})^{d/2+o}(r+d/2+o)^{d/2+o}. \tag{68}$$

The proof of this claim is also deferred to after the proof of Lemma 10.3. To finish the proof, we return to eq. (66) to bound $\sum_{(r,o)\text{–QEC tuple } w} \left|\langle w|\widehat{n}_{z_1}\dots\widehat{n}_{z_v}|u\rangle\right|$. We explain the derivation of each subequation after the statement.

$$\sum_{(r,o)\text{–QEC tuple } w} \left|\langle w|\widehat{n}_{z_1}\dots\widehat{n}_{z_v}|u\rangle\right| \tag{69a}$$

$$\leq \sum_{d\geq 0}\left(\max_{\substack{(r,o)\text{–QEC tuple } w\\ |w-u|=2d}}\left|\langle w|\widehat{n}_{z_1}\dots\widehat{n}_{z_v}|u\rangle\right|\cdot \#\left\{w : \substack{(r,o)\text{–QEC tuple } w\\ |w-u|=2d}\right\}\right) \tag{69b}$$

$$\leq v!(2r)^v\sum_{d=0}^{v}(2r)^{d/2}\frac{\ell^{v-d/2}}{(2^n)^v}\left((2^{n+1})^{d/2+o}(r+1+d/2+o)^{d/2+o}\right) \tag{69c}$$

$$= v!(2r)^{3v/2}\left(\frac{\ell}{2^n}\right)^v(2^n)^o\sum_{d=0}^{v}\left(\sqrt{\frac{2^{n+1}}{\ell}}\right)^d(r+1+d/2+o)^{d/2+o} \tag{69d}$$

$$\leq v!(2r)^{3v/2}\left(\frac{\ell}{2^n}\right)^v(2^n)^o\sum_{d=0}^{v}\left(\sqrt{\frac{2^{n+1}}{\ell}}\right)^d(r+1+v/2+o)^{d/2+o} \tag{69e}$$

$$= v!(2r)^{3v/2}\left(\frac{\ell}{2^n}\right)^v(2^n)^o(r+v/2+o)^o\sum_{d=0}^{v}\left(\sqrt{\frac{2^{n+1}}{\ell}}\right)^d(r+1+v/2+o)^{d/2} \tag{69f}$$

$$\leq 2v!(2r)^{3v/2}\frac{\ell^v}{(2^n)^{v-o}}\left(\sqrt{\frac{2^{n+1}}{\ell}}\right)^v(r+1+v/2+o)^{v/2+o} \tag{69g}$$

$$\leq 2^{2v+1}v!(r^{3v/2})(r+1+v/2+o)^{v/2+o}\frac{\ell^{v/2}}{(2^n)^{v/2-o}}\,. \tag{69h}$$

Here, to derive eq. (69c), we use the bounds in Claim 10.4 and Claim 10.5. Eq. (69d) factors out the terms that do not depend on d, eq. (69e) uses that $d \leq v$, and eq. (69f) again factors out terms that do not depend on d. Then we observe that the sum in eq. (69f) is a geometric series; since the summand

$$\frac{\sqrt{2^{n+1}}}{\ell}(r+1+v/2+o) \geq 2 \tag{70}$$

and, therefore, the series sums to at most twice the largest term in the series (i.e., when $d = v$), giving eq. (69g). Re-arranging terms again gives us the bound in eq. (69h) Setting $o = v/4$, we finish the proof, as we have a bound of

$$\left\|\mathsf{QEC}_{(r,o)}\cdot\widehat{n}_{z_1}\dots\widehat{n}_{z_v}\cdot\mathsf{QEC}_{(r,o)}\right\|_{\text{op}} \leq \max_{(r,o)\text{–QEC tuple } u}\sum_{(r,o)\text{–QEC tuple } w}\left|\langle w|\widehat{n}_{z_1}\dots\widehat{n}_{z_v}|u\rangle\right| \tag{71a}$$

$$\leq \frac{2^{2v+1}v!(r^{3v/2})(r+3v/4+1)^{3v/4}\ell^{v/2}}{(2^n)^{v/4}} \tag{71b}$$

$$\leq 2\left(\frac{4v(r^3+vr^2)\sqrt{\ell}}{2^{n/4}}\right)^v\,. \tag{71c}$$

Here, we use the crude upper bounds $r^{3/2} \le r^2$ and $3v/4 + 1 \le v$ to remove the constants in the expression. □

Now we prove the additional claims.

Claim (Claim 10.4 restated). *Fix two (r, o)-quasi-even condensate tuples u, w such that $|u - w| = 2d$ with $\ell \ge 2d$. Then the following holds for all distinct $z_1, \dots, z_v$.*

$$\left|\langle u|\widehat{n}_{z_1} \dots \widehat{n}_{z_v}|w\rangle\right| \le \begin{cases} v!(2r)^{v+d/2}\dfrac{\ell^{v-d/2}}{2^{nv}} & \text{if } d \le v, \\ 0 & \text{if } d > v. \end{cases} \tag{72}$$

Proof. Recall that $\widehat{a}_{z_i} = \frac{1}{\sqrt{2^n}}\sum_w (-1)^{w\cdot z_i}\widetilde{a}_w$. Then we can write

$$\begin{aligned} |\langle u|\widehat{n}_{z_1} \dots \widehat{n}_{z_v}|w\rangle| &= |\langle u|\widehat{a}^\dagger_{z_1} \dots \widehat{a}^\dagger_{z_v}\widehat{a}_{z_1} \dots \widehat{a}_{z_v}|w\rangle| && (73a)\\ &= \left|\frac{1}{2^{nv}}\sum_{\substack{\alpha_1,\dots,\alpha_v\\ \beta_1,\dots,\beta_v}}\left(\prod_{i=1}^{v}(-1)^{z_i\cdot(\alpha_i\oplus\beta_i)}\right)\langle u|\widetilde{a}^\dagger_{\alpha_1} \dots \widetilde{a}^\dagger_{\alpha_v}\widetilde{a}_{\beta_1} \dots \widetilde{a}_{\beta_v}|w\rangle\right| && (73b)\\ &\le \frac{1}{2^{nv}}\sum_{\substack{\alpha_1,\dots,\alpha_v\\ \beta_1,\dots,\beta_v}}\left|\langle u|\widetilde{a}^\dagger_{\alpha_1} \dots \widetilde{a}^\dagger_{\alpha_v}\widetilde{a}_{\beta_1} \dots \widetilde{a}_{\beta_v}|w\rangle\right|. && (73c) \end{aligned}$$

Next, if $d > v$, observe that any term in eq. (73c) corresponds to subtracting 1 from momentum modes $\beta_1, \dots, \beta_v$, and adding 1 to momentum modes $\alpha_1, \dots, \alpha_v$ starting from the quasi-even condensate. Furthermore, the term is non-zero if any only if they correspond to a sequence of additions and subtractions that map w to u. By assumption, u and w differ by $2d > 2v$ edits in terms of the 1-norm, and so there is no sequence of v many additions $\alpha_1, \dots, \alpha_v$ and subtractions $\beta_1, \dots, \beta_v$ mapping w to u, and thus every term in the sum is 0.

Now we switch to the case when $d \le v$. We first bound each of the non-zero terms in the sum. Fix a choice $\alpha_1, \dots, \alpha_v$ and $\beta_1, \dots, \beta_v$, and consider the term

$$\left|\langle u|\,\widetilde{a}^\dagger_{\alpha_1} \dots \widetilde{a}^\dagger_{\alpha_v}\widetilde{a}_{\beta_1} \dots \widetilde{a}_{\beta_v}\,|w\rangle\right|. \tag{74}$$

As noted before, this is only non-zero if subtracting 1 from the modes $\beta_1, \dots, \beta_v$ and adding 1 to the modes $\alpha_1, \dots, \alpha_v$ maps from the Fock state w to u. To bound the term, we notice that when we apply an annihilation operator to a mode z, the norm can increase multiplicatively by at most $\sqrt{w_z}$, and whenever $z \ne 0$, we can bound w_z by r, the total number of non-zero bosons. Next, since w, u both are exactly ℓ boson states, at most half of their difference can be accounted for by the 0-mode. Since w and u differ by $2d$, therefore at least d of the combined collection of β's and α's must be non-zero, meaning at most $2v - d$ of them can be 0. The $\le 2v - d$ annihilation operators corresponding to the 0-mode have operator norm bounded by $\sqrt{\ell}$, and the remaining d operators are bounded by $\sqrt{r}$ (when restricting to the subspace of (r, o)-quasi-even condensates). Thus, we have the following upper bound:

$$\left|\langle u|\,\widetilde{a}^\dagger_{\alpha_1} \dots \widetilde{a}^\dagger_{\alpha_v}\widetilde{a}_{\beta_1} \dots \widetilde{a}_{\beta_v}\,|w\rangle\right| \le \ell^{v-d/2}r^{d/2}. \tag{75}$$

The final step in the proof is to bound the total number of non-zero terms in the sum. To do this, we make two observations. We first note that we can bound the number of choices of $\beta_1, \dots, \beta_v$ that correspond

to a non-zero term in the sum by $\binom{r+1}{v}$. This is because whenever β_i acts on a mode that is unoccupied in w, it maps $|w\rangle$ to 0, and there are only $r+1$ many occupied modes in w.

Then, we notice that for any u, w and $\beta_1, \cdots, \beta_v$, there is a unique multiset of creation operators $\{\alpha_1, \cdots, \alpha_v\}$ such that $\langle u| \widetilde{a}^\dagger_{\alpha_1} \dots \widetilde{a}^\dagger_{\alpha_v} \widetilde{a}_{\beta_1} \dots \widetilde{a}_{\beta_v} |w\rangle \neq 0$. The number of possible ordered lists of $\alpha_1, \cdots, \alpha_v$ is then at most $v!$. Combining our bounds, we have

$$|\langle u|\widehat{n}_{z_1} \dots \widehat{n}_{z_v}|w\rangle| \leq \frac{1}{2^{nv}} \sum_{\substack{\alpha_1,\dots,\alpha_v \\ \beta_1,\dots,\beta_v}} \left|\langle u|\widetilde{a}^\dagger_{\alpha_1} \dots \widetilde{a}^\dagger_{\alpha_v} \widetilde{a}_{\beta_1} \dots \widetilde{a}_{\beta_v}|w\rangle\right| \tag{76a}$$

$$\leq \frac{1}{2^{nv}} \ell^{v-d/2} r^{d/2} \sum_{\substack{\alpha_1,\dots,\alpha_v \\ \beta_1,\dots,\beta_v}} \delta\left(\left|\langle u| \widetilde{a}^\dagger_{\alpha_1} \dots \widetilde{a}^\dagger_{\alpha_v} \widetilde{a}_{\beta_1} \dots \widetilde{a}_{\beta_v} |w\rangle\right| \neq 0\right) \tag{76b}$$

$$= \frac{1}{2^{nv}} \ell^{v-d/2} r^{d/2} v! \binom{r+1}{v} \tag{76c}$$

$$\leq v!(2r)^{v+d/2} \frac{\ell^{v-d/2}}{2^{nv}} . \tag{76d}$$

Here, we apply our bound on the magnitude of $\left|\langle u|\widetilde{a}^\dagger_{\alpha_1} \dots \widetilde{a}^\dagger_{\alpha_v} \widetilde{a}_{\beta_1} \dots \widetilde{a}_{\beta_v}|w\rangle\right|$, then apply our bound on the number of non-zero terms in the sum, and finally re-arrange terms to get the desired expression. □

Finally, we prove the second of the additional claims. The proof of the lemma is more combinatorial in nature than the other proofs in this section, so we provide some intuition before the proof.

One way to bound the number of quasi-even condensates at distance $2d$ from a fixed condensate u would be to think about tuples e (for "error" relative to u) with sum 0 and 1-norm $2d$ (representing the distance between u and another quasi-even condensate) and upper bound the ways to construct e. Of course, simply taking all vectors of norm $2d$ would imply an upper bound of $\sim (2^n)^{2d}$, i.e., to add d units of positive difference, and d units of negative distance to any of the 2^n entries. However, this does not use the fact that our condensates are mostly paired up. To take advantage of this, we might imagine that we split $e = 2e_{\text{even}} + e_{\text{odd}}$, where e_{odd} is a vector consisting of $\{-1, 0, +1\}$ entries. Here, we again run into a number of problems. First, even ignoring e_{odd}, the number of ways to create e_{even} is still roughly $\sim (2^n)^d$ (i.e., add $d/2$ positive and negative units to the 2^n units), and secondly, there are many choices for e_{even} and e_{odd}, because anytime an entry of e_{odd} is $+1$, we can add 1 to the corresponding entry of e_{even} and switch e_{odd} to be -1. To fix these issues requires first noticing that we can only negative units to entries of u that are non-zero, reducing the number of ways to generate e_{even} to $\sim (2^n)^{d/2} \cdot r^{d/2}$, and describing a "canonical" way to assign e_{odd}. These complications require some careful and lengthy accounting, which the following claim handles.

Claim (Claim 10.5 restated)**.** *For every (r, o)-quasi-even condensate tuple u and $d \geq 0$, the number of (r, o)-quasi-even condensate tuples w that are at exactly distance $2d$ from u is upper bounded by*

$$(2^{n+1})^{d/2+o}(r + d/2 + o)^{d/2+o} . \tag{77}$$

Proof. The proof will proceed as follows: We will first define a more abstract counting problem having to do with placing balls in bins with constraints. We will then show that the answer to this counting problem upper bounds the number of (r, o)-quasi-even condensates at distance d from u, and further show that

the answer is upper bounded by eq. (68).

Defining the counting problem. For every tuple u, define $\mathrm{pos}(u)$ to be the indices of u that have non-zero entries. The counting problem that we consider is the number of ways to assign balls into 2^n bins such that:

1. Each ball is labeled with an integer -2, -1, +1, or +2 and the balls of a particular label are identical.
2. There are exactly $\lfloor d/2 \rfloor$ many $+2$ and $\lfloor d/2 \rfloor$ many -2 balls, and exactly o many $+1$ and o many -1 balls.
3. For any bin outside of $\mathrm{pos}(u)$, the sum of the labels of the balls placed in this bin must be ≥ 0.

We first prove that the number of assignments of balls to bins is an upper bound for the number of (r, o)-quasi-even condensates at distance $2d$ from a given quasi-even condensate u by constructing an injective map from quasi-even condensates to assignments.

Constructing the injective map. Fix a quasi-even condensate w at distance $2d$ from u and let $e = w - u$ be the entry-wise difference of u and w. Note that because u and w have the same number of bosons, the sum of entries of e must be 0. Divide e into positive and negative components—i.e., such that

$$e_+ \geq 0, e_- \geq 0 \text{ such that } e = e_+ - e_-. \tag{78}$$

Define $e_{\mathrm{even}} \stackrel{\mathrm{def}}{=} 2\left(\lfloor e_+/2 \rfloor - \lfloor e_-/2 \rfloor\right)$, where operations like the floor and dividing by 2 act entry-wise on the tuple, and define $e_{\mathrm{odd}} \stackrel{\mathrm{def}}{=} e - e_{\mathrm{even}}$. We first make some remarks about e_{even} and e_{odd}. For any entry e_x,

$$(e_{\mathrm{odd}})_x = \begin{cases} 0 & \text{if } e_x \text{ is even,} \\ \mathrm{sgn}(e_x) & \text{if } e_x \text{ is odd.} \end{cases} \tag{79}$$

We further note that because we started with two quasi-even condensates, the number of non-zero entries in e_{odd} is upper bounded by $2o$. Finally, we note that the tuple e_{even} always satisfies $|e_{\mathrm{even}}| \leq 2\lfloor d/2 \rfloor$, meaning that if d is not even, e_{even} accounts for strictly less than $d/2$ of the distance between w and u.

Now consider the following assignment of balls to bins. Without loss of generality, assume that the sum of entries of e_{even} is ≤ 0; the positive case follows by a similar logic.

1. For each bin x, if $(e_{\mathrm{even}})_x \geq 0$, add $(e_{\mathrm{even}})_x/2$ many $+2$ balls to the bin, and if it $e'_x < 0$, add $(e_{\mathrm{even}})_x/2$ many -2 balls to the bin.
2. Let b_+ and b_- be the number of $+2$ and -2 balls assigned in the previous step, respectively. Note, $b_+ \leq b_-$. Add $\lfloor d/2 \rfloor - b_-$ many $+2$ and the same number of -2 balls to the 0-th bin. This assigns all the -2 balls.
3. $b_{\mathrm{rem}} \stackrel{\mathrm{def}}{=} b_- - b_+ \geq 0$ is the number of $+2$ that have not been assigned (rem stands for "remainder"). Assign an additional $+2$ ball to the bins corresponding to the first (in lexicographical ordering) b_{rem} entries of e_{odd} that equal $+1$. Then, define e_{balanced} to be equal to e_{odd} where we subtract 2 from the entries affected by this step.
4. At this step, e_{balanced} is a tuple where every entry is $\in \{-1, 0, +1\}$ because the previous step only subtracts 2 from entries in e_{odd} that were $+1$. Furthermore, we have that the sum of entries in e_{balanced} is 0 and there are at most $2o$ many non-zero entries of e_{balanced}. Thus, there are at most

$o' \le o$ many $+1$ and o' many -1 entries of e_{balanced}. Thus, we can assign o' many $+1$ and -1 balls according to non-zero entries of e_{balanced}, and assign the remaining $o - o'$ many $+1$ and -1 balls to the 0-th bin.

One can verify that this is a valid assignment. Finally, if we take the assignment of balls that this procedure outputs and add up the labels of the balls, we recover the vector $e = w - u$. Since for any two quasi-even condensates w and w', this difference vector can not be the same, we have that our assignment is injective, and thus the number of assignments upper bounds the number of quasi-even condensates at distance exactly $2d$.

Bounding the combinatorial identity. Now, we upper bound the number of assignments of the balls to bins using the stars and bars counting trick[a]. We can over-count this by (a) assigning the $+2$ and $+1$ balls to $2^n - 1$ bins, and (b) assigning the -2 and -1 balls to the $r + 1$ many bins corresponding to $\mathrm{pos}(u)$, as well as the at most $d/2 + o$ bins that the $+2$ and $+1$ balls were assigned to. We have the following bounds on (a) and (b)

$$(a) \le \binom{\frac{d}{2} + k + 2^n}{\frac{d}{2}, k, 2^n} = \frac{\left(\frac{d}{2} + k + 2^n\right)!}{\left(\frac{d}{2}\right)! k! (2^n)!} \le \left(2^{n+1}\right)^{\frac{d}{2}+k} \le \left(2^{n+1}\right)^{\frac{d}{2}+o} \tag{80a}$$

$$(b) \le \binom{r + 1 + \frac{d}{2} + o}{r + 1, \frac{d}{2}, o} \le (r + 1 + d/2 + o)^{\frac{d}{2}+o}. \tag{80b}$$

Here, we are ignoring the effect of taking $\lfloor d/2 \rfloor$, since it only decreases the quantities. Taking the product of (a) and (b) upper bounds the number of ways to assign the balls to bins, which completes the proof. □

[a] *Stars and bars* is a combinatorial technique for counting the number of ways to partition a identical items ("stars") into b distinct bins [Wik25]. The number of solutions is $\binom{a+b-1}{b-1}$. This can be derived by observing that any placement is equivalent to placing a ("stars") and $b - 1$ "bars" in a row, with the number of stars between two bars denoting the number of stars to place in the corresponding bin. The identified bijection proves that the number of ways is equal to selecting $b - 1$ locations for the bars among $a + b - 1$.

Part IV
Polynomial-query algorithms generate quasi-even condensates

In this part, we complete the proof of the sampling probability upper bound by showing that polynomial-query algorithms are almost entirely supported on quasi-even condensates.

11 A bosonic compressed oracle technique

In this section, we describe a natural compressed oracle technique for a family of random sparse functions, like S, and identify a natural basis in which queries to U (when (S, U) are sampled from the distribution Strong) has a simple purified description in terms of bosonic hopping operators.

11.1 Initial bosonic state

We first need to compute a purification of the choice of multiset S. As mentioned in the introduction, one technique is to sample a uniformly random multiset of size ℓ is preparing ℓ bosons initialized in the 0-momentum mode and then measuring in the position basis. Therefore, a purification of the oracle S is simply ℓ bosons in the 0-momentum mode:

$$\frac{1}{\sqrt{\ell!}}\left(\widetilde{a}_0^\dagger\right)^\ell |\mathrm{vac}\rangle\,. \tag{81}$$

Claim 11.1. *The state above is equal to the uniform superposition over all multisets of ℓ elements, where each element is sampled uniformly at random, independent of the other elements.*

Proof. We can write the purification of a uniformly random multiset of ℓ elements in the Fock basis as follows:

$$\sum_{\substack{\ell_1,\ldots,\ell_{2^n}\in\mathbb{Z}_{\geq 0}\\ \sum_{x\in\{0,1\}^n}\ell_x=\ell}} \sqrt{p_{\ell_0,\ldots,\ell_{2^n-1}}}\,|\ell_0,\ldots,\ell_{2^n-1}\rangle\,, \tag{82}$$

where $p_{\ell_0,\ldots,\ell_{2^n-1}} = \frac{1}{2^{n\ell}}\frac{\ell!}{\prod_{x\in\{0,1\}^n}\ell_x!}$ is the probability of measuring getting ℓ_x copies of x when sampling ℓ uniformly random strings. We can also write out the result of applying ℓ creation operators in the 0-momentum mode as follows:

$$\frac{1}{\sqrt{\ell!}}\left(\widetilde{a}_0^\dagger\right)^\ell |\mathrm{vac}\rangle = \frac{1}{\sqrt{\ell!2^{n\ell}}}\sum_{x_1,\ldots,x_\ell\in\{0,1\}^n}\left(\prod_{i=1}^{\ell}\widehat{a}_{x_i}^\dagger\right)|\mathrm{vac}\rangle \tag{83a}$$

$$= \frac{1}{\sqrt{\ell!2^{n\ell}}}\sum_{\substack{\ell_0,\ldots,\ell_{2^n-1}\in\mathbb{Z}_{\geq 0}\\ \sum_{x\in\{0,1\}^n}\ell_x=\ell}}\frac{\ell!}{\prod_{x\in\{0,1\}^n}\ell_x!}\left(\prod_{i=1}^{\ell}\widehat{a}_{x_i}^\dagger\right)|\mathrm{vac}\rangle \tag{83b}$$

$$= \frac{1}{\sqrt{2^{n\ell}}} \sum_{\substack{\ell_0,\dots,\ell_{2^n-1}\in\mathbb{Z}_{\geq 0} \\ \sum_{x\in\{0,1\}^n}\ell_x=\ell}} \sqrt{\frac{\ell!}{\prod_{x\in\{0,1\}^n}\ell_x!}}\,|\ell_0,\dots,\ell_{2^n-1}\rangle\ . \tag{83c}$$

We first use the fact that, when sampling a random collection $x_1,\dots,x_\ell$, the Fock basis state $|\ell_0,\dots,\ell_{2^n-1}\rangle$ appears exactly $\ell!/\prod_{x\in\{0,1\}^n}\ell_x!$ times. In the final line, we use the fact that by the definition of the creation operators, $|\ell_0,\dots,\ell_{2^n-1}\rangle = \frac{1}{\prod_{x\in\{0,1\}^n}\ell_x!}\prod_{x\in\{0,1\}^n}\widehat{a}_x^{\dagger\ell_x}|\mathrm{vac}\rangle$. Since these two states are equal, we have completed the proof. □

11.2 Purified state of algorithm and oracle registers

Next, we introduce the purification of the state of a query algorithm querying U, where U is sampled from the distribution Strong, Definition 7.2. For this, we will take inspiration from the compressed oracle technique introduced by Zhandry [Zha19]. To construct the compressed oracle, we write out the following purification of the initial state of the system for both S and U; we assume the state is expressed on purifying registers S and U.

$$|\mathrm{init}\rangle_{\mathsf{SU}} \stackrel{\mathrm{def}}{=} \frac{1}{\sqrt{\ell!}}\left(\widetilde{a}_0^\dagger\right)^\ell |\mathrm{vac}\rangle_{\mathsf{S}} \otimes |\perp\rangle_{\mathsf{U}}^{\otimes 2^n}\ . \tag{84}$$

Here, U is divided into $\bigotimes_{y\in\{0,1\}^n}\mathsf{U}_y$, where each single-qubit register U_y is initially in the state $|\perp\rangle$. We also define the initial state restricted to the S or U registers,

$$|\mathrm{init}_S\rangle_{\mathsf{S}} \stackrel{\mathrm{def}}{=} \frac{1}{\sqrt{\ell!}}\left(\widetilde{a}_0^\dagger\right)^\ell |\mathrm{vac}\rangle_{\mathsf{S}} \quad \text{and} \quad |\mathrm{init}_U\rangle_{\mathsf{U}} \stackrel{\mathrm{def}}{=} |\perp\rangle_{\mathsf{U}}^{\otimes 2^n}\ , \tag{85}$$

We then define the following isometry acting on registers US:

$$\begin{aligned}\mathcal{V}_1 \stackrel{\mathrm{def}}{=} \sum_S |\mathrm{tt}_S\rangle\langle\mathrm{tt}_S|_{\mathsf{S}} \otimes \bigotimes_{y\in\{0,1\}^n} \Bigg(&\left(\sqrt{1-\frac{1}{2}\mathrm{e}^{-\kappa\gamma_y^{(S)}}}\,|0\rangle + \sqrt{\frac{1}{2}\mathrm{e}^{-\kappa\gamma_y^{(S)}}}\,|1\rangle\right)\langle\perp|_{\mathsf{U}_y} \\ &+ \left(\sqrt{\frac{1}{2}\mathrm{e}^{-\kappa\gamma_y^{(S)}}}\,|0\rangle - \sqrt{1-\frac{1}{2}\mathrm{e}^{-\kappa\gamma_y^{(S)}}}\,|1\rangle\right)\langle\top|_{\mathsf{U}_y}\Bigg).\end{aligned} \tag{86}$$

Note that this is a unitary and only acts on US. Finally, we define a second isometry acting on registers AU, with A acting as the algorithm's query register.

$$\mathcal{V}_2 \stackrel{\mathrm{def}}{=} \sum_{\substack{y\in\{0,1\}^n \\ b\in\{0,1\}}} \sum_U (-1)^{b\cdot U(y)}\,|b,y\rangle\langle b,y|_{\mathsf{A}} \otimes |\mathrm{tt}_U\rangle\langle\mathrm{tt}_U|_{\mathsf{U}}\,. \tag{87}$$

Then, we have the following:

Lemma 11.2. *For all query algorithms $\mathcal{A}$,*

$$\mathrm{Tr}_{\mathsf{US}}\left[\mathcal{A}^{\mathcal{V}_1^\dagger\cdot\mathcal{V}_2\cdot\mathcal{V}_1}(|0,\mathrm{init}\rangle\langle 0,\mathrm{init}|)\right] = \mathop{\mathbb{E}}_{S,U}\left[\mathcal{A}^U(|0\rangle\langle 0|)\right]\ . \tag{88}$$

Proof. Since $\mathcal{V}_1$ only acts on US, it commutes with the unitaries that $\mathcal{A}$ applies. Moreover, the $\mathcal{V}_1$ from one query cancels out the $\mathcal{V}_1^\dagger$ from the next query. The result is that only the inner-most and outer-most

$\mathcal{V}_1$ and $\mathcal{V}_1^\dagger$ are left. Thus, we can rewrite the left side of the equation as

$$\mathrm{Tr}_{\mathsf{US}}[\mathcal{V}_1^\dagger \mathcal{A}^{\mathcal{V}_2}(|0\rangle\langle 0| \otimes \mathcal{V}_1 |\mathrm{init}\rangle\langle \mathrm{init}| \mathcal{V}_1^\dagger)\mathcal{V}_1] = \mathrm{Tr}_{\mathsf{US}}[\mathcal{A}^{\mathcal{V}_2}(|0\rangle\langle 0| \otimes \mathcal{V}_1 |\mathrm{init}\rangle\langle \mathrm{init}| \mathcal{V}_1^\dagger)] \,. \tag{89}$$

Then we note that applying $\mathcal{V}_1$ to $|\mathrm{init}\rangle$ yields exactly the following state:

$$\mathcal{V}_1 |\mathrm{init}\rangle = \frac{1}{\sqrt{\ell! \cdot 2^{n\ell}}} \sum_{s_1,\ldots,s_\ell \in \{0,1\}^n} \bigotimes_{y \in \{0,1\}^n} \left(\sqrt{1 - \frac{1}{2}\mathrm{e}^{-\kappa\gamma_y^{(S)}}} |0\rangle + \sqrt{\frac{1}{2}\mathrm{e}^{-\kappa\gamma_y^{(S)}}} |1\rangle \right)_{\mathsf{U}_y} \otimes \widehat{a}_{s_1}^\dagger \ldots \widehat{a}_{s_\ell}^\dagger |\mathrm{vac}\rangle \,. \tag{90}$$

From here, it is clear to see that tracing out the US register of $\mathcal{A}$ yields a random U and S according to the distribution Strong, and that $\mathcal{A}$ querying $\mathcal{V}_2$ yields an identical mixed state to $\mathcal{A}$ querying a U (where (S, U) are sampled according to Strong). □

Next, we make a simplification to this oracle that gives the oracle a nicer form.

Corollary 11.3. *Define the following Kraus operators acting on* S*, parametrized by* y*.*

$$E_0^{(y)} \stackrel{\mathrm{def}}{=} \sum_S (1 - \mathrm{e}^{-\kappa\gamma_y^{(S)}}) |\mathrm{tt}_S\rangle\langle \mathrm{tt}_S|_{\mathsf{S}} \quad \textit{and} \quad E_1^{(y)} \stackrel{\mathrm{def}}{=} \sum_S \sqrt{\mathrm{e}^{-\kappa\gamma_y^{(S)}}(2 - \mathrm{e}^{-\kappa\gamma_y^{(S)}})} |\mathrm{tt}_S\rangle\langle \mathrm{tt}_S|_{\mathsf{S}} \,. \tag{91}$$

Then define $\mathcal{O}$ *to be the following unitary acting on registers* AUS*.*

$$\mathcal{O} \stackrel{\mathrm{def}}{=} \sum_{y \in \{0,1\}^n, b \in \{0,1\}} |b, y\rangle\langle b, y|_{\mathsf{A}} \otimes \left(\widetilde{Z}_{\mathsf{U}_y} \otimes \left(E_0^{(y)}\right)_{\mathsf{S}} + \widetilde{X}_{\mathsf{U}_y} \otimes \left(E_1^{(y)}\right)_{\mathsf{S}} \right)^b \,, \tag{92}$$

where $\widetilde{X}$ *and* $\widetilde{Z}$ *are the usual Pauli operators except in the* $|\bot\rangle$ *and* $|\top\rangle$ *basis (as opposed to the* $|0\rangle$ *and* $|1\rangle$ *basis). Note the exponent of* b *acting on the unitary applied to the* US *registers. Then,*

$$\mathrm{Tr}_{\mathsf{US}}[\mathcal{A}^{\mathcal{O}}(|0, \mathrm{init}\rangle\langle 0, \mathrm{init}|)] = \mathbb{E}_{S,U}\left[\mathcal{A}^U(|0\rangle\langle 0|)\right] \,. \tag{93}$$

Proof. Observe that $(E_0^{(y)})^2 + (E_1^{(y)})^2 = \mathrm{id}$, $0 \preceq E_0^{(y)} \preceq \mathrm{id}$, and $0 \preceq E_1^{(y)} \preceq \mathrm{id}$. From Lemma 11.2, for any algorithm $\mathcal{A}$, querying $\mathcal{V}_1^\dagger \cdot \mathcal{V}_2 \cdot \mathcal{V}_1$ is identical to querying U after tracing out US. To prove the corollary, we show that $\mathcal{V}_1^\dagger \cdot \mathcal{V}_2 \cdot \mathcal{V}_1$ is equal to $\mathcal{O}$. Writing it out, we will get the following:

$$\mathcal{V}_1^\dagger \cdot \mathcal{V}_2 \cdot \mathcal{V}_1 \tag{94a}$$

$$= \sum_{S,y} |1, y\rangle\langle 1, y|_{\mathsf{A}} \otimes \Bigg(\left((1 - \mathrm{e}^{-\kappa\gamma_y^{(S)}}) |\bot\rangle + \sqrt{\mathrm{e}^{-\kappa\gamma_y^{(S)}}(2 - \mathrm{e}^{-\kappa\gamma_y^{(S)}})} |\top\rangle \right) \langle\bot|_{\mathsf{U}_y} + \left(\sqrt{\mathrm{e}^{-\kappa\gamma_y^{(S)}}(2 - \mathrm{e}^{-\kappa\gamma_y^{(S)}})} |\bot\rangle - (1 - \mathrm{e}^{-\kappa\gamma_y^{(S)}}) |\top\rangle \right) \langle\top|_{\mathsf{U}_y} \Bigg) \otimes |\mathrm{tt}_S\rangle\langle \mathrm{tt}_S|_{\mathsf{S}} + \sum_y |0, y\rangle\langle 0, y|_{\mathsf{A}} \otimes \mathrm{id}_{\mathsf{US}} \tag{94b}$$

$$= \sum_{S,y} |1, y\rangle\langle 1, y|_{\mathsf{A}} \otimes \left((1 - \mathrm{e}^{-\kappa\gamma_y^{(S)}}) \cdot \widetilde{Z} + \sqrt{\mathrm{e}^{-\kappa\gamma_y^{(S)}}(2 - \mathrm{e}^{-\kappa\gamma_y^{(S)}})} \cdot \widetilde{X} \right)_{\mathsf{U}_y} \otimes |\mathrm{tt}_S\rangle\langle \mathrm{tt}_S|_{\mathsf{S}} + \sum_y |0, y\rangle\langle 0, y| \otimes \mathrm{id}_{\mathsf{US}} \,. \tag{94c}$$

Here, $\widetilde{Z} = |\perp\rangle\langle\perp| - |\top\rangle\langle\top|$ and $\widetilde{X} = |\top\rangle\langle\perp| + |\perp\rangle\langle\top|$ are the Pauli Z and X operators in the $|\top\rangle$ and $|\perp\rangle$ basis. Direct calculation shows that this unitary squares to the identity as expected. Then we can rewrite this using the definition of $E_0^{(y)}$ and $E_1^{(y)}$ as follows:

$$\mathcal{V}_1^\dagger \cdot \mathcal{V}_2 \cdot \mathcal{V}_1 = \sum_{y,b} (|b,y\rangle\langle b,y|_{\mathsf{A}} \otimes \mathrm{id}_{\mathsf{B}}) \otimes \left(\widetilde{Z}_{\mathsf{U}_y} \otimes \left(E_0^{(y)}\right)_{\mathsf{S}} + \widetilde{X}_{\mathsf{U}_y} \otimes \left(E_1^{(y)}\right)_{\mathsf{S}}\right)^b . \tag{95}$$

Applying Lemma 11.2 completes the proof. □

This implies that we can write the purified state of any algorithm $\mathcal{A}$ that only queries U in a simple form. Here we assume the query algorithm can be expressed as a repeating sequence of oracle queries and unitaries A_{A} (without loss of generality, we can assume the same unitary A is applied each time):

$$\underbrace{A\mathcal{O}A\mathcal{O}\dots\mathcal{O}}_{T \text{ times}} A\,|0\rangle_{\mathsf{A}}\,|\mathrm{init}\rangle_{\mathsf{US}} \tag{96}$$

where there are T queries to the oracle U. The query access to controlled-U described by eq. (92) can be further abbreviated as

$$\mathcal{O} = \sum_{\substack{y\in\{0,1\}^n \\ b\in\{0,1\}}} |b,y\rangle\langle b,y|_{\mathsf{A}} \otimes \left(\sum_{x\in\{0,1\}} \underbrace{\left(\widetilde{Z}^{1-x} + \widetilde{X}^x\right)_{\mathsf{U}_y}}_{\stackrel{\text{def}}{=}(K^x)_{\mathsf{U}_y}} \otimes (E_x^{(y)})_{\mathsf{S}}\right)^b , \tag{97}$$

where we define K^x for notational simplicity. Note the exponent b in the large parenthetical. Then, an alternating sequence of algorithm and queries can be calculated as follows. Given $\mathbf{y} = (y_1,\dots,y_T)$, $\mathbf{b} = (b_1,\dots,b_T)$, let $\mathbf{y}^{\mathbf{b}} = (y_i : b_i = 1)$ be the ordered tuple of length $|\mathbf{b}| \stackrel{\text{def}}{=} \sum_i b_i$ corresponding to indices of $\mathbf{y}$ where $\mathbf{b}$ is 1, with $y_i^{\mathbf{b}}$ being the i'th entry of this tuple. Then given $\mathbf{x} = (x_1,\dots,x_{|\mathbf{b}|})$, define the abbreviations:

$$(K^{\mathbf{x}})_{\mathbf{y},\mathbf{b}} \stackrel{\text{def}}{=} \prod_{i=1}^{|\mathbf{b}|} (K^{x_i})_{\mathsf{U}_{y_i^{\mathbf{b}}}}, \qquad E_{\mathbf{x}}^{(\mathbf{y},\mathbf{b})} \stackrel{\text{def}}{=} \prod_{i=1}^{|\mathbf{b}|} E_{x_i}^{(y_i^{\mathbf{b}})}, \qquad A_{\mathbf{y},\mathbf{b}} \stackrel{\text{def}}{=} \prod_{i=T}^{1} A\cdot|b_i,y_i\rangle\langle b_i,y_i| . \tag{98}$$

Then, the post-query state will be

$$|\psi_{\mathrm{PQ}}\rangle = |\psi_{\mathrm{Post-Query}}\rangle \stackrel{\text{def}}{=} \sum_{\substack{\mathbf{y}\in\{0,1\}^{nT} \\ \mathbf{b}\in\{0,1\}^T}} \left[\left(A_{\mathbf{y},\mathbf{b}}\right)_{\mathsf{A}} \otimes \left(\sum_{\mathbf{x}\in\{0,1\}^{|\mathbf{b}|}} (K^{\mathbf{x}})_{\mathbf{y},\mathbf{b}} \otimes E_{\mathbf{x}}^{(\mathbf{y},\mathbf{b})}\right)\right] A\,|0\rangle_{\mathsf{A}}\,|\mathrm{init}\rangle_{\mathsf{US}} . \tag{99}$$

The abbreviated notation will prove itself useful when we start approximating the oracle queries with polynomials. Note that in the rest of the section, we drop the register indices when they are clear from context, and whenever an operator acts as identity on a register, we omit the ⊗id part of the operator. Having developed an expression for the state after applying T queries, we now press forward and study how the post-query state (expressed in eq. (99)) is almost entirely supported on quasi-even condensates. This will be the content of the next section. By doing so, we can appeal to Theorem 10.2 to conclude a probability upper bound.

11.3 Action of the oracle in the momentum basis

We see that $\mathcal{O}$ consists of Kraus operators $E_0^{(y)}$ and $E_1^{(y)}$ that can be thought of as functions of a map $|\mathrm{tt}_S\rangle \mapsto \gamma_y^{(S)} |\mathrm{tt}_S\rangle$. We now examine the action of this map directly before examining the action of $\mathcal{O}$. By doing so, we will build a natural framework for understanding how to study maps that are more complex functions of $\gamma_y^{(S)}$ such as $|\mathrm{tt}_S\rangle$ to either $E_0^{(y)} |\mathrm{tt}_S\rangle$ or $E_1^{(y)} |\mathrm{tt}_S\rangle$. This action, perhaps surprisingly, has a natural interpretation in the momentum space. To see this, we define the "single y-momentum hopping operator" $\widetilde{\mathbf{G}}_y$ and the "double y-momentum hopping operator" $\widetilde{\mathbf{H}}_y$ as[11]:

$$\widetilde{\mathbf{G}}_y \stackrel{\text{def}}{=} \frac{1}{\sqrt{\ell}} \sum_{x \in \{0,1\}^n} \widetilde{a}^\dagger_{x \oplus y} \widetilde{a}_x. \tag{100a}$$

$$\widetilde{\mathbf{H}}_y \stackrel{\text{def}}{=} \frac{1}{\ell} \sum_{x,x' \in \{0,1\}^n} \widetilde{a}^\dagger_{x \oplus y} \widetilde{a}^\dagger_{x' \oplus y} \widetilde{a}_x \widetilde{a}_{x'}. \tag{100b}$$

We also refer to these as single-hopping or double-hopping operators when y is clear from context. (In this work, we never use the analogous *position* hopping operators.) Morally, for a bosonic system $|\psi\rangle$, we can interpret $\widetilde{\mathbf{G}}_y$ as adding momentum y (in superposition) to each boson and $\widetilde{\mathbf{H}}_y |\psi\rangle$ as adding momentum y (in superposition) to each pair of bosons. Observe that action by $\widetilde{\mathbf{H}}_y |\psi\rangle$ increases the momentum of the entire system by $2y = 0$. Therefore, the total momentum of the system is conserved under any function of $\widetilde{\mathbf{H}}_y$. Second, note that the single- and double-momentum hopping operators are related:

$$\widetilde{\mathbf{G}}_y^2 = \frac{1}{\ell} \sum_{x,x' \in \{0,1\}^n} \widetilde{a}^\dagger_{x \oplus y} \widetilde{a}_x \widetilde{a}^\dagger_{x' \oplus y} \widetilde{a}_{x'} \tag{101a}$$

$$= \frac{1}{\ell} \sum_{x,x' \in \{0,1\}^n} \widetilde{a}^\dagger_{x \oplus y} \left(\widetilde{a}^\dagger_{x' \oplus y} \widetilde{a}_x + \delta_{x,x' \oplus y} \right) \widetilde{a}_{x'} \tag{101b}$$

$$= \frac{1}{\ell} \sum_{x,x' \in \{0,1\}^n} \widetilde{a}^\dagger_{x \oplus y} \widetilde{a}^\dagger_{x' \oplus y} \widetilde{a}_x \widetilde{a}_{x'} + \frac{1}{\ell} \sum_{x \in \{0,1\}^n} \widetilde{a}^\dagger_x \widetilde{a}_x \tag{101c}$$

$$= \widetilde{\mathbf{H}}_y + \frac{\widetilde{N}}{\ell}. \tag{101d}$$

Within the subspace of ℓ bosons,

$$\widetilde{\mathbf{G}}_y^2 = \widetilde{\mathbf{H}}_y + \mathrm{id}. \tag{102}$$

We prove the following lemma relating the hopping operators to the map $|\mathrm{tt}_S\rangle \mapsto \gamma_y^{(S)} |\mathrm{tt}_S\rangle$.

Lemma 11.4 (Single hopping twice applies $\gamma_y^{(S)}$)**.** *For all $y \in \{0,1\}^n$, the following holds.*

$$\widetilde{\mathbf{G}}_y^2 \widehat{a}^\dagger_{s_1} \dots \widehat{a}^\dagger_{s_\ell} |\mathrm{vac}\rangle = \gamma_y^{(S)} \widehat{a}^\dagger_{s_1} \dots \widehat{a}^\dagger_{s_\ell} |\mathrm{vac}\rangle \ . \tag{103}$$

Proof. We can directly compute the action of $\widetilde{\mathbf{G}}_y$ on a position basis state to relate it to $\gamma_y^{(S)}$.

$$\widetilde{\mathbf{G}}_y \widehat{a}^\dagger_{s_1} \dots \widehat{a}^\dagger_{s_\ell} |\mathrm{vac}\rangle = \frac{1}{\sqrt{\ell}} \sum_{x \in \{0,1\}^n} \widetilde{a}^\dagger_{x \oplus y} \widetilde{a}_x \widehat{a}^\dagger_{s_1} \dots \widehat{a}^\dagger_{s_\ell} |\mathrm{vac}\rangle \tag{104a}$$

[11] We use normalized versions of the single and double y-momentum hopping operators. This is because we are only considering systems of ℓ bosons. A situation with a variable number of bosons might use the definitions without the $1/\sqrt{\ell}$ and $1/\ell$ constants, respectively.

$$= \frac{1}{\sqrt{\ell N^\ell}} \sum_{x\in\{0,1\}^n} \sum_{t_1,\ldots,t_\ell} (-1)^{t_1\cdot s_1+\ldots+t_\ell\cdot s_\ell} \widetilde{a}^\dagger_{x\oplus y} \widetilde{a}_x \widetilde{a}^\dagger_{t_1} \ldots \widetilde{a}^\dagger_{t_\ell} |\mathrm{vac}\rangle \tag{104b}$$

$$= \frac{1}{\sqrt{\ell N^\ell}} \sum_{i=1}^{\ell} \sum_{t_1,\ldots,t_\ell} (-1)^{t_1\cdot s_1+\ldots+t_\ell\cdot s_\ell} \widetilde{a}^\dagger_{t_i\oplus y} \left(\prod_{k\,:\,k\neq i} \widetilde{a}^\dagger_{t_k}\right) |\mathrm{vac}\rangle \tag{104c}$$

$$= \frac{1}{\sqrt{\ell N^\ell}} \sum_{i=1}^{\ell} \sum_{t_1,\ldots,t_\ell} (-1)^{y\cdot s_i} (-1)^{t_1\cdot s_1+\ldots+t_\ell\cdot s_\ell} \widetilde{a}^\dagger_{t_1} \ldots \widetilde{a}^\dagger_{t_\ell} |\mathrm{vac}\rangle \tag{104d}$$

$$= \left(\frac{1}{\sqrt{\ell}} \sum_{i=1}^{\ell} (-1)^{y\cdot s_i}\right) \widehat{a}^\dagger_{s_1} \ldots \widehat{a}^\dagger_{s_\ell} |\mathrm{vac}\rangle \ . \tag{104e}$$

Here, in the second line we applied the definition of the Fourier basis creation operators, $\widehat{a}^\dagger_s = \frac{1}{\sqrt{N}} \sum_t (-1)^{s\cdot t} \widetilde{a}^\dagger_t$, and in the third line we applied the commutation relations between the creation and annihiliation operators. Since we have shown that $\widetilde{\mathbf{G}}_y$ is diagonal in the position basis, acting by $\widetilde{\mathbf{G}}^2_y$ simply squares the coefficient, so recalling the definition of $\gamma^{(S)}_y$ (defined in eq. (28)) completes the proof. □

As an immediate corollary, we have that within the subspace of ℓ bosons, both $\widetilde{\mathbf{G}}_y$ and $\widetilde{\mathbf{H}}_y$ are diagonal in the position Fock basis, and that the operator that maps $|\mathrm{tt}_S\rangle$ to $\gamma^{(S)}_y |\mathrm{tt}_S\rangle$ is in fact equal to $\widetilde{\mathbf{G}}^2_y = \widetilde{\mathbf{H}}_y + \mathrm{id}$. The formulation of this map in terms of both the single- and double-momentum hopping operators will be convenient. In some cases, one formulation will be more useful than the other.

The double-momentum hopping operator is particularly nice to analyze. Recall our definition of Con_r as the projector onto states that are r-condensates. Con_r is a projector diagonal in the momentum Fock basis and with the projector capturing all momentum Fock basis states that have a total of ℓ bosons and at least $\ell - r$ bosons in the 0-momentum mode. Then, it is easy to observe that

Fact 11.5 (Hopping operators preserve condensates)**.** *For any $r \geq 0$ and $k \geq 0$,*

$$(\mathrm{id} - \mathsf{Con}_{r+2k}) \cdot \widetilde{\mathbf{H}}_{y_k} \ldots \widetilde{\mathbf{H}}_{y_1} \cdot \mathsf{Con}_r = 0. \tag{105}$$

Second, for any polynomial $p : \mathbb{R}^{2^n} \to \mathbb{R}$, let M_p be the operator $p(\gamma^{(S)}_0, \ldots, \gamma^{(S)}_{2^n-1}) |\mathrm{tt}_S\rangle\langle\mathrm{tt}_S|$. Then,

$$\left(\mathrm{id} - \mathsf{Con}_{r+2\deg p}\right) \cdot M_p \cdot \mathsf{Con}_r = 0. \tag{106}$$

Proof. The first equation follows as each term of $\widetilde{\mathbf{H}}_y$ can at most move 2 bosons from the 0-momentum mode. For the second equation, observe that the action M_p can be expressed as

$$p\left(\widetilde{\mathbf{G}}^2_0, \ldots, \widetilde{\mathbf{G}}^2_{2^n-1}\right) = p\left(\widetilde{\mathbf{H}}_0 + \mathrm{id}, \ldots, \widetilde{\mathbf{H}}_{2^n-1} + \mathrm{id}\right) \tag{107}$$

due to eq. (104) which can then be expressed as a linear combination of terms of the form $\widetilde{\mathbf{H}}_{y_{\leq \deg p}} \ldots \widetilde{\mathbf{H}}_{y_1}$. Combining with the first equation completes the proof. □

Next, we prove norm bounds on the Hamiltonians $\widetilde{\mathbf{G}}_y$ and $\widetilde{\mathbf{H}}_y$ on the subspaces of ℓ bosons as well as the subspace Con_r. By the action of creation and annihilation operators on Fock states calculated in eq. (45a), on the subspace of ℓ bosons, $\|\widetilde{\mathbf{G}}_y\| \leq \sqrt{\ell}$ and, therefore, $\|\widetilde{\mathbf{H}}_y\| \leq \ell - 1$. Unfortunately, these norms are too large for our polynomial approximations to handle. Luckily, both norms are significantly smaller within the space of condensates.

Fact 11.6 (Norm of hopping operators for condensates)**.** *For $y \neq 0$ and all $r \geq 0$, the action of the single- and double-momentum hopping operator has small norm on the space of r-condensates. More specifically,*

$$\left\| \widetilde{\mathbf{G}}_y \cdot \mathsf{Con}_r \right\|_{\mathrm{op}} \leq \sqrt{r} + \sqrt{2+4r}\,, \text{ and} \tag{108a}$$

$$\left\| \widetilde{\mathbf{H}}_y \cdot \mathsf{Con}_r \right\|_{\mathrm{op}} \leq 9r + 9\,. \tag{108b}$$

Proof. Let us define

$$\widetilde{M}_y \stackrel{\text{def}}{=} \frac{1}{\sqrt{\ell}} \Big(\widetilde{a}_y^\dagger \widetilde{a}_0 + \widetilde{a}_0^\dagger \widetilde{a}_y \Big), \quad \text{and} \quad \widetilde{M}'_y \stackrel{\text{def}}{=} \frac{1}{\sqrt{\ell}} \sum_{x \notin \{0,y\}} \widetilde{a}_{x\oplus y}^\dagger \widetilde{a}_x\,, \tag{109}$$

and therefore bounding the norm of $\widetilde{\mathbf{G}}_y \cdot \mathsf{Con}_r$ can be achieved by bounding the norms of both $\widetilde{M}_y \cdot \mathsf{Con}_r$ and $\widetilde{M}'_y \cdot \mathsf{Con}_r$. By construction, $\widetilde{M}_y$ only depends on the bosons in the 0- and y-momentum modes and $\widetilde{M}'_y$ only depends on the other momentum modes. Observe that $\widetilde{\mathbf{G}}_y$ and $\widetilde{M}_y$ both commute with $\widetilde{n}_0 + \widetilde{n}_y$. This can be verified by direct calculation using the commutation relations given in eq. (49). Therefore, $\widetilde{M}_y$ preserves the sum of the number of bosons in the 0- and y-momentum modes. Equivalently, if we divide the bosonic subspace of ℓ bosons into orthogonal subspaces based on the value of $\widetilde{n}_0 + \widetilde{n}_y$, we find that $\widetilde{M}_y$ is block-diagonal with respect to this direct sum. Therefore, we can restrict our analysis to states where there are a total of L bosons combined in the 0- and y-momentum modes for $\ell - L \leq r$. Furthermore, since $\widetilde{M}_y$ only depends on the bosons in the 0- and y-momentum modes, for computing the spectral norm of $\widetilde{M}_y \cdot \mathsf{Con}_r$ we can restrict our analysis to states of a total of L bosons entirely contained in the 0- and y-momentum modes[12]. Every such state can be expressed as a superposition of states $|L-j, j\rangle$ representing the number of bosons in the 0- and y-momentum modes. Then, we can then calculate the effect of $\widetilde{M}_y$:

$$\sqrt{\ell} \cdot \widetilde{M}_y |\phi\rangle = \sum_{j=0}^{L} \left(\widetilde{a}_y^\dagger \widetilde{a}_0 + \widetilde{a}_0^\dagger \widetilde{a}_y \right) \alpha_j |L-j, j\rangle \tag{112a}$$

$$= \sum_{j=0}^{L} \alpha_j \left(\sqrt{(L-j)(j+1)} |L-j-1, j+1\rangle + \sqrt{(L-j+1)j} |L-j+1, j-1\rangle \right) \tag{112b}$$

$$\begin{aligned} &= \left(\sum_{j'=1}^{L} \alpha_{j'-1} \sqrt{(L-j'+1)j'} |L-j', j'\rangle \right) \\ &\quad + \left(\sum_{j''=0}^{L-1} \alpha_{j''+1} \sqrt{(L-j'')(j''+1)} |L-j'', j''\rangle \right) \end{aligned} \tag{112c}$$

[12] To formalize this, observe that the Hilbert space corresponding to Con_r can be factorized as

$$\mathsf{Con}_r = \bigoplus_{L=\ell-r}^{\ell} \left(\mathcal{H}_{\{0,y\}}^{(L)} \otimes \mathcal{H}_{\text{rest}}^{(\ell-L)} \right) \tag{110}$$

where $\mathcal{H}_{\{0,y\}}^{(L)}$ is the Fock space of the 0- and y-momentum modes restricted to containing exactly L bosons and $\mathcal{H}_{\text{rest}}^{(\ell-L)}$ is the Fock space for the remainder of the modes restricted to $\ell - L$ bosons. Since $\widetilde{M}_y$ commutes with $\widetilde{n}_0 + \widetilde{n}_y$ and does not depend on the state of the other modes, we can also factorize $\widetilde{M}_y$ as

$$\widetilde{M}_y = \bigoplus_{L=\ell-r}^{\ell} \left(\widetilde{M}_y^{(L)} \otimes \mathrm{id}_{\text{rest}} \right). \tag{111}$$

Therefore, we can restrict to analyzing only $\widetilde{M}_y^{(L)}$ for every L.

$$= \sum_{j=0}^{L} \left(\alpha_{j-1} \sqrt{(L-j+1)j} + \alpha_{j+1} \sqrt{(L-j)(j+1)} \right) |L-j, j\rangle \tag{112d}$$

Above, in the first line, we are applying the annihilation and creation operators using eq. (45a). Then we broke the expression into two terms, and did a change of variable $j' = j+1$ and $j'' = j-1$, with $\alpha_{-1} = \alpha_{L+1} = 0$. Then we relabeled $j'' \to j$ and $j' \to j$. Now we can bound the norm of $\widetilde{M}_y|\phi\rangle$, which is:

$$\left\| \widetilde{M}_y |\phi\rangle \right\|^2 = \frac{1}{\ell} \sum_{j=0}^{L} \left| \alpha_{j-1} \sqrt{(L-j+1)j} + \alpha_{j+1} \sqrt{(L-j)(j+1)} \right|^2 \tag{113a}$$

$$\leq \frac{2}{\ell} \sum_{j=0}^{L} \left| \alpha_{j-1} \sqrt{(L-j+1)j} \right|^2 + \left| \alpha_{j+1} \sqrt{(L-j)(j+1)} \right|^2 \tag{113b}$$

$$= \frac{2}{\ell} \sum_{j=0}^{L} |\alpha_{j-1}|^2 (L-j+1)j + |\alpha_{j+1}|^2 (L-j)(j+1) \tag{113c}$$

$$= \frac{2}{\ell} \left(\sum_{j'=0}^{L} |\alpha_{j'}|^2 (L-j')(j'+1) \right) + \frac{2}{\ell} \left(\sum_{j''=0}^{L} |\alpha_{j''}|^2 (L-j''+1)j'' \right) \tag{113d}$$

$$= \frac{2}{\ell} \sum_{j=0}^{L} |\alpha_j|^2 (L + 2jL - 2j^2) \tag{113e}$$

$$\leq \frac{2}{\ell} \sum_{j=0}^{L} |\alpha_j|^2 (L + 2rL) \tag{113f}$$

$$= 2L(1+2r)/\ell \tag{113g}$$

$$\leq 2 + 4r\,. \tag{113h}$$

Here, we first used that $|a+b|^2 \leq 2|a|^2 + 2|b|^2$, then broke up the sum and performed a change of variable $j' = j-1$ and $j'' = j+1$. Then we relabeled $j'' \to j$ and $j' \to j$, and used the bound $j \leq r$, $-j^2 \leq 0$, $L \leq \ell$, and finally that $\sum_j |\alpha_j|^2 = 1$. We also use the fact that $(L-j)(j+1) + (L-j+1)j = L - 2Lj - 2j^2$. This gives the bound on $\|\widetilde{M}_y \cdot \mathsf{Con}_r\|_{\mathrm{op}}$.

Similarly, to bound $\|\widetilde{M}'_y \cdot \mathsf{Con}_r\|_{\mathrm{op}}$ observe that since $\widetilde{M}'_y$ only depends on the modes outside of 0 and y and there are at most r bosons in said modes, then by concavity of the square root function and eq. (45a), the total norm is at most $r/\sqrt{\ell} \leq r/\sqrt{r} \leq \sqrt{r}$ for $r \leq \ell$. This proves the bound for $\|\widetilde{M}'_y \cdot \mathsf{Con}_r\|_{\mathrm{op}}$ for $r \leq \ell$. When $r \geq \ell$ we can use the fact that $\left\|\widetilde{\mathbf{G}}_y\right\| \leq \sqrt{\ell} \leq \sqrt{r}$.

To get the bound on $\|\widetilde{\mathbf{H}}_y \cdot \mathsf{Con}_r\|_{\mathrm{op}}$, we note that $\widetilde{\mathbf{G}}_y|\phi\rangle$ is an $(r+1)$-condensate, so we can apply the bound on $\|\widetilde{\mathbf{G}}_y \cdot \mathsf{Con}_r\|_{\mathrm{op}}$ twice:

$$\left\| \widetilde{\mathbf{H}}_y |\phi\rangle \right\| \leq 1 + \left\| \widetilde{\mathbf{G}}_y^2 |\phi\rangle \right\| \tag{114a}$$

$$\leq 1 + \left(\sqrt{r} + \sqrt{2+4r} \right) \left(\sqrt{r+1} + \sqrt{2+4(r+1)} \right) \tag{114b}$$

$$\leq 9r + 9\,. \tag{114c}$$

Here in the first line we use the fact that $\widetilde{\mathbf{H}}_y = \widetilde{\mathbf{G}}_y^2 - \mathrm{id}$ and the triangle inequality, and then take the crude upper bound of $9r + 9$. □

12 Proving the condensate property

In this section, we prove that, for any T-query algorithm, the purified post-query state of the algorithm is mostly supported on condensates, i.e., that most of the bosons (in the momentum basis) remain in the 0-momentum mode. Recall that the actual (purified) post-query state of an algorithm A is given by eq. (99),

$$|\psi_{\mathrm{PQ}}\rangle = \sum_{\mathbf{y},\mathbf{b}} \left[\left(A_{\mathbf{y},\mathbf{b}}\right)_{\mathsf{A}} \otimes \left(\sum_{\mathbf{x}} (K^{\mathbf{x}})_{\mathbf{y},\mathbf{b}} \otimes e_{\mathbf{x}} \left(\widetilde{\mathbf{G}}^2_{y_1^{\mathbf{b}}}, \ldots, \widetilde{\mathbf{G}}^2_{y^{\mathbf{b}}_{|\mathbf{b}|}} \right) \right) \right] A\,|0\rangle_{\mathsf{A}}\,|\mathrm{init}\rangle_{\mathsf{US}}\ . \tag{115}$$

Our primary goal in this section is to show that $\left\| |\psi_{\mathrm{PQ}}\rangle - \mathsf{Con}_R \cdot |\psi_{\mathrm{PQ}}\rangle \right\|$ is very small for some suitably large (but polynomial in n) R. However, it turns out that we will be able to prove a much stronger statement: that restricting the operators $\widetilde{\mathbf{G}}$ to Con_R subspace does not change the state much. Formally we define the condensate sandwiched post-query states as follows.

Definition 12.1 (Condensate sandwiched post-query states)**.** *For integers R, $r \geq 0$, and a quantum adversary A making T-queries to an oracle, define the (R, r)-sandwiched state to be,*

$$|\widetilde{\psi}_{R,r}\rangle \stackrel{\mathrm{def}}{=} \sum_{\mathbf{y},\mathbf{b}} \left[\left(A_{\mathbf{y},\mathbf{b}}\right)_{\mathsf{A}} \otimes \left(\sum_{\mathbf{x}} (K^{\mathbf{x}})_{\mathbf{y},\mathbf{b}} \otimes \prod_{i=1}^{|\mathbf{b}|} \mathsf{Con}_r \cdot e_{x_i} \left(\mathsf{Con}_R \cdot \widetilde{\mathbf{G}}^2_{y_i^{\mathbf{b}}} \cdot \mathsf{Con}_R \right) \cdot \mathsf{Con}_r \right) \right] A\,|0\rangle_{\mathsf{A}}\,|\mathrm{init}\rangle_{\mathsf{US}}\,. \tag{116}$$

It is clear from the definition that restricting $|\widetilde{\psi}_{R,r}\rangle$ to the Con_r subspace does not change the state at all, so showing that $|\widetilde{\psi}_{R,r}\rangle$ is close to $|\psi_{\mathrm{PQ}}\rangle$ will accomplish our goal. Our main theorem is the following.

Theorem 12.2 (Sandwiching theorem)**.** *For every $\iota > 0$, there exist integers $r = O(n^2 T^5 \ln^3(T) \ln^2(1/\iota))$ and $R = O(n^3 T^6 \ln^4(T) \ln^3(1/\iota))$, such that*

$$\left\| |\psi_{\mathrm{PQ}}\rangle - |\widetilde{\psi}_{R,r}\rangle \right\| \leq \iota. \tag{117}$$

The main technique used in the proof is a clever polynomial approximation to the exponential function. To see what we mean by polynomial approximation in this context, we define function post-query states as follows.

Definition 12.3 (Function post-query state)**.** *Let $\mathcal{A}$ be a T-query quantum algorithm that implements unitary A between queries. Given a family of functions $f_{\mathbf{x}} = f_{(x_1,\ldots,x_T)}$, we define the f-post-query state*

$$|\psi_f\rangle \stackrel{\mathrm{def}}{=} \sum_{\mathbf{y},\mathbf{b}} \left[\left(A_{\mathbf{y},\mathbf{b}}\right)_{\mathsf{A}} \otimes \left(\sum_{\mathbf{x}} (K^{\mathbf{x}})_{\mathbf{y},\mathbf{b}} \otimes f_{\mathbf{x}} \left(\widetilde{\mathbf{G}}^2_{y_1^{\mathbf{b}}}, \ldots, \widetilde{\mathbf{G}}^2_{y^{\mathbf{b}}_{|\mathbf{b}|}} \right) \right) \right] A\,|0\rangle_{\mathsf{A}}\,|\mathrm{init}\rangle_{\mathsf{US}}\,. \tag{118}$$

Evaluating the functions $f(X)$ for a diagonalizable matrix X happens by applying diagonalizing X and applying f to the diagonal entries. Intuitively, the function post-query state corresponds to the (not necessarily normalized) purified state of an algorithm querying the U oracle, if U_y was sampled according to $f(\gamma_y^{(S)})$. To relate these states to our original post-query state, we can write down the functions corresponding to $|\psi_{\mathrm{PQ}}\rangle$.

Definition 12.4 (Kraus operator eigenvalue functions)**.** *Define the pair of functions e_0 and e_1 as follows,*

$$e_0(\gamma) \stackrel{\text{def}}{=} 1 - \mathrm{e}^{-\kappa\gamma}, \tag{119a}$$

$$e_1(\gamma) \stackrel{\text{def}}{=} \sqrt{\mathrm{e}^{-\kappa\gamma}(2 - \mathrm{e}^{-\kappa\gamma})} = \sqrt{2}\mathrm{e}^{-\kappa\gamma/2}\sqrt{1 - \mathrm{e}^{-\kappa\gamma/2}}\,. \tag{119b}$$

Note that we can express the Kraus operators $E_x^{(y)}$ as $\sum_S e_x(\gamma_y^{(S)})\,|\mathrm{tt}_S\rangle\langle\mathrm{tt}_S|$ to simplify the expressions in eq. (91).

For all $B \in \mathbb{Z}_{\geq 0}$ and $\mathbf{x} = (x_1, \ldots, x_B)$, we also use the notational shorthand $e_{\mathbf{x}}(\gamma_1, \ldots, \gamma_B) \stackrel{\text{def}}{=} \prod_{i=1}^{B} e_{x_i}(\gamma_i)$. Using these definition, we can express the post-query state of any algorithm (eq. (99)) as the function state corresponding to $e_{\mathbf{x}}$,

$$|\psi_{\mathrm{PQ}}\rangle = |\psi_e\rangle\ . \tag{120}$$

We can now state the notion of polynomial approximation that we want.

Theorem 12.5 (Polynomial approximation of post-query state)**.** *For every $\iota > 0$, there exists a family of polynomials* $\mathrm{AKraus}_{\mathbf{x}}$ *(that implicitly depends on ι) in the operators $\widetilde{\mathbf{G}}_{y_i}^2$ with*

$$\deg(\mathrm{AKraus}_{\mathbf{x}}) = O(n^2 T^5 \ln(T) \ln^2(1/\iota))\,, \tag{121}$$

such that we can bound the distance between the true post-query state and the polynomial approximation by

$$\big\| |\psi_{\mathrm{PQ}}\rangle - |\psi_{\mathrm{AKraus}}\rangle \big\| \leq \iota\,. \tag{122}$$

12.1 Overview of the polynomial approximation

Our starting point for the polynomial approximation is the Taylor series approximation of the exponential. Roughly speaking, taking the d'th level truncation to the Taylor series is a good approximation for inputs up to $O(d)$, and has error that scales as $O(e^x)$ for x much larger than d. Since the best bound we have on the norm of $\widetilde{\mathbf{G}}_y$ is $\left\|\widetilde{\mathbf{G}}_y\right\|_{\mathrm{op}} \leq \sqrt{\ell}$, this would seem to imply that, if we want to apply the Taylor series approximation to the exponential functions in the post-query state, we must take the degree of the approximation to be very high.

However, we are not trying to approximate $\widetilde{\mathbf{G}}_y^2$ on all inputs, but rather only on the distribution of $\gamma_y^{(S)}$ induced by $|\mathrm{init}_S\rangle$. Since $\gamma_y^{(S)}$ is a sum of random ± 1 variables, standard concentration inequalities tell us that the probability that $\gamma_y^{(S)}$ is much larger than x scales as e^{-x}, which exactly matches the error in the Taylor series. Indeed, in the $T = 1$ case, this argument (sometimes called chaining) can be used to show that the Taylor series with degree $O(T \log(1/\varepsilon))$ yields a good polynomial approximation state to the original post-query state.

Going beyond $T = 1$, the state we are trying to approximate now has factors that look like

$$\exp\left(\widetilde{\mathrm{G}}^2_{y_1} + \ldots + \widetilde{\mathrm{G}}^2_{y_T}\right) . \tag{123}$$

From the same logic as before, if the factors $\gamma^{(S)}_{y_i}$ were independent of each other, we would expect that the Taylor series provide a good approximation to the function $\exp\left(\left(\widetilde{\mathrm{G}}_{y_1} + \ldots + \widetilde{\mathrm{G}}_{y_T}\right)/T\right)$, i.e., the average of the $\widetilde{\mathrm{G}}_{y_i}$ operators. This motivates the idea, originally proposed by Narayan [Nar24], of first writing $\exp(-z)$ as $(\exp(-z/T))^T$, and then taking the Taylor series approximation for $\exp(-z/T)$.

Let us first examine what happens when we approximate $\exp(-z/s)$. Instead of applying the tail bound to $\gamma^{(S)}_y$, our chaining argument can be applied to the operators $\widetilde{\mathrm{G}}^2_y$ themselves. We use the observation (proven in Fact 11.6) that when restricted to the r-condensate subspace, $\widetilde{\mathrm{G}}^2_y$ has bounded norm *and* does not increase the number of non-zero bosons by much. When we apply this to the definition of the Taylor series, we find that when we use the Taylor series to approximate $e^{-z/s} = \sum_{m=0}^{\infty} \frac{(-z)^m}{s^m m!}$ for an appropriately chosen s depending on r, the factor of $\frac{1}{s^m m!}$ beats down the norm of $\widetilde{\mathrm{G}}^2_y \cdot \mathsf{Con}_r$, allowing us to prove a strong bound on the approximation error.

To go from our approximation of $\exp(-z/s)$ to an approximation of $\exp(-z)$, we would need to take s copies of $\exp(-z/s)$, starting from the Con_0 subspace. However, every time we apply $\exp(-z/s)$, the number of non-zero bosons increases, and we must increase the degree of our polynomial approximation of the next copy of $\exp(-z/s)$. It turns out that if we take s copies, the final polynomial approximation degree actually explodes exponentially in s. To fix this, we notice that instead of simply taking s copies of $\exp(-z/s)$, we can actually approximate the function z^s with a Chebyshev polynomial of degree $\sqrt{s}$. Applying this approximation allows us to take fewer copies of $\exp(-z/s)$, which lets us arrive at a low-degree approximation to our post-query state.

Remark 12.6. *Using* $\mathrm{e}^z = (\mathrm{e}^{-z/s})^s$*, approximating* $\mathrm{e}^{-z/s}$ *via a truncated Taylor series and* z^s *by a truncated Chebyshev series is identical to a construction of a flat approximation due to Narayanan [Nar24]. Using this idea, Narayanan achieves an exponential improvement over the original flat approximation of the exponential function in Ref. [BLMT24]. An earlier version of our manuscript used Narayanan's result as a black box and combined the flat approximation with tail bounds on* $\gamma^{(S)}_y$ *in the position basis.*

Once we have established the polynomial approximations, the proof of Theorem 12.2 works by going back and forth between the polynomial approximation and the original Kraus operators. As the action of the $\widetilde{\mathrm{G}}^2_y$ Hamiltonians cannot move more than 2 bosons from the 0-momentum mode, we can move/insert the projectors Con_r virtually anywhere in the polynomial approximation for sufficiently large r. We first use this to move the projectors Con_r between the Kraus operators and then exploit that the strategy to prove Theorem 12.5 also works if we start in Con_r instead of Con_0 by choosing a degree that depends on r polynomially. Moreover, we find that the polynomial approximation remains valid after replacing $\widetilde{\mathrm{G}}^2_y$ by the sandwiched $\mathsf{Con}_R \cdot \widetilde{\mathrm{G}}^2_y \cdot \mathsf{Con}_R$ in a single Kraus operator. This observation roughly allows us to approximate $\mathsf{Con}_r \cdot e_x(\widetilde{\mathrm{G}}^2_y) \cdot \mathsf{Con}_r$ by a polynomial depending on r and ι, move Con_R into the argument for large enough R, and finally reinsert the original Kraus operator.

12.2 Polynomial approximations to the Kraus operators

We will introduce two polynomial approximations in this section, the first of which is the truncated Taylor expansion of the exponential, given below.

$$\mathrm{Taylor}_d(z) \stackrel{\text{def}}{=} \sum_{j=0}^{d} \frac{z^j}{j!}\,. \tag{124}$$

Then, we have the following lemma.

Lemma 12.7 (Truncated Taylor condensate approximation)**.** *Let W be an operator such that there exists a constant $M > 0$ such that for all integers $m > 0$, the action of W on the subspace of m-condensates is bounded in norm by Mm. Additionally, assume that W maps the subspace of m-condensates into the subspace of $m+2$-condensates for all m. I.e., $\forall m > 0$, W satisfies both of the following:*

$$\|\mathsf{Con}_m \cdot W \cdot \mathsf{Con}_m\|_{\mathrm{op}} \leq Mm \quad \textit{and,} \tag{125a}$$

$$W \cdot \mathsf{Con}_m = \mathsf{Con}_{m+2} \cdot W \cdot \mathsf{Con}_m = \mathsf{Con}_{m+2} \cdot W \cdot \mathsf{Con}_{m+2} \cdot \mathsf{Con}_m\,. \tag{125b}$$

Then, for all $r \geq 0$, $s \geq 3M$, $1/\mathrm{e} \geq \varepsilon > 0$, and $d \geq 4\ln(1/\varepsilon) + r$, we have

$$\left\| \left(\mathrm{Taylor}_d\left(-W/s\right) - \exp(-W/s)\right) \cdot \mathsf{Con}_r \right\|_{\mathrm{op}} \leq \varepsilon\,. \tag{126}$$

Proof. Expanding out the Taylor series for $\mathrm{e}^{-W/s}$, we have

$$\begin{aligned}
\left\| \left(\mathrm{Taylor}_d\left(-W/s\right) - \exp(-W/s)\right) \mathsf{Con}_r \right\|_{\mathrm{op}} &= \left\| \sum_{j=d+1}^{\infty} \frac{1}{j!} s^{-j} (-W)^j \cdot \mathsf{Con}_r \right\|_{\mathrm{op}} && \text{(127a)}\\
&\leq \sum_{j=d+1}^{\infty} \frac{1}{j!} s^{-j} \left\| W^j \cdot \mathsf{Con}_r \right\|_{\mathrm{op}} && \text{(127b)}\\
&= \sum_{j=d+1}^{\infty} \frac{1}{j!} s^{-j} \left\| \left(\prod_{k=j}^{1} \mathsf{Con}_{2k+r} W \mathsf{Con}_{2k+r} \right) \mathsf{Con}_r \right\|_{\mathrm{op}} && \text{(127c)}\\
&\leq \sum_{j=d+1}^{\infty} \frac{1}{j!} s^{-j} \prod_{k=j}^{1} \left\| \mathsf{Con}_{2k+r} W \mathsf{Con}_{2k+r} \right\|_{\mathrm{op}} && \text{(127d)}\\
&\leq \sum_{j=d+1}^{\infty} \frac{1}{j!} \left(\prod_{k=j}^{1} (k + r/2) \right) \left(\frac{M}{s} \right)^j && \text{(127e)}\\
&= \sum_{j=d+1}^{\infty} \binom{r/2 + j}{j} \left(\frac{M}{s} \right)^j, && \text{(127f)}
\end{aligned}$$

where we first used the triangle inequality, then used eq. (125b) to resolve eq. (127c), sub-multiplicativity of the operator norm in eq. (127d), and eq. (125a) for proving eq. (127e). Now, we bound eq. (127f). Consider

the following

$$S(d) \stackrel{\text{def}}{=} \sum_{j=d+1}^{\infty} \underbrace{\binom{r/2+j}{j}\left(\frac{M}{s}\right)^{j}}_{\stackrel{\text{def}}{=} w_j}. \tag{128}$$

We will bound this from above by a geometric series, which requires bounding the ratio of successive terms. We find

$$\frac{w_{j+1}}{w_j} = \left(\frac{M}{s}\right)\frac{\binom{r/2+j+1}{j+1}}{\binom{r/2+j}{j}} = \left(\frac{M}{s}\right)\frac{r/2+j+1}{j+1} = \left(\frac{M}{s}\right)\left(1+\frac{r}{2(j+1)}\right). \tag{129}$$

For $j \geq d$, this is at most

$$Q \stackrel{\text{def}}{=} \left(\frac{M}{s}\right)\left(1+\frac{r}{2(d+1)}\right). \tag{130}$$

Since $M/s \leq 1/3$ and $r/2(d+1) \leq 1/2$, $Q \leq 1/2$. Therefore, the series $S(d)$ is bounded by

$$S(d) = \sum_{j=d+1}^{\infty} a_j \leq w_{d+1}\sum_{k=0}^{\infty} Q^k = 2w_{d+1} = 2\cdot\binom{r/2+d+1}{d+1}\left(\frac{M}{s}\right)^{d+1} \leq 2^{r/2+d+2}\cdot 3^{-(d+1)}. \tag{131}$$

For $d \geq 4\ln(1/\varepsilon) + r$, this can be upper bounded by ε. Since $S(d)$ equals the error in eq. (127), this completes the proof. □

The second family of polynomials we consider is the family of Chebyshev polynomials of the first kind. These are typically denoted using the symbol T_k, but since this choice would overload the character T, we use the following notation,

$$\mathrm{Cheby}_k(z) = \cos(k\arccos(z)), \quad z \in [-1,1]. \tag{132}$$

The Chebyshev polynomials form a basis for the space of polynomials, and, therefore, there exist coefficients $a_k^{(s)}$ such that the function $z \mapsto z^s$ in the basis of Chebyshev polynomials is given by

$$z^s = \sum_{k=0}^{\infty} a_k^{(s)}\mathrm{Cheby}_k(z). \tag{133}$$

Denote the truncated Chebyshev expansion of x^s by

$$\mathrm{TCheby}_{s,d}(z) \stackrel{\text{def}}{=} \sum_{k=0}^{d} a_k^{(s)}\mathrm{Cheby}_k(z). \tag{134}$$

We will make use of the following facts about Chebyshev polynomials.

Fact 12.8 (Theorem 3.3 in [SV+14]). *Then, $|a_k^{(s)}| \leq 1$ for all $k \geq 0$ and*

$$\left|z^s - \mathrm{TCheby}_{s,d}(z)\right| \leq 2\cdot\exp\left(-\frac{d^2}{2s}\right) \text{ for all } x \in [-1,1]. \tag{135}$$

Fact 12.9 (Bounds on the coefficients of the Chebyshev polynomials [Nar24])**.** *For all $k \geq 0$, the coefficients of* Cheby_k *in the monomial basis are bounded by* $(1+\sqrt{2})^k$.

Then, we have the following polynomial approximation:

Lemma 12.10 (Flat approximations of exponential functions)**.** *Let W be a PSD operator $W \succeq 0$, satisfying the predicates of Lemma 12.7. Then, for any $\varepsilon > 0$, $r \geq 0$, there is a polynomial, which we call* $\mathrm{Flat}_\varepsilon = \mathrm{Flat}_{\varepsilon,M,r}$ *(but we drop the subscripts M, r), of degree at most* $100M\ln(1/\varepsilon)(r+1)$ *such that*

$$\left\|(\exp(-W) - \mathrm{Flat}_\varepsilon(W)) \cdot \mathsf{Con}_r\right\|_{\mathrm{op}} \leq \varepsilon\,. \tag{136}$$

Proof. We employ the fact that $W \succeq 0$ to ensure that the spectrum of $\exp(-W/w)$ is contained in $[0,1]$. To apply Lemma 12.7, fix $w \stackrel{\mathrm{def}}{=} 36M^2\ln(2/\varepsilon)$, and $d' \stackrel{\mathrm{def}}{=} \left\lceil\sqrt{72M^2\ln^2(2/\varepsilon)}\right\rceil$. Then,

$$\exp(-W)\cdot\mathsf{Con}_r = (\exp(-W/w))^w \cdot \mathsf{Con}_r \approx_{\varepsilon/2} \mathrm{TCheby}_{w,d'}(\exp(-W/w))\cdot\mathsf{Con}_r \tag{137}$$

in operator norm for all $d' \geq \sqrt{2w\ln(2/\varepsilon)}$ by Fact 12.8. Here, we use the following notation: $X \approx_\varepsilon X'$ is used to denote that $\|X - X'\|_{\mathrm{op}} \leq \varepsilon$. Now let b_k be the coefficient of the monomial z^k in the decomposition of the polynomial $\mathrm{TCheby}_{w,d'}(z)$—i.e.,

$$\mathrm{TCheby}_{w,d'}(z) = \sum_k b_k z^k\,. \tag{138}$$

For our choice of w and d', we have

$$d' = \left\lceil\sqrt{2w\ln(2/\varepsilon)}\right\rceil, \tag{139}$$

and thus we can bound for all $k \leq d'$,

$$\frac{k}{w} \leq \frac{d'}{w} \leq \frac{\sqrt{4w\ln(2/\varepsilon)}}{w}\,. \tag{140}$$

Now we fix $d \stackrel{\mathrm{def}}{=} (4\left((9M+1)\ln(1/\varepsilon) + \ln(2)\right) + r)$. By Lemma 12.7 that whenever

$$d \geq 4\ln\left(2\cdot\left(\max_k |b_k|\right)\cdot d'/\varepsilon\right) + r \text{ and } s \stackrel{\mathrm{def}}{=} \sqrt{\frac{w}{4\ln(2/\varepsilon)}} \geq 3M\,, \tag{141}$$

we have that

$$\mathrm{TCheby}_{w,d'}(\exp(-W/w))\cdot\mathsf{Con}_r = \sum_{k=1}^{d'} b_k\cdot\exp(-W/w)^k\cdot\mathsf{Con}_r \tag{142a}$$

$$= \sum_{k=1}^{d'} b_k\cdot\exp\left(-\frac{k}{w}W\right)\cdot\mathsf{Con}_r \tag{142b}$$

$$\approx_{\varepsilon/2} \sum_{k=1}^{d'} b_k\cdot\mathrm{Taylor}_d\left(-\frac{k}{w}W\right)\cdot\mathsf{Con}_r\,, \tag{142c}$$

Here, we recall that Taylor_d is the truncated Taylor series for the exponential. In the final step, we use the fact that W satisfies the conditions required of the operator, and recall that we chose $w = 36M^2 \ln(2/\varepsilon)$, combined with eq. (140), to satisfy the constraint on s required for Lemma 12.7. The next part of the proof is to show that d satisfies the required bound. We will take the following upper bound on d' for sufficiently small ε,

$$d' = \left\lceil \sqrt{72M^2 \ln^2(2/\varepsilon)} \right\rceil \leq 9M \ln(1/\varepsilon)\,. \tag{143}$$

Observe that, by definition,

$$b_k = \sum_{k'=k}^{d'} a_{k'}^{(s)} \cdot [\text{Coefficient of } z^{k'} \text{ in } \mathrm{Cheby}_{k'}(z)]\,, \tag{144}$$

and by Fact 12.9, for all k, the coefficients b_k can be bounded by

$$|b_k| \leq \left(|a_1^{(s)}| + \cdots + |a_{d'}^{(s)}| \right) \left(1 + \sqrt{2}\right)^{d'} \leq d'(2.5)^{d'}\,. \tag{145}$$

Taking the logarithm and applying the choice from eq. (143) for any $d' \geq 10$, we have that $\ln(|b_k|) \leq \ln\left(d'\right) + d' \ln(2.5)$, this is equivalent to

$$\begin{aligned} \ln(|b_k|) + \ln\left(d'\right) &\leq 2\ln\left(d'\right) + d' \ln(2.5)\,. && \text{(146a)} \\ &\leq 2\ln\left(d'\right) + 9M \ln(1/\varepsilon) \cdot 0.89 && \text{(146b)} \\ &\leq 9M \ln(1/\varepsilon)\,. && \text{(146c)} \end{aligned}$$

Plugging into eq. (141), we find that d satisfies the necessary lower bound. Now, take $\mathrm{Flat}_\varepsilon(z)$ to be the polynomial defined to be

$$\mathrm{Flat}_\varepsilon(z) \stackrel{\text{def}}{=} \sum_{k=1}^{d'} b_k \cdot \mathrm{Taylor}_d\left(-\frac{k}{w} \cdot z\right)\,. \tag{147}$$

So $\mathrm{Flat}_\varepsilon$ is a degree $\leq d$ polynomial. Where again, $d = (4\left((9M+1)\ln(1/\varepsilon) + \ln(2)\right) + r)$, and the coefficients b_k come from $\mathrm{TCheby}_{w,d'}$ for our choice of w and d' from the beginning of the proof. It has degree at most

$$\deg(\mathrm{Flat}_\varepsilon) \leq d \leq (4\left((9M+1)\ln(1/\varepsilon) + \ln(2)\right) + r) \leq 100M \ln(1/\varepsilon)(r+1)\,. \tag{148}$$

Here, we use loose upper bounds. This polynomial appears on the right-hand side of eq. (142c), then

$$\mathrm{TCheby}_{w,d'}(\exp(-W/w)) \cdot \mathsf{Con}_r \approx_{\varepsilon/2} \mathrm{Flat}(W) \cdot \mathsf{Con}_r\,. \tag{149}$$

The previous equation and eq. (137) combine via triangle inequality for the lemma statement. □

We can now finally consider the operators $\prod_{i=1}^{T} e_{x_i}(\widetilde{\mathrm{G}}_{y_i}^2)$, where, recall, $e_{x_i}(z)$ are terms

$$\begin{aligned} e_0(z) &\stackrel{\text{def}}{=} 1 - \mathrm{e}^{-\kappa z}, \quad \text{and} && \text{(150a)} \\ e_1(z) &\stackrel{\text{def}}{=} \sqrt{\mathrm{e}^{-\kappa z}(2 - \mathrm{e}^{-\kappa z})} = \sqrt{2}\mathrm{e}^{-\kappa z/2}\sqrt{1 - \mathrm{e}^{-\kappa z}/2}\,, && \text{(150b)} \end{aligned}$$

to match the definitions of the Kraus operators $E_0^{(y_i)}$ and $E_1^{(y_i)}$ (Corollary 11.3). Towards the proof of Theorem 12.2, we prove the following more general lemma:

Lemma 12.11. *Let $W_1, \ldots, W_T \succeq 0$ be PSD operators that each satisfy the predicates of Lemma 12.7 and pairwise commute. For any $r \geq 0$ and every $\mathbf{x} \in \{0,1\}^T$, there exists a multivariate polynomial $\mathrm{AKraus}_{\varepsilon,\mathbf{x}}$ such that*

$$\left\| \left(\prod_{i=0}^{T} e_{x_i}(W_i) - \mathrm{AKraus}_{\varepsilon,\mathbf{x}}(W_1, \ldots, W_T) \right) \cdot \mathsf{Con}_r \right\|_{\mathrm{op}} \leq \varepsilon \,. \tag{151}$$

Moreover, $\mathrm{AKraus}_{\varepsilon,\mathbf{x}}$ is of degree $O(MT^3 \ln(T) \ln^2(1/\varepsilon)(r+1))$.

Proof. We begin with a single PSD operator W that satisfies the predicates of Lemma 12.7. We replace the function $z \mapsto \sqrt{1 - z/2}$ in the definition of e_1 with the truncated binomial expansion[13]:

$$\mathrm{TSqrt}\,(z) \stackrel{\text{def}}{=} \sum_{k=0}^{d''} \binom{1/2}{k} \left(\frac{-1}{2} \right)^k z^k \,, \tag{152}$$

with the choice of $d'' = 4 + \frac{3}{2}(\ln(T) + \ln(1/\varepsilon))$. The nomenclature TSqrt$(\cdot)$ comes from it being a truncated Taylor expansion. Recall that the square-root function does not have a good Taylor series expansion about 0, which is why we build it from $z \mapsto \sqrt{1 - z/2}$. We define the following polynomial approximations to make progress towards approximation $e_0(z)$ and $e_1(z)$.

$$p_0(z) \stackrel{\text{def}}{=} 1 - z^2 \,, \tag{153a}$$

$$p_1(z) \stackrel{\text{def}}{=} \sqrt{2} z \cdot \mathrm{TSqrt}(z^2) \,. \tag{153b}$$

Then $e_x(z)$ should be approximated by $p_x(e^{-\kappa z/2})$ whenever TSqrt is a good approximation of $\sqrt{1 - z/2}$.

Claim 12.12. *It follows that*

$$\left\| \prod_{i=1}^{T} e_{x_i}(W_i) - \prod_{i=1}^{T} p_{x_i}\left(\exp(-\kappa W_i/2)\right) \right\|_{\mathrm{op}} \leq \varepsilon/2 \,. \tag{154}$$

Proof. Next, observe that by construction for PSD W,

$$p_0(\exp(-\kappa W/2)) = e_0(W). \tag{155}$$

The challenge is to prove a similar statement for p_1 and e_1. Here, we will achieve an approximation.

$$\left\| p_1(\exp(-\kappa W/2)) - e_1(W) \right\|_{\mathrm{op}} = \tag{156a}$$

$$\left\| \sqrt{2} \exp(-\kappa W/2) \left(\sqrt{1 - \frac{1}{2}\exp(-\kappa W)} - \mathrm{TSqrt}(\exp(-\kappa W)) \right) \right\|_{\mathrm{op}} \tag{156b}$$

$$\leq \sqrt{2} \left\| \exp(-\kappa W/2) \right\|_{\mathrm{op}} \sum_{k=d''}^{\infty} \left| \binom{1/2}{k} \right| \cdot \frac{1}{2^k} \cdot \left\| \exp(-\kappa W)^k \right\|_{\mathrm{op}} \tag{156c}$$

$$\leq \frac{\sqrt{2}}{2^{d''}} \sum_{k=0}^{\infty} \frac{1}{2^k} \tag{156d}$$

[13] The usual binomial expansion $(1+a)^r = \sum_k \binom{r}{k} a^k$ also extends to real-valued r. Here, we are using $r = 1/2$ and $a = -z/2$, and truncating to $k \leq d''$.

$$\leq \frac{\varepsilon}{4T}\,, \tag{156e}$$

where we only use the facts that $\left|\binom{1/2}{k}\right| \leq 1$ and $\left\|\left(\mathrm{e}^{-\kappa W}\right)^k\right\|_{\mathrm{op}} \leq 1$ as W is PSD. So far we have produced approximations of single terms. We now lift these approximations to products of terms using a hybrid argument. We define hybrids B_j:

$$B_j \stackrel{\mathrm{def}}{=} \prod_{k=T-j}^{T} e_{x_k}(W_k) \prod_{k=1}^{T-j-1} p_{x_i}(\exp(-\kappa W_k/2))\,. \tag{157}$$

The hybrids combine to yield a total bound of

$$\left\|\prod_{i=1}^{T} e_{x_i}(W_i) - \prod_{i=1}^{T} p_{x_i}\left(\exp(-\kappa W_i/2)\right)\right\|_{\mathrm{op}} \tag{158a}$$

$$\leq \sum_{j=1}^{T} \left\|B_j - B_{j-1}\right\|_{\mathrm{op}} \tag{158b}$$

$$\leq \sum_{j=1}^{T} \left\|e_{x_{T-j}}(W_{T-j}) - p_{x_{T-j}}(\exp(-\kappa W_{T-j}/2))\right\|_{\mathrm{op}} \prod_{k=1}^{T-j} \left\|p_{x_k}(\exp(-\kappa W_k/2))\right\|_{\mathrm{op}} \tag{158c}$$

$$\leq T \cdot \frac{\varepsilon}{4T} \cdot \left(1 + \frac{\varepsilon}{4T}\right)^{T-j} \tag{158d}$$

$$\leq \frac{\varepsilon}{2}, \tag{158e}$$

Here, we use the facts that $\left\|e_{x_k}(W_k)\right\|_{\mathrm{op}} \leq 1$, and that $e_0(W_i) = p_0(W_i)$, and from eq. (156) we know that $\left\|e_1(W_i) - p_1(W_i)\right\|_{\mathrm{op}} \leq \varepsilon/4T$. We also used that via the triangle inequality,

$$\left\|p_{x_k}(\exp(-\kappa W_k/2))\right\|_{\mathrm{op}} \leq \frac{\varepsilon}{4T} + \left\|e_{x_k}(W_k)\right\|_{\mathrm{op}} \leq 1 + \frac{\varepsilon}{4T}\,. \tag{159}$$

Lastly, we resolved eq. (158e) using

$$\left(1 + \frac{\varepsilon}{4T}\right)^{T-j} \leq \left(1 + \frac{\varepsilon}{4T}\right)^{T} \leq \mathrm{e}^{\frac{\varepsilon}{4T}T} \leq \mathrm{e}^{1/4} \leq 2\,. \tag{160}$$

□

We then observe that $\prod_{i=1}^{T} p_{x_i}(\exp(-\kappa W_i/2))$ is a polynomial in the variables $\exp(-\kappa W_k/2)$ of degree at most $3d''T$ due to T terms of the form $z\mathrm{TSqrt}(z^2)$ which will yield a degree of $(1+2d'') \leq 3d''$. For any collection of J many operators W_{i_j}, we can write

$$\prod_{j=1}^{J} \exp\left(-\kappa W_{i_j}/2\right) = \exp\left(-\kappa \sum_{j=1}^{J} W_{i_j}/2\right) \tag{161}$$

as by assumption, the W_i pairwise commute. Notice also that $\sum_{j=1}^{J} W_{i_j}$ satisfies eq. (125b) (because each term maps the subspace of m-condensates to the subspace of $m+2$-condensates) and by the triangle inequality, we have that

$$\left\|\mathsf{Con}_m \cdot \left(\frac{\kappa}{2} \sum_{j=1}^{J} W_{i_j}\right) \cdot \mathsf{Con}_m\right\|_{\mathrm{op}} \leq \kappa JMm \leq JMm\,. \tag{162}$$

In the last line, we dropped the κ as it is a constant < 1 (think $1/10$). Next, we take the univariate polynomials $p_x(z)$ defined in eq. (153) and write them in the monomial basis in terms of families of coefficients $\{c_k^{(0)}\}$ and $\{c_k^{(1)}\}$:

$$p_x(z) = \sum_{k=0}^{3d''} c_k^{(x)} z^k. \tag{163}$$

This lets us construct multivariate polynomials $\mathrm{AKraus}_{\varepsilon,\mathbf{x}}$ in terms of the $\mathrm{Flat}_\varepsilon$ polynomials built in Theorem 12.10. We define $\mathrm{AKraus}_{\varepsilon,\mathbf{x}}$ for $\mathbf{x} = (x_1, \dots, x_T)$ as

$$\mathrm{AKraus}_{\varepsilon,\mathbf{x}}(z_1, \dots, z_T) \stackrel{\text{def}}{=} \sum_{k_1,\dots,k_T=0}^{3d''} \left(\prod_{i=1}^{T} c_{k_i}^{(x_i)} \cdot \mathrm{Flat}_{\frac{\varepsilon}{2\cdot 2^T}} \left(-\kappa \sum_{i=1}^{T} k_i z_i / 2 \right) \right), \tag{164}$$

Using the fact that in the sum in the argument of $\mathrm{Flat}_{\frac{\varepsilon}{2\cdot 2^T}}$, we are summing T many operators with coefficients that are integers and range from 0 to $3d''$, the largest value of J which we need to consider in eq. (162) is upper bounded by $J \le 3d''T$.

The polynomials $\mathrm{AKraus}_{\varepsilon,\mathbf{x}}$ are designed to approximate the polynomials p_0 and p_1 in the exponential function when evaluated on operators $(W_1, \dots, W_T)$, which we prove in the following claim.

Claim 12.13. *It follows that*

$$\left\| \left(\prod_{i=1}^{T} p_{x_i}(\exp(-\kappa W_i/2)) - \mathrm{AKraus}_{\varepsilon,\mathbf{x}}(W_1, \dots, W_T) \right) \cdot \mathsf{Con}_r \right\|_{\mathrm{op}} \le \varepsilon/2\,. \tag{165}$$

Proof. We use that for any monomial $\prod_{j=1}^{J} x_{i_j}$, we can apply Theorem 12.10 to achieve

$$\prod_{j=1}^{J} \exp(-\kappa W_{i_j}/2) \cdot \mathsf{Con}_r \approx_{\frac{\varepsilon}{2\cdot 2^T}} \mathrm{Flat}_{\frac{\varepsilon}{2\cdot 2^T}} \left(\sum_{j=1}^{J} \kappa W_{i_j}/2 \right) \cdot \mathsf{Con}_r\,. \tag{166}$$

Moreover, we used that the absolute coefficients of $\prod_{i=1}^{T} p_{x_i}$ add up to at most 2^T. To see this, observe that the absolute coefficients of p_0 add up to 2 Moreover, all coefficients in $\mathrm{TSqrt}(z^2)$ are negative except the first, which equals 1. Therefore, the absolute coefficients of p_1 add up to

$$\sum_{k_i=0}^{3d''} \left| c_{k_i}^{(x_i)} \right| = \sqrt{2}(2 - \mathrm{TSqrt}(1)) \le \sqrt{2}(2 - \sqrt{1 - 1/2})) \le 2. \tag{167}$$

Finally, we can apply a triangle inequality to find

$$\left\| \left(\prod_{i=1}^{T} p_{x_i}(\exp(-\kappa W_i/2)) - \mathrm{AKraus}_{\varepsilon,\mathbf{x}}(W_1, \dots, W_T) \right) \cdot \mathsf{Con}_r \right\|_{\mathrm{op}} \tag{168a}$$

$$\le \sum_{k_1,\dots,k_T=0}^{3d''} \left| \prod_{i=1}^{T} c_{k_i}^{(x_i)} \right| \cdot \left\| \left(\mathrm{Flat}_{\frac{\varepsilon}{2\cdot 2^T}} \left(-\kappa \sum_{i=1}^{T} k_i W_i/2 \right) - \exp(-\kappa \sum_{i=1}^{T} k_i W_i/2) \right) \mathsf{Con}_r \right\| \tag{168b}$$

$$\le \sum_{k_1,\dots,k_T=0}^{3d''} \left| \prod_{i=1}^{T} c_{k_i}^{(x_i)} \right| \cdot \frac{\varepsilon}{2 \cdot 2^T} \tag{168c}$$

$$\leq \sum_{k_1,\ldots,k_T=1}^{3d''} \prod_{i=1}^{T} \left| c_{k_i}^{(x_i)} \right| \cdot \frac{\varepsilon}{2 \cdot 2^T} \tag{168d}$$

$$= \prod_{i=1}^{T} \sum_{k=0}^{3d''} \left| c_{k_i}^{(x_i)} \right| \cdot \frac{\varepsilon}{2 \cdot 2^T} \tag{168e}$$

$$\leq 2^T \frac{\varepsilon}{2 \cdot 2^T} \tag{168f}$$

$$\leq \frac{\varepsilon}{2} . \tag{168g}$$

□

Now we bound the distance of $\mathrm{AKraus}_{\varepsilon,\mathbf{x}}$ from the product of exponentials via a middleman:

$$\left\| \left(\prod_{i=0}^{T} e_{x_i}(W_i) - \mathrm{AKraus}_{\varepsilon,\mathbf{x}}(W_1, \ldots, W_T) \right) \cdot \mathsf{Con}_r \right\|_{\mathrm{op}} \tag{169a}$$

$$\leq \left\| \left(\prod_{i=0}^{T} e_{x_i}(W_i) - \prod_{i=0}^{T} p_{x_i}(\exp(-\kappa W_i/2)) \right) \cdot \mathsf{Con}_r \right\|_{\mathrm{op}} + \left\| \left(\prod_{i=0}^{T} p_{x_i}(\exp(-\kappa W_i/2)) - \mathrm{AKraus}_{\varepsilon,\mathbf{x}}(W_1, \ldots, W_T) \right) \cdot \mathsf{Con}_r \right\|_{\mathrm{op}} \tag{169b}$$

$$\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon , \tag{169c}$$

from a combination of Claim 12.12 and Claim 12.13. The overall degree of $\mathrm{AKraus}_{\varepsilon,\mathbf{x}}$ is then given by $J \leq 3d''T$ and $d'' = 4 + \frac{3}{2}(\ln(T) + \ln(1/\varepsilon))$ to yield

$$\deg(\mathrm{AKraus}_{\varepsilon,\mathbf{x}}) \leq T \deg \mathrm{Flat}_{\frac{\varepsilon}{2 \cdot 2^T}, JM, r} \tag{170a}$$

$$\leq 100 JMT(\ln(1/\varepsilon) + T + 1)(r+1) \tag{170b}$$

$$\leq 300 d'' M T^3 \ln(1/\varepsilon)(r+1) \tag{170c}$$

$$\leq O(MT^3 \ln(T) \ln^2(1/\varepsilon)(r+1)) , \tag{170d}$$

completing the proof. □

Theorem 12.5 will follow from the following lemma (Lemma 12.14) combined with Lemma 12.11 for the choice of $W_i = \widetilde{\mathbf{G}}^2_{y_i}$, where y_i corresponds to the query of the verification algorithm. The following lemma applies triangle inequalities to bound the difference in the overall states when the oracle action is lightly replaced.

Lemma 12.14. *Let $B_{\mathbf{x},\mathbf{y},\mathbf{b}}$ and $B'_{\mathbf{x},\mathbf{y},\mathbf{b}}$ be two families of (possibly unnormalized) operators acting on the* S *register. Define*

$$|\psi_B\rangle \stackrel{\mathrm{def}}{=} \sum_{\mathbf{y},\mathbf{b}} \left[\left(A_{\mathbf{y},\mathbf{b}}\right)_{\mathsf{A}} \otimes \left(\sum_{\mathbf{x}} (K^{\mathbf{x}})_{\mathbf{y},\mathbf{b}} \otimes B_{\mathbf{x},\mathbf{y},\mathbf{b}} \right) \right] A |0\rangle_{\mathsf{A}} |\mathrm{init}\rangle_{\mathsf{US}} \tag{171}$$

and $|\psi_{B'}\rangle$ similarly. Then for all suitably large n,

$$\| |\psi_B\rangle - |\psi_{B'}\rangle \| \leq 2^{2nT} \max_{\mathbf{x},\mathbf{y},\mathbf{b}} \left\| (B_{\mathbf{x},\mathbf{y},\mathbf{b}} - B'_{\mathbf{x},\mathbf{y},\mathbf{b}}) |\mathrm{init}_S\rangle \right\| . \tag{172}$$

Proof. We apply the triangle inequality twice (once for the sum over $\mathbf{y}$ and then for the sum over $\mathbf{x}$). We use that $\|K^{x_i}\|_{\mathrm{op}} \leq 2$ and that $\||y_i\rangle\langle y_i| \cdot A\|_{\mathrm{op}} \leq 1$ as well as the submultiplicativity of the Schatten ∞-norm.

$$
\begin{aligned}
&\| |\psi_B\rangle - |\psi_{B'}\rangle\| \\
&= \left\| \sum_{\substack{\mathbf{y}\in(\{0,1\}^n)^T \\ \mathbf{b}\in\{0,1\}^T}} \left[\left(A_{\mathbf{y},\mathbf{b}}\right)_{\mathsf{A}} \otimes \left(\sum_{\mathbf{x}\in\{0,1\}^{|\mathbf{b}|}} (K^{\mathbf{x}})_{\mathbf{y},\mathbf{b}} \otimes \left(B_{\mathbf{x},\mathbf{y},\mathbf{b}} - B'_{\mathbf{x},\mathbf{y},\mathbf{b}}\right) \right) \right] A|0\rangle_{\mathsf{A}} |\mathrm{init}\rangle_{\mathsf{US}} \right\| && (173a) \\
&\leq \sum_{\substack{\mathbf{y}\in(\{0,1\}^n)^T \\ \mathbf{b}\in\{0,1\}^T}} \left\| \left[A_{\mathbf{y},\mathbf{b}} \cdot A|0\rangle \otimes \left(\sum_{\mathbf{x}\in\{0,1\}^{|\mathbf{b}|}} (K^{\mathbf{x}})_{\mathbf{y},\mathbf{b}} |\perp\rangle^{\otimes 2^n} \otimes \left(B_{\mathbf{x},\mathbf{y},\mathbf{b}} - B'_{\mathbf{x},\mathbf{y},\mathbf{b}}\right) |\mathrm{init}_S\rangle \right) \right] \right\| && (173b) \\
&= \sum_{\substack{\mathbf{y}\in(\{0,1\}^n)^T \\ \mathbf{b}\in\{0,1\}^T}} \left[\left\|A_{\mathbf{y},\mathbf{b}} \cdot A|0\rangle\right\| \cdot \left\| \left(\sum_{\mathbf{x}\in\{0,1\}^{|\mathbf{b}|}} (K^{\mathbf{x}})_{\mathbf{y},\mathbf{b}} |\perp\rangle^{\otimes 2^n} \otimes \left(B_{\mathbf{x},\mathbf{y},\mathbf{b}} - B'_{\mathbf{x},\mathbf{y},\mathbf{b}}\right) |\mathrm{init}_S\rangle \right) \right\| \right] && (173c) \\
&\leq \sum_{\substack{\mathbf{y}\in(\{0,1\}^n)^T \\ \mathbf{b}\in\{0,1\}^T}} \left[\left\|A_{\mathbf{y},\mathbf{b}} \cdot A|0\rangle\right\| \cdot \sum_{\mathbf{x}\in\{0,1\}^{|\mathbf{b}|}} \left\| \left((K^{\mathbf{x}})_{\mathbf{y},\mathbf{b}} |\perp\rangle^{\otimes 2^n} \otimes \left(B_{\mathbf{x},\mathbf{y},\mathbf{b}} - B'_{\mathbf{x},\mathbf{y},\mathbf{b}}\right) |\mathrm{init}_S\rangle \right) \right\| \right] && (173d) \\
&= \sum_{\substack{\mathbf{y}\in(\{0,1\}^n)^T \\ \mathbf{b}\in\{0,1\}^T}} \left[\left\|A_{\mathbf{y},\mathbf{b}} \cdot A|0\rangle\right\| \cdot \left(\sum_{\mathbf{x}\in\{0,1\}^{|\mathbf{b}|}} \left\| (K^{\mathbf{x}})_{\mathbf{y},\mathbf{b}} |\perp\rangle^{\otimes 2^n} \right\| \cdot \left\| \left(\left(B_{\mathbf{x},\mathbf{y},\mathbf{b}} - B'_{\mathbf{x},\mathbf{y},\mathbf{b}}\right) |\mathrm{init}_S\rangle \right) \right\| \right) \right] && (173e) \\
&\leq 2^T \sum_{\substack{\mathbf{y}\in(\{0,1\}^n)^T \\ \mathbf{b}\in\{0,1\}^T \\ \mathbf{x}\in\{0,1\}^{|\mathbf{b}|}}} \left\| \left(B_{\mathbf{x},\mathbf{y},\mathbf{b}} - B'_{\mathbf{x},\mathbf{y},\mathbf{b}}\right) |\mathrm{init}_S\rangle \right\| && (173f) \\
&\leq 2^{3T+nT} \max_{\mathbf{y},\mathbf{b},\mathbf{x}} \left\| \left(B_{\mathbf{x},\mathbf{y},\mathbf{b}} - B'_{\mathbf{x},\mathbf{y},\mathbf{b}}\right) |\mathrm{init}_S\rangle \right\| && (173g) \\
&\leq 2^{2nT} \max_{\mathbf{y},\mathbf{x},\mathbf{b}} \left\| \left(B_{\mathbf{x},\mathbf{y},\mathbf{b}} - B'_{\mathbf{x},\mathbf{y},\mathbf{b}}\right) |\mathrm{init}_S\rangle \right\| . && (173h)
\end{aligned}
$$

Here, eq. (173b) and eq. (173d) are applications of the triangle inequality, eq. (173c) and eq. (173e) are applying that norm of a tensor product is the tensor product of the norms, and eq. (173f) comes from bounding $\left\|A_{\mathbf{y},\mathbf{b}} A|0\rangle\right\| \leq 1$, and $\|(K^{\mathbf{x}})_{y,b} |\perp\rangle^{\otimes 2^n}\| \leq 2^T$ using the bound on $\left\|K_{x_i}\right\|_{\mathrm{op}}$ and the submultiplicativity of the operator norm. □

Finally, we proof Theorem 12.5, which we re-state here.

Theorem (Polynomial approximation of post-query state, Theorem 12.5, restated)**.** *For every $\iota > 0$, there exists a family of polynomials* $\mathrm{AKraus}_{\mathbf{x}}$ *(that implicitly depends on ι) in the operators $\widetilde{\mathbf{G}}^2_{y_i}$ with*

$$\deg\left(\mathrm{AKraus}_{\mathbf{x}}\right) = O(n^2 T^5 \ln(T) \ln^2(1/\iota)), \tag{174}$$

such that we can bound the distance between the true post-query state and the polynomial approximation by

$$\left\| |\psi_{\mathrm{PQ}}\rangle - |\psi_{\mathrm{AKraus}}\rangle \right\| \leq \iota . \tag{175}$$

Proof. Since $|\text{init}_S\rangle = \mathsf{Con}_0 |\text{init}_S\rangle$, we can directly apply Lemma 12.11 for the operators $W_i = \widetilde{\mathbf{G}}^2_{y_i}$ with $r = 0$. Notice that the $\widetilde{\mathbf{G}}^2_{y_i}$ are PSD, pairwise commute as they are all diagonal in the position basis, and satisfy eq. (125b) as well as eq. (125a) with a constant M by Fact 11.6. The result then follows from Lemma 12.14 by choosing ε such that $\varepsilon = \iota/(2^{2nT})$. □

12.3 Extending the approximation to the sandwiched operator

In this subsection, we prove Theorem 12.2, which we first re-state.

Theorem (Sandwiching theorem, Theorem 12.2, restated)**.** *For every $\iota > 0$, there exist integers $r = O(n^2 T^5 \ln^3(T) \ln^2(1/\iota))$ and $R = O(n^3 T^6 \ln^4(T) \ln^3(1/\iota))$, such that*

$$\left\| |\psi_{\text{PQ}}\rangle - |\widetilde{\psi}_{R,r}\rangle \right\| \leq \iota. \tag{176}$$

Proof. The proof consists of two hybrid arguments on the S register. We will use the approximation in Lemma 12.11 in both hybrids. The first hybrid uses Lemma 12.11 starting from a 0-condensate to obtain an approximation $\text{AKraus}_{\varepsilon,\mathbf{x}_{T-j}}(\widetilde{\mathbf{G}}^2_{y_1}, \dots \widetilde{\mathbf{G}}^2_{y_{T-j}})$ of the operators $\prod_{i=1}^{T-j} e_{x_i}(\widetilde{\mathbf{G}}^2_{y_i})$ restricted to Con_0. Here, we define the substring $\mathbf{x}_k = (x_k, \dots, x_1)$. Note that these approximation are for products of $e_{x_i}(\widetilde{\mathbf{G}}^2_{y_i})$ of varying lengths. By Lemma 12.11 with error $\varepsilon/2T$, we have that the degree of $\text{AKraus}_{\frac{\varepsilon}{2T},\mathbf{x}_{T-j}}$ is bounded by the following for all j,

$$\deg(\text{AKraus}_{\frac{\varepsilon}{2T},\mathbf{x}_{T-j}}) = O(T^3 \ln(T) \ln^2(2T/\varepsilon)) = O(T^3 \ln^3(T) \ln^2(1/\varepsilon)). \tag{177}$$

Now set $r = 2\deg(\text{AKraus}_{\frac{\varepsilon}{2T},\mathbf{x}}) \geq 2\max_j \deg(\text{AKraus}_{\frac{\varepsilon}{2T},\mathbf{x}_{T-j}})$, and define the following hybrids:

$$D_j \stackrel{\text{def}}{=} \prod_{i=T}^{T-j+1} \mathsf{Con}_r \cdot e_{x_i}(\widetilde{\mathbf{G}}^2_{y_i}) \prod_{i=1}^{T-j} e_{x_i}(\widetilde{\mathbf{G}}^2_{y_i}) |\text{init}_S\rangle \ . \tag{178}$$

Note that the first $T - j$ terms commute so the order of expansion of the product does not matter. Then because the initial state is a 0-condensate,

$$\left\| \left(\prod_{i=T}^{1} \mathsf{Con}_r \cdot e_{x_i}(\widetilde{\mathbf{G}}^2_{y_i}) \cdot \mathsf{Con}_r - \prod_{i=1}^{T} e_{x_i}(\widetilde{\mathbf{G}}^2_{y_i}) \right) |\text{init}_S\rangle \right\| = \|D_T - D_0\| \leq \sum_{j=1}^{T} \|D_j - D_{j-1}\|. \tag{179}$$

Then, we have

$$\begin{aligned}
\|D_j - D_{j-1}\| &\leq \left\| (\text{id} - \mathsf{Con}_r) \cdot \prod_{i=1}^{T-j} e_{x_i}(\widetilde{\mathbf{G}}^2_{y_i}) |\text{init}_S\rangle \right\| && \text{(180a)} \\
&\leq \frac{\varepsilon}{2T} + \left\| (\text{id} - \mathsf{Con}_r) \cdot \text{AKraus}_{\varepsilon,\mathbf{x}_{T-j}}(\widetilde{\mathbf{G}}^2_{y_1}, \dots, \widetilde{\mathbf{G}}^2_{y_{T-j}}) |\text{init}_S\rangle \right\| && \text{(180b)} \\
&\leq \frac{\varepsilon}{2T} + \left\| (\text{id} - \mathsf{Con}_r) \cdot \mathsf{Con}_r \cdot \text{AKraus}_{\varepsilon\mathbf{x}_{T-j}}(\widetilde{\mathbf{G}}^2_{y_1}, \dots, \widetilde{\mathbf{G}}^2_{y_{T-j}}) |\text{init}_S\rangle \right\| && \text{(180c)} \\
&\leq \frac{\varepsilon}{2T}. && \text{(180d)}
\end{aligned}$$

Here we use the fact that $\mathrm{AKraus}_{\varepsilon,\mathbf{x}_{T-j}}$ is a polynomial of degree at most $r/2$ in $\widetilde{\mathbf{G}}_y^2$ and therefore moves at most r bosons from the 0-momentum mode, so the state after applying $\mathrm{AKraus}_{\varepsilon,\mathbf{x}_{T-j}}$ to the initial state is an r-condensate. Next, since the condensate projectors commute, we have that for all $m \geq 0$ that

$$\left\|\mathsf{Con}_m \cdot \mathsf{Con}_R \cdot \widetilde{\mathbf{G}}_y^2 \cdot \mathsf{Con}_R \cdot \mathsf{Con}_m\right\|_{\mathrm{op}} = \left\|\mathsf{Con}_R \cdot \mathsf{Con}_m \cdot \widetilde{\mathbf{G}}_y^2 \cdot \mathsf{Con}_m \cdot \mathsf{Con}_R\right\|_{\mathrm{op}} \leq \left\|\mathsf{Con}_m \cdot \widetilde{\mathbf{G}}_y^2 \cdot \mathsf{Con}_m\right\|_{\mathrm{op}}. \quad (181)$$

Therefore, by Fact 11.6, the operator $W = \mathsf{Con}_R \cdot \widetilde{\mathbf{G}}_y^2 \cdot \mathsf{Con}_R$ satisfies eq. (125a) for all $m > 0$ with a constant $M = O(1)$. Moreover, it is easy to check that W is also PSD and satisfies eq. (125b) (again, because Con_R and Con_m commute). Hence, from Lemma 12.11 for $T = 1$ (here, T is the parameter in Lemma 12.11, not the number of queries as in the rest of the paper) and a constant $M = O(1)$, there exists a polynomial $\mathrm{AKraus}_{\frac{\varepsilon}{4T},x}$ such that:

$$\mathrm{AKraus}_{\frac{\varepsilon}{4T},x}(\mathsf{Con}_R \cdot \widetilde{\mathbf{G}}_y^2 \cdot \mathsf{Con}_R) \cdot \mathsf{Con}_r \approx_{\varepsilon/4T} e_x(\mathsf{Con}_R \cdot \widetilde{\mathbf{G}}_y^2 \cdot \mathsf{Con}_R) \cdot \mathsf{Con}_r, \text{ and} \quad (182a)$$

$$\mathrm{AKraus}_{\frac{\varepsilon}{4T},x}(\widetilde{\mathbf{G}}_y^2) \cdot \mathsf{Con}_r \approx_{\varepsilon/4T} e_x(\widetilde{\mathbf{G}}_y^2) \cdot \mathsf{Con}_r\,, \quad (182b)$$

where the degree of $\mathrm{AKraus}_{\frac{\varepsilon}{4T},x}$ will be $O((r+1)\ln(4T/\varepsilon))$. Since this holds for all $R \geq r$, it will hold for $R = r + 2\deg(\mathrm{AKraus}_{\frac{\varepsilon}{4T},x}) = r + O((r+1)\ln(4T/\varepsilon))$. We now proceed with the second hybrid. Let C_j be defined as

$$C_j \stackrel{\text{def}}{=} \prod_{i=T}^{T-j+1} \mathsf{Con}_r \cdot e_{x_i}(\mathsf{Con}_R \cdot \widetilde{\mathbf{G}}_{y_i}^2 \cdot \mathsf{Con}_R) \cdot \mathsf{Con}_r \cdot \prod_{i=T-j}^{0} \mathsf{Con}_r \cdot e_{x_i}(\widetilde{\mathbf{G}}_{y_i}^2) \cdot \mathsf{Con}_r\,. \quad (183)$$

Then $\|C_T - C_0\|$ equals

$$\left\|\prod_{i=T}^{1} \mathsf{Con}_r \cdot e_{x_i}(\mathsf{Con}_R \cdot \widetilde{\mathbf{G}}_{y_i}^2 \cdot \mathsf{Con}_R) \cdot \mathsf{Con}_r - \prod_{i=T}^{1} \mathsf{Con}_r \cdot e_{x_i}(\widetilde{\mathbf{G}}_{y_i}^2) \cdot \mathsf{Con}_r\right\|_{\mathrm{op}} \leq \sum_{j=1}^{T} \left\|C_j - C_{j-1}\right\|_{\mathrm{op}}. \quad (184)$$

Then, we find

$$\begin{aligned}
&\left\|C_j - C_{j-1}\right\|_{\mathrm{op}} && (185a)\\
&\quad\leq \left\|\mathsf{Con}_r \cdot e_{x_{T-j+1}}(\mathsf{Con}_R \cdot \widetilde{\mathbf{G}}_{y_{T-j+1}}^2 \cdot \mathsf{Con}_R) \cdot \mathsf{Con}_r - \mathsf{Con}_r \cdot e_{x_{T-j+1}}(\widetilde{\mathbf{G}}_{y_{T-j+1}}^2) \cdot \mathsf{Con}_r\right\|_{\mathrm{op}} && (185b)\\
&\quad\leq \left\|\left(e_{x_{T-j+1}}(\mathsf{Con}_R \cdot \widetilde{\mathbf{G}}_{y_{T-j+1}}^2 \cdot \mathsf{Con}_R) - \mathrm{AKraus}_{\frac{\varepsilon}{4T},x_{T-j+1}}(\mathsf{Con}_R \cdot \widetilde{\mathbf{G}}_{y_{T-j+1}}^2 \cdot \mathsf{Con}_R)\right) \cdot \mathsf{Con}_r\right\|_{\mathrm{op}}\\
&\qquad\qquad + \left\|\left(e_{x_{T-j+1}}\left(\widetilde{\mathbf{G}}_{y_{T-j+1}}^2\right) - \mathrm{AKraus}_{\frac{\varepsilon}{4T},x_{T-j+1}}\left(\widetilde{\mathbf{G}}_{y_{T-j+1}}^2\right)\right) \cdot \mathsf{Con}_r\right\|_{\mathrm{op}} && (185c)\\
&\qquad\qquad + \left\|\begin{array}{l}\mathsf{Con}_r \cdot \mathrm{AKraus}_{\frac{\varepsilon}{4T},x_{T-j+1}}\left(\widetilde{\mathbf{G}}_{y_{T-j+1}}^2\right) \cdot \mathsf{Con}_r\\ \quad -\mathsf{Con}_r \cdot \mathrm{AKraus}_{\frac{\varepsilon}{4T},x_{T-j+1}}\left(\mathsf{Con}_R \cdot \widetilde{\mathbf{G}}_{y_{T-j+1}}^2 \cdot \mathsf{Con}_R\right) \cdot \mathsf{Con}_r\end{array}\right\|_{\mathrm{op}}\\
&\quad\leq \frac{\varepsilon}{4T} + \frac{\varepsilon}{4T} + 0 && (185d)\\
&\quad\leq \frac{\varepsilon}{2T}\,, && (185e)
\end{aligned}$$

for any $R \geq r + 2\deg(\mathrm{AKraus}_{\frac{\varepsilon}{4T},x_{T-j}})$. Here, in eq. (185c), observe that when the polynomials $\mathrm{AKraus}_{\frac{\varepsilon}{4T},x}$ are being invoked, the state remains a $r + 2\deg(\mathrm{AKraus}_{\frac{\varepsilon}{4T}x_{T-j}}) \leq R$-condensate so the final term in the

3-fold expression is 0. We are also using the definitions of C_j and C_{j-1}, together with standard properties of the operator norm, and the polynomial approximations from eq. (182a) and Lemma 12.11. The result then follows from choosing $\varepsilon = \iota/(2^{2nT})$ and applying Lemma 12.14. We then upper bound $\text{polylog}(T) \leq O(T)$ for notational convenience. □

13 Proving the quasi-even property

In this section, we prove that the purified state of any verification algorithm making T-queries to the U oracle is also incredibly close to a $v/4$-quasi-even state, i.e., that the state is almost entirely supported on momentum Fock states that have at most $v/4$ modes occupied by an odd number of bosons. The proof follows the general framework of [HM23] for unconstrained search. At a high level, the authors first define a sequence of projectors $\Pi_0 \preceq \Pi_1 \preceq ...$, and then show that querying their oracle maps a state supported on Π_k to a state that it mostly supported to Π_k but with an extremely small support on Π_{k+1}.

Our proof will utilize their framework, with the projectors QE_o playing the role of the Π_k from [HM23]. The main goal of the section will be to argue that the double-hopping operator is very unlikely to create unpaired bosons. At a high level, the idea is to notice that the double-hopping operator $\widetilde{\mathbf{H}}_y$ picks two momentum modes x and x', and moves two bosons from modes x and x' to $x \oplus y$ and $x' \oplus y$, with a normalization that corresponds roughly to the square root of product of the number of bosons that occupy modes x, x', $x \oplus y$, and $x' \oplus y$. Since the "weight" of a pair of modes corresponds to the number of bosons in the modes, when we apply the double-hopping operator to a condensate, most of the mass of the post-operation state corresponds to moving bosons in or out of the 0-momentum mode. This mostly preserves the "oddness" of the quasi-even condensate.

A subtle difference between our analysis is the fact that the oracle from [HM23] can only output states from Π_{k+1} when given a state in Π_k as input, which means that concluding the proof only requires a simple inductive argument. Oracle queries in our setting will roughly correspond to applying $\mathrm{e}^{-\kappa\widetilde{\mathbf{H}}_y}$ for some y. While $\widetilde{\mathbf{H}}_y$ itself only "moves" 4 bosons, the Taylor series of $\mathrm{e}^{-\kappa\widetilde{\mathbf{H}}_y}$ includes arbitrarily high powers of $\widetilde{\mathbf{H}}_y$, although we can bound these tails using a convenient expansion of the exponential from perturbation theory.

13.1 The double-hopping operator is almost always paired on condensates

The central premise of this next subsection is that the action of the double-hopping operator on condensates is dominated by the action of hopping two bosons into or out of the 0-momentum mode. This is because the action of a creation or annihilation operator on a Fock state is proportional to the number of bosons in said mode. Since almost all the bosons are in the 0-momentum mode, it follows that the behavior of the 0-momentum mode dominates the actions. However, it is important to note that ignoring the other modes would only produce an approximation that would be too coarse for the sampling probability upper bounds we need to achieve. Nevertheless, we can show that on a quasi-even condensate, the action will, with almost certainty, produce another quasi-even condensate where the "oddness" is unlikely to be changed. We emphasize that the "invariance" we prove is only true if the original state is a condensate. To see

otherwise, consider a state where the ℓ bosons are paired up in $\ell/2$ distinct momentum modes. Then the action of a double-hopping operator will almost certainly increase the "oddness" of the state by 2.

Concretely, in this subsection, we show that applying one double-hopping operator to a quasi-even condensate produces another quasi-even condensate, where the oddness is unlikely to be changed. This section captures the main idea behind the proof of the quasi-even property.

We first define the y-double difference set of a tuple w and prove a bound on the number of elements of the double-difference set. For simplicity, we will use the following notation:

$$\Delta^{(y)}_{x,x'} \stackrel{\text{def}}{=} 1_{x\oplus y} + 1_{x'\oplus y} - 1_x - 1_{x'} \in \mathbb{Z}^{2^n}. \tag{186}$$

With this, we define the double difference set.

Definition 13.1 (Double difference set)**.** *Given $y \in \{0,1\}^n$ and $y \neq 0^n$, define the y-double difference set of a tuple w as follows.*

$$\mathrm{diff}^2_y(w) \stackrel{\text{def}}{=} \left\{ u \in \mathbb{Z}^{2^n}_{\geq 0} \;:\; \exists\,(x,x') \text{ with } x \notin \{x', x'\oplus y\} \text{ and } w = u + \Delta^{(y)}_{x,x'} \right\}. \tag{187}$$

Furthermore, when $u \in \mathrm{diff}^2_y(w)$, we write $(x,x') = \mathrm{diff}^2_y(u,w)$ to represent the choice of x, x' that map from w to u. Without loss of generality, we take x to be the lower of the two in lexicographical ordering (since $x \neq x'$ by assumption).

We make the following observation.

Claim 13.2. *For every (R,o)-quasi-even condensate u, and every $y \in \{0,1\}^n$ and $y \neq 0^n$,*

$$\left|\mathrm{diff}^2_y(u)\right| \leq (R+1)^2. \tag{188}$$

Proof. Since u is a condensate with at most R bosons not in the 0-momentum mode, there are at most $R+1$ ways to choose x and $R+1$ was to choose x' so that the resulting tuple has non-negative entries after adding $\Delta^{(y)}_{x,x'}$, because they have to be taken from modes of u that are strictly positive. We will typically take the crude upper bound of $(R+1)^2 \leq 2R^2$ for suitably large R. □

Recall the definition of $\mathsf{QEC}_{r,o}$ as the projector onto momentum Fock states that are $\leq r$ condensates with $\leq o$ many non-zero modes having an odd number of bosons. And Con_r and QE_o are the projectors only ensuring the former and latter properties, respectively. Lastly, $\mathsf{QEC}_{r,=o}$ and $\mathsf{QE}_{=o}$ are the projectors requiring exactly o many non-zero modes.

Lemma 13.3. *For every $R \in [\ell]$, $y \in \{0,1\}^n$, with $y \neq 0^n$, the following holds.*

$$\left\| \sum_{o\in[R]} (\mathrm{id} - \mathsf{QE}_{=o}) \cdot \widetilde{\mathbf{H}}_y \cdot \mathsf{QEC}_{(R,=o)} \right\|_{\mathrm{op}} \leq \frac{R^5}{\sqrt{\ell}}. \tag{189}$$

Proof. We argue that the operator norm is small by showing that for every input state $|\psi\rangle$, the norm of the state shrinks by an appropriate factor. Since there are at most R choices of o that we are considering, it will suffice to bound the norm of $(\mathrm{id} - \mathsf{QE}_o)\cdot\widetilde{\mathbf{H}}_y|\psi\rangle$ for a state $|\psi\rangle$ supported on $\mathsf{QEC}_{(R,=o)}$ for a fixed o and

then apply the Cauchy-Schwarz inequality at the end. We can write $|\psi\rangle$ as a superposition of momentum Fock states $|u\rangle$, where u is a $(R, = o)$-quasi-even condensate, as follows,

$$|\psi\rangle = \sum_{u \in \mathsf{QEC}_{(R,=o)}} \alpha_u |u\rangle \ . \tag{190}$$

Expanding out the definition of the double-hopping operator, after applying $\widetilde{\mathbf{H}}_y$, we have the following state.

$$\widetilde{\mathbf{H}}_y |\psi\rangle = \sum_{u \in \mathsf{QEC}_{(R,=o)}} \alpha_u \left(\frac{1}{\ell} \sum_{x,x' \in \{0,1\}^n} \sqrt{u_x u_{x'} (u_{x\oplus y} + 1)(u_{x'\oplus y} + 1)} \, |u + \Delta^{(y)}_{x,x'}\rangle \right) . \tag{191}$$

Applying the projector onto states that have $o' \neq o$ many odd entries, will remove all terms in the sum that correspond to $x = x'$ or $x = x' \oplus y$, as these never change the number of odd indices. Thus, we can re-group the terms and use the definition of the double difference set to get the following.

$$\begin{aligned}
&(\mathrm{id} - \mathsf{QE}_{=o}) \cdot \widetilde{\mathbf{H}}_y |\psi\rangle && \text{(192a)}\\
&= \sum_{u \in \mathsf{QEC}_{(R,=o)}} \left(\frac{\alpha_u}{\ell} \sum_{x,x' \in \{0,1\}^n} \delta(u + \Delta^{(y)}_{x,x'} \notin \mathsf{QE}_{=o}) \sqrt{u_x u_{x'} (u_{x\oplus y} + 1)(u_{x'\oplus y} + 1)} \, |u + \Delta^{(y)}_{x,x'}\rangle \right) && \text{(192b)}\\
&= \sum_{u \in \mathsf{QEC}_{(R,=o)}} \frac{\alpha_u}{\ell} \sum_{\substack{x,x' \in \{0,1\}^n \\ x \notin \{x', x'\oplus y\}}} \delta(u + \Delta^{(y)}_{x,x'} \notin \mathsf{QE}_{=o}) \sqrt{u_x u_{x'} (u_{x\oplus y} + 1)(u_{x'\oplus y} + 1)} \, |u + \Delta^{(y)}_{x,x'}\rangle && \text{(192c)}\\
&= \frac{1}{\ell} \sum_{w \in \mathsf{QEC}_{(R+2,o+2)}} \delta(w \notin \mathsf{QE}_{=o}) \left(\sum_{\substack{u \in \mathsf{QEC}_{(R,=o)} \\ u \in \mathrm{diff}^2_y(w) \\ (x,x') = \mathrm{diff}^2_y(u,w)}} 2\alpha_u \sqrt{u_x u_{x'} (u_{x\oplus y} + 1)(u_{x'\oplus y} + 1)} \right) |w\rangle \ . && \text{(192d)}
\end{aligned}$$

Here, we first applied the definition of $\widetilde{\mathbf{H}}_y$ and used the fact that $\mathsf{QE}_{=o}$ is diagonal in the momentum Fock basis. Then we use the fact that whenever $x = x'$ or $x' \oplus y$, the number of odd indices is preserved, which means we can remove those terms from the sum. Then we regrouped terms that have the same value of $u + \Delta^{(y)}_{x,x'}$, noting that every pair $(x, x') = \mathrm{diff}^2_y(u, w)$ appears twice in the sum over x and x' in $\widetilde{\mathbf{H}}_y$, which gives us the multiplicative factor of 2.

We now bound the squared norm of this state, using the fact that the $|w\rangle$ are orthogonal to each other, and the number of u in $\mathrm{diff}^2_y(w)$ is bounded.

$$\begin{aligned}
&\left\| (\mathrm{id} - \mathsf{QE}_{=o}) \cdot \widetilde{\mathbf{H}}_y |\psi\rangle \right\|^2 && \text{(193a)}\\
&= \frac{1}{\ell^2} \left\| \sum_{w \in \mathsf{QEC}_{(R+2,o+2)}} \delta(w \notin \mathsf{QE}_{=o}) \left(\sum_{\substack{u \in \mathsf{QEC}_{(R,=o)} \\ u \in \mathrm{diff}^2_y(w) \\ (x,x') = \mathrm{diff}^2_y(u,w)}} 2\alpha_u \sqrt{u_x u_{x'} (u_{x\oplus y} + 1)(u_{x'\oplus y} + 1)} \right) |w\rangle \right\|^2 && \text{(193b)}
\end{aligned}$$

$$= \frac{4}{\ell^2} \sum_{w \in \mathsf{QEC}_{(R+2,o+2)}} \delta(w \notin \mathsf{QE}_{=o}) \left| \sum_{\substack{u \in \mathsf{QEC}_{(R,=o)} \\ u \in \mathrm{diff}_y^2(w) \\ (x,x')=\mathrm{diff}_y^2(u,w)}} \alpha_u \sqrt{u_x u_{x'} (u_{x\oplus y}+1)(u_{x'\oplus y}+1)} \right|^2 \tag{193c}$$

$$\leq \frac{4}{\ell^2} \sum_{w \in \mathsf{QEC}_{(R+2,o+2)}} \sum_{\substack{u \in \mathsf{QEC}_{(R,=o)} \\ u \in \mathrm{diff}_y^2(w)}} 2R^2 \cdot u_x u_{x'} (u_{x\oplus y}+1)(u_{x'\oplus y}+1) \cdot |\alpha_u|^2 \tag{193d}$$

$$\leq \frac{8}{\ell^2} \sum_{w \in \mathsf{QEC}_{(R+2,o+2)}} \sum_{u \in \mathrm{diff}_y^2(w)} \ell R^5 \cdot |\alpha_u|^2 \tag{193e}$$

$$\leq \frac{8}{\ell^2} \sum_u |\alpha_u|^2 \ell R^5 \left| \mathrm{diff}_y^2(u) \right| \tag{193f}$$

$$\leq \frac{16R^7}{\ell} . \tag{193g}$$

Here, in the first line, we expand the definition of the state from eq. (192d). Then we use the definition of the squared norm with the fact that $|w\rangle$ are orthogonal to each other. Then we use Cauchy-Schwarz to bound the square of the sum by $2R^2$ times the sum of the squares, as there are at most $(R+3) \leq 2R^2$ elements of $\mathrm{diff}_y^2(w)$, applying Theorem 13.2 to a $R+2$ condensate. In the second to last line, we switch the order of the sums and use the fact that diff_y^2 is reflexive — i.e., $w \in \mathrm{diff}_y^2(u) \Leftrightarrow u \in \mathrm{diff}_y^2(w)$. Then, we again use the fact that there are at most $2R^2$ elements in $\mathrm{diff}_y^2(u)$ and that the $|\alpha_u|^2$ sum to 1. Going from eq. (193d) to the next line, we note that by the definition of diff_y^2, for every $(x, x') \in \mathrm{diff}_y^2(w)$, at most one of x, x', $x \oplus y$ and $x' \oplus y$ can be 0 since $y \neq 0^n$. Therefore, at least three of the entries u_x, $u_{x'}$, $u_{x\oplus y}+1$ and $u_{x'\oplus y}+1$ can be bounded from above by R, the number of non-zero bosons, and the fourth one can be bounded by ℓ, the total number of bosons. Therefore, for all $x, x' \in \mathrm{diff}_y^2(u, w)$, $u_x u_{x'}(u_{x\oplus y}+1)(u_{x'\oplus y}+1) \leq \ell R^3$.

Taking the square root of this expression gives us the following bound for all o.

$$\left\| (\mathrm{id} - \mathsf{QE}_{=o}) \cdot \widetilde{\mathbf{H}}_y \cdot \mathsf{QEC}_{(R,=o)} \right\|_{\mathrm{op}} \leq \frac{4R^{7/2}}{\sqrt{\ell}} . \tag{194}$$

Since o can only be as large as R, with one more application of the triangle inequality, we can bound the norm of the sum over all o from 0 to R by $4R^{9/2}/\sqrt{\ell}$. To make the numbers cleaner, we take $R^5/\sqrt{\ell}$ as an upper bound. □

We note that the above bound holds for $\widetilde{\mathbf{H}}_y$, but for the next section we will want to apply it to $\widetilde{\mathbf{G}}_y = \widetilde{\mathbf{H}}_y + \mathrm{id}$, which is the positive semi-definite operator that corresponds to applying $\gamma_y^{(S)}$ in the position Fock basis.

Corollary 13.4. *For all $R \in [\ell]$ and $y \in \{0,1\}^n$ with $y \neq 0^n$, the following holds*

$$\left\| \sum_{o \in [R]} (\mathrm{id} - \mathsf{QE}_{=o}) \cdot \widetilde{\mathbf{G}}_y^2 \cdot \mathsf{QEC}_{(R,=o)} \right\|_{\mathrm{op}} \leq \frac{R^5}{\sqrt{\ell}} . \tag{195}$$

Proof. Using, $\widetilde{\mathbf{G}}_y^2 = \widetilde{\mathbf{H}}_y + \mathrm{id}$, we can apply Lemma 13.3 in a straight-forward way as follows.

$$\left\| \sum_{o\in[R]} (\mathrm{id} - \mathsf{QE}_{=o}) \cdot \widetilde{\mathbf{G}}_y^2 \cdot \mathsf{QEC}_{(R,=o)} \right\|_{\mathrm{op}} \tag{196a}$$

$$= \left\| \sum_{o\in[R]} (\mathrm{id} - \mathsf{QE}_{=o}) \cdot \widetilde{\mathbf{H}}_y \cdot \mathsf{QEC}_{(R,=o)} + \sum_{o\in[R]} (\mathrm{id} - \mathsf{QE}_{=o}) \cdot \mathsf{QE}_{=o} \cdot \mathrm{Con}_R \right\|_{\mathrm{op}} \tag{196b}$$

$$= \left\| \sum_{o\in[R]} (\mathrm{id} - \mathsf{QE}_{=o}) \cdot \widetilde{\mathbf{H}}_y \cdot \mathsf{QEC}_{(R,=o)} \right\|_{\mathrm{op}} \tag{196c}$$

$$\leq \frac{R^5}{\sqrt{\ell}}. \tag{196d}$$

Here, we simply expand out the definition of $\widetilde{\mathbf{G}}_y^2$ and use the fact that $\mathsf{QE}_{=o}$ is a projector and therefore $(\mathrm{id} - \mathsf{QE}_{=o})\mathsf{QE}_{=o} = 0$. □

13.2 Dyson series expansion of the exponential

The previous section demonstrates that applying the double-hopping operator on a quasi-even condensate is likely to preserve the number of odd entries. Unfortunately, although applying the double-hopping operator does not yield a state with high overlap on states with $\neq o$ odd entries, it does potentially increase the norm of the input state within the $= o$ subspace. This is problematic to a naive inductive argument because after applying many double-hopping operators, one might have to scale the bound from Lemma 13.3 by the growing norm of the state.

On the other hand, we expect that this is not a fundamental problem, because the operators that the algorithm applies are unitary, and even more so, operators of the form $\exp(-\widetilde{\mathbf{G}}_y^2)$ clearly have bounded operator norm. In this section, we show how to apply the Dyson series for the exponential function to lift Lemma 13.3 to certain functions of the double-hopping operator.

Fact 13.5 (Application of Duhamel's principle)**.** *The following identity holds*[a] *for all matrices A and V, and $t \geq 0$,*

$$\exp(-t \cdot (A + V)) = \exp(-t \cdot A) - \int_{0\leq s\leq t} \mathrm{d}s \, \exp(-(t-s) \cdot (A+V)) \cdot V \cdot \exp(-s \cdot A) \,. \tag{197}$$

Applying this identity repeatedly gives us the following Dyson series for a small perturbation to A.

Fact 13.6 (Dyson's formula for the exponential function)**.** *For all matrices A and V and $\kappa \geq 0$, we can express* $\exp(-\kappa \cdot (A+V)) - \exp(-\kappa \cdot A)$ *as*

$$\sum_{k=1}^{\infty} (-1)^k \int_{0\leq s_1\leq \ldots \leq s_k \leq \kappa} \mathrm{d}\mathbf{s} \, \exp(-(\kappa - s_k)A) \cdot \underbrace{V \cdot \exp(-(s_k - s_{k-1})A) \cdot V \ldots V}_{k \text{ times}} \exp(-s_1 A) \,. \tag{198}$$

The goal of Dyson's formula is to expand the exponential $\exp(-\kappa \cdot (A+V))$ in such a way that the expansion only includes products of the term V and exponentials of A. The use case is when V is a small perturbation applied with A, and Dyson's formula allows us to expand the exponential and apply bounds we know about V. In mathematical physics, A is sometimes called the "free" part of the Hamiltonian, and V the "interacting" part.

[a] To prove this statement, construct a function f such that $f(0) = \mathrm{e}^{-tA}$ and $f(t) = \mathrm{e}^{-t(A+V)}$. Then the proof follows by observing $f(t) = f(0) + \int_0^t f'(s)\mathrm{d}s$.

We apply this to show that the exponential of the double-hopping operator does not change the number of odd entries, except with low probability.

Lemma 13.7. *The following bound holds for all $\kappa \le 1$, $R \le \ell^{1/10}/2$, and $1 \le d \le R$.*

$$\left\| \sum_{o \in [R]} \mathsf{QE}_{\ge o+d} \cdot \exp\left(-\kappa \left(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R\right)\right) \cdot \mathsf{QE}_o \right\|_{\mathrm{op}} \le \left(\frac{2R^5}{\sqrt{\ell}}\right)^{d/4} . \tag{199}$$

Proof. We use Fact 13.6 applied to the following decomposition of $\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R$,

$$\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R = \underbrace{\sum_{o \in [R]} \mathsf{QEC}_{(R,=o)} \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{QEC}_{(R,=o)}}_{A} + \underbrace{\sum_{\substack{o,o' \in [R] \\ o \neq o'}} \mathsf{QEC}_{(R,=o)} \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{QEC}_{(R,=o')}}_{V} . \tag{200}$$

We note that V is equal to $\mathsf{Con}_R \sum_{o'} (\mathrm{id} - \mathsf{QE}_{=o'}) \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{QE}_{=o'}$, whose operator norm is bounded by Corollary 13.4. We further note that using the fact that $\widetilde{\mathsf{G}}_y^2$ only moves 2 bosons (and therefore affects at most 4 momentum modes), and applying the definition of A, we have that

$$V \cdot \mathsf{QE}_o = \mathsf{QE}_{o+4} \cdot V \cdot \mathsf{QE}_o \quad \text{and} \quad A \cdot \mathsf{QE}_o = \mathsf{QE}_o \cdot A \cdot \mathsf{QE}_o . \tag{201}$$

Repeatedly pushing the QE_o operator through, this implies that $\mathsf{QE}_{\ge o+d} \cdot \mathrm{e}^{-\kappa A} \cdot \mathsf{QE}_o = 0$ and for all $d' < d/4$ and $0 \le s_1 \le \ldots \le s_{d'} \le \kappa$,

$$\mathsf{QE}_{\ge o+d} \cdot \exp\left(-(\kappa - s_{d'}) \cdot A\right) \cdot \underbrace{V \ldots V}_{d' \text{ times}} \cdot \exp\left(-s_1 \cdot A\right) \cdot \mathsf{QE}_o = 0 . \tag{202}$$

Now applying Fact 13.6 with A and V as defined in the equation above, we have the following bound on the norm for a fixed o,

$$\left\| \mathsf{QE}_{\ge o+d} \cdot \exp\left(-\kappa \left(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R\right)\right) \cdot \mathsf{QE}_o \right\|_{\mathrm{op}} \tag{203a}$$

$$= \left\| \sum_{k \ge d/4} (-1)^k \int_{0 \le s_1 \le \ldots \le s_k \le \kappa} \mathrm{d}\mathbf{s}\, \mathsf{QE}_{\ge o+d} \cdot \exp\left(-(\kappa - s_k) \cdot A\right) \cdot V \ldots V \cdot \exp\left(-s_1 \cdot A\right) \cdot \mathsf{QE}_o \right\|_{\mathrm{op}} \tag{203b}$$

$$\le \sum_{k \ge d/4} \int_{0 \le s_1 \le \ldots \le s_k \le \kappa} \mathrm{d}\mathbf{s} \left\| \mathsf{QE}_{\ge o+d} \cdot \exp\left(-(\kappa - s_k) \cdot A\right) \cdot V \ldots V \cdot \exp\left(-s_1 \cdot A\right) \cdot \mathsf{QE}_o \right\|_{\mathrm{op}} \tag{203c}$$

$$\leq \sum_{k\geq d/4} \int_{0\leq s_1\leq \ldots \leq s_k\leq \kappa} \mathrm{d}\mathbf{s}\, \|V\|_{\mathrm{op}}^k \tag{203d}$$

$$\leq \sum_{k\geq d/4} \left(\frac{R^5}{\sqrt{\ell}}\right)^k = \frac{1}{1-R^5/\sqrt{\ell}} \left(\frac{R^5}{\sqrt{\ell}}\right)^{\lceil d/4\rceil} \tag{203e}$$

$$\leq 2\left(\frac{R^5}{\sqrt{\ell}}\right)^{d/4} \tag{203f}$$

$$\leq \left(\frac{2R^5}{\sqrt{\ell}}\right)^{d/4}. \tag{203g}$$

Here, we first apply Dyson's formula. Recall that $\widetilde{\mathbf{G}}_y^2$ can only increase the number of odd indices by 4. Moreover, by definition, A does not change the number of odd indices and hence neither does any exponential in A. Therefore, since the expression is sandwiched between $\mathsf{QE}_{\geq o+d}$ and QE_o, the only way to get a non-zero output is to be a product where the number of V terms is at least $d/4$. This means that (1) we can eliminate $\exp(-\kappa\cdot A)$ arising from Dyson's formula, as well as (2) start the sum at $k \geq d/4$. Then we apply the triangle inequality to move the norm into the sum and integral. Then we use the fact that because $\widetilde{\mathbf{G}}_y^2$ is a positive operator (and sandwiching a positive operator by projectors yields a positive operator), $\exp(-t\cdot A)$ has operator norm at most 1. We then use the fact that the operator norm is sub-multiplicative and Corollary 13.4 bounds the operator norm of V by $\frac{R^5}{\sqrt{\ell}}$. Finally, we use the fact that $\int_{0\leq s_1\leq\ldots\leq s_k\leq\kappa} \mathrm{d}s = \kappa^k/k! \leq 1$ whenever $\kappa \leq 1$ and bound the geometric series. □

13.3 Recursive bounds on the oddness of products of operators

The previous lemma bounds the probability that a state having o odd indices ends up in the subspace of momentum Fock states with $o+d$ many odd indices, for every d. We can apply a generalization of the stars-and-bars style counting argument to show that the probability of applying the operator many times does not increase the number of odd indices either.

Lemma 13.8. *Let $\Pi_0 \preceq \Pi_1 \preceq \ldots$ be a sequence of projectors on a Hilbert space $\mathcal{H}$. Let $A_1, \ldots, A_t$ be a family of operators on $\mathcal{H}$ such that $\|A_i\|_{\mathrm{op}} \leq 1$. Furthermore, suppose that for all integers $a, b \geq 0$, $\|(\mathrm{id}-\Pi_{a+b})A_i\Pi_a\|_{\mathrm{op}} \leq \varepsilon^{b+1}$. Then for all integers $\lambda \geq 0$ and states $|\psi\rangle$ such that $\Pi_0|\psi\rangle = |\psi\rangle$,*

$$\|(\mathrm{id}-\Pi_\lambda)A_t\ldots A_1|\psi\rangle\| \leq \binom{t+\lambda}{t-1}\varepsilon^{\lambda+1}. \tag{204}$$

Proof. The statement for $t=1$ is trivially true. Next define $\Pi_{-1}=0$ and for $j=0,\ldots,\lambda$, $\Delta_j = \Pi_j - \Pi_{j-1}$. Observe that $\Delta_j = \Pi_j(\mathrm{id}-\Pi_{j-1})$ by containment of the projectors. Then, write

$$\begin{aligned}\|(\mathrm{id}-\Pi_\lambda)A_{t+1}\ldots A_1|\psi\rangle\| \leq{}& \underbrace{\sum_{j=0}^{\lambda} \left\|(\mathrm{id}-\Pi_\lambda)A_{t+1}\Delta_j A_t\ldots A_1\Pi_0|\psi\rangle\right\|}_{(A)} \\ &+ \underbrace{\|(\mathrm{id}-\Pi_\lambda)A_{t+1}(\mathrm{id}-\Pi_\lambda)A_t\ldots A_1\Pi_0|\psi\rangle\|}_{(B)}\end{aligned} \tag{205}$$

We handle each of the terms separately. To bound (B), we can use induction to bound

$$\|(\mathrm{id}-\Pi_\lambda)A_{t+1}(\mathrm{id}-\Pi_\lambda)A_t \dots A_1\Pi_0\,|\psi\rangle\| \le \|(\mathrm{id}-\Pi_\lambda)A_{t+1}(\mathrm{id}-\Pi_\lambda)\|_{\mathrm{op}} \cdot \|(\mathrm{id}-\Pi_\lambda)A_t \dots A_1\Pi_0\|_{\mathrm{op}} \tag{206a}$$

$$\le \underbrace{\|A_{t+1}\|}_{\le 1} \cdot \binom{t+\lambda}{t-1}\varepsilon^{\lambda+1}. \tag{206b}$$

For the terms in (A), by induction,

$$\sum_{j=0}^{\lambda} \|(\mathrm{id}-\Pi_\lambda)A_{t+1}\Delta_j A_t \dots A_1\Pi_0\|_{\mathrm{op}} = \sum_{j=0}^{\lambda} \|(\mathrm{id}-\Pi_\lambda)A_{t+1}\Pi_j(\mathrm{id}-\Pi_{j-1})A_t \dots A_1\Pi_0\|_{\mathrm{op}} \tag{207a}$$

$$\le \sum_{j=0}^{\lambda} \|(\mathrm{id}-\Pi_\lambda)A_{t+1}\Pi_j\|_{\mathrm{op}} \cdot \|(\mathrm{id}-\Pi_{j-1})A_t \dots A_1\Pi_0\|_{\mathrm{op}} \tag{207b}$$

$$\le \sum_{j=0}^{\lambda} \varepsilon^{\lambda-j+1} \cdot \binom{t+j-1}{t-1} \cdot \varepsilon^j \tag{207c}$$

$$= \binom{t+\lambda}{t}\varepsilon^{\lambda+1}. \tag{207d}$$

Adding up all the terms

$$(A)+(B) \le \left(\binom{t+\lambda}{t} + \binom{t+\lambda}{t-1}\right)\varepsilon^{\lambda+1} = \binom{t+\lambda+1}{t}\varepsilon^{\lambda+1}. \tag{208}$$

completing the inductive proof. □

13.4 Wrapping up the proof

In this section, we apply the previous two lemmas to show that algorithms that make polynomial in n queries to the U oracle are exponentially close to the $\mathsf{QE}_{v/4}$ subspace.

Lemma 13.9. *The following bound holds for all $y \in \{0,1\}^n$, with $y \neq 0^n$, $\kappa \le 1$, $R \le \ell^{1/10}/2$ and $1 \le d \le R$,*

$$\left\| \sum_o \mathsf{QE}_{\ge o+d} \cdot \sqrt{1 - \frac{1}{2}\exp\left(-\kappa\left(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R\right)\right)} \cdot \mathsf{QE}_o \right\|_{\mathrm{op}} \le \left(\frac{64R^5}{\sqrt{\ell}}\right)^{d/4}. \tag{209}$$

Proof. We perform the proof by applying the previous two lemmas to the polynomial expansion of the square root. The Taylor/binomial expansion of the square root is given by

$$\sqrt{1+x} = \sum_{k=0}^{\infty} \binom{1/2}{k} x^k. \tag{210}$$

Applying it to the expression in eq. (209), we have the following

$$\sqrt{1 - \frac{1}{2}\exp\left(-\kappa\left(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R\right)\right)} = \sum_{k=0}^{\infty} \binom{1/2}{k}\left(\frac{-1}{2}\right)^k \exp\left(-\kappa\left(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R\right)\right)^k. \tag{211}$$

We can bound the norm in eq. (209) using Lemma 13.8, with $\lambda = d-1$, projectors $\Pi_0 = \mathsf{QE}_o \preceq \Pi_1 = \mathsf{QE}_{o+1} \preceq \ldots \preceq \mathsf{QE}_{o+d-1}$, and $A_1 = \ldots = A_t = \exp\big(-\kappa\big(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R\big)\big)$.

$$\left\| \mathsf{QE}_{\geq o+d} \cdot \sqrt{1 - \frac{1}{2}\exp\left(-\kappa\left(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R\right)\right)} \cdot \mathsf{QE}_o \right\|_{\text{op}} \tag{212a}$$

$$= \left\| \sum_{k=0}^{\infty} \binom{1/2}{k} \left(\frac{-1}{2}\right)^k \cdot \mathsf{QE}_{\geq o+d} \cdot \exp\left(-\kappa\left(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R\right)\right)^k \cdot \mathsf{QE}_o \right\|_{\text{op}} \tag{212b}$$

$$\leq \sum_{k=1}^{\infty} \left|\binom{1/2}{k}\right| \left(\frac{1}{2}\right)^k \cdot \left\| \mathsf{QE}_{\geq o+d} \cdot \exp\left(-\kappa\left(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R\right)\right)^k \cdot \mathsf{QE}_o \right\|_{\text{op}} \tag{212c}$$

$$\leq \left(\frac{2R^5}{\sqrt{\ell}}\right)^{d/4} \sum_{k=1}^{\infty} \left|\binom{1/2}{k}\right| \left(\frac{1}{2}\right)^k \binom{k+d}{k-1} \tag{212d}$$

$$\leq \left(\frac{2R^5}{\sqrt{\ell}}\right)^{d/4} \cdot 2^{d+1} \tag{212e}$$

$$\leq \left(\frac{64R^5}{\sqrt{\ell}}\right)^{d/4} . \tag{212f}$$

In the first line, we apply the Taylor expansion of the square root from before, noting that the norm in the $k=0$ term in eq. (212c) is $\left\|\mathsf{QE}_{\geq o+d} \cdot \mathsf{QE}_o\right\|_{\text{op}} = 0$. Then, we apply the triangle inequality. In the third line, we apply Lemma 13.8. together with the bound of $(2R^5/\sqrt{\ell})^{d/4}$ from Lemma 13.7. For the final line, we bound the infinite sum by

$$\sum_{k=1}^{\infty} \left|\binom{1/2}{k}\right| \left(\frac{1}{2}\right)^k \binom{k+d}{k-1} \leq \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k \binom{k+d}{k-1} \tag{213a}$$

$$= \frac{1}{2} \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k \binom{(d+2)+k-1}{k} \tag{213b}$$

$$= \frac{1}{2} \sum_{k=0}^{\infty} \left(-\frac{1}{2}\right)^k \binom{-(d+2)}{k} \tag{213c}$$

$$= \frac{1}{2} \left(1 - \frac{1}{2}\right)^{-d-2} \tag{213d}$$

$$= 2^{d+1} . \tag{213e}$$

Here, in the first line, we use the fact that $\left|\binom{1/2}{k}\right| \leq 1$ for all $k \geq 1$. Then we re-index the sum to start from $k=0$, and write $k+d+1 = (d+2)+k-1$. Then we use the equality $\binom{-a}{b} = (-1)^b \binom{a+b-1}{b}$, with $a = d+2$ and $b = k$. Finally, we use the binomial expansion, $(1-x)^n = \sum_{k=0}^{\infty} \binom{n}{k} x^k$, with $n = -(d+2)$ and $x = -1/2$. □

Recall, $e_0(z)$ and $e_1(z)$ were defined as

$$e_0(z) \stackrel{\text{def}}{=} 1 - \mathrm{e}^{-\kappa z}, \quad \text{and} \tag{214a}$$

$$e_1(z) \stackrel{\text{def}}{=} \sqrt{\mathrm{e}^{-\kappa z}(2-\mathrm{e}^{-\kappa z})} = \sqrt{2}\mathrm{e}^{-\kappa z/2}\sqrt{1-\mathrm{e}^{-\kappa z}/2}\,. \tag{214b}$$

Since the Kraus operators E_0 and E_1 are defined in terms of the expressions given in Lemma 13.7 and Lemma 13.9, we can combine these two lemmas to show that the two Kraus operators E_0 and E_1 also do not increase the number of odd entries by too much, and we then have the following lemma.

Lemma 13.10 (Single query preserves quasi-evenness)**.** *For all o, $R \le \ell^{1/10}/2$, $1 \le d \le R$, and operators A with $\|A\|_{\mathrm{op}} \le 1$ acting on A, the following inequality holds.*

$$\left\| \mathsf{QE}_{\ge o+d} \left(\sum_y A \cdot |y\rangle\langle y| \otimes \begin{pmatrix} \widetilde{X}_{\mathsf{U}_y} \otimes e_1(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R) \\ +\widetilde{Z}_{\mathsf{U}_y} \otimes e_0(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R) \end{pmatrix} \right) \mathsf{QE}_o \right\|_{\mathrm{op}} \le \left(\frac{2^{14}R^5 d}{\sqrt{\ell}} \right)^{d/4}. \tag{215}$$

Here, we have omitted id_{A} on the projectors QE_o and $\mathsf{QE}_{\ge o+d}$.

Proof. We expand out the operator norm as follows.

$$\left\| \mathsf{QE}_{\ge o+d} \left(\sum_y A \cdot |y\rangle\langle y| \otimes \begin{pmatrix} \widetilde{X}_{\mathsf{U}_y} \otimes e_1(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R) \\ +\widetilde{Z}_{\mathsf{U}_y} \otimes e_0(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R) \end{pmatrix} \right) \mathsf{QE}_o \right\|_{\mathrm{op}} \tag{216a}$$

$$= \left\| \sum_y A \cdot |y\rangle\langle y| \otimes \mathsf{QE}_{\ge o+d} \begin{pmatrix} \widetilde{X}_{\mathsf{U}_y} \otimes e_1(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R) \\ +\widetilde{Z}_{\mathsf{U}_y} \otimes e_0(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R) \end{pmatrix} \mathsf{QE}_o \right\|_{\mathrm{op}} \tag{216b}$$

$$\le \left\| \sum_y |y\rangle\langle y| \otimes \mathsf{QE}_{\ge o+d} \begin{pmatrix} \widetilde{X}_{\mathsf{U}_y} \otimes e_1(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R) \\ +\widetilde{Z}_{\mathsf{U}_y} \otimes e_0(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R) \end{pmatrix} \mathsf{QE}_o \right\|_{\mathrm{op}} \tag{216c}$$

$$\le \max_y \left\| \mathsf{QE}_{\ge o+d} \begin{pmatrix} \widetilde{X}_{\mathsf{U}_y} \otimes e_1(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R) \\ +\widetilde{Z}_{\mathsf{U}_y} \otimes e_0(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R) \end{pmatrix} \mathsf{QE}_o \right\|_{\mathrm{op}} \tag{216d}$$

$$\begin{aligned} &\le \max_y \left\| \mathsf{QE}_{\ge o+d} \left(\widetilde{X}_{\mathsf{U}_y} \otimes e_1(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R) \right) \mathsf{QE}_o \right\|_{\mathrm{op}} \\ &\quad + \max_y \left\| \mathsf{QE}_{\ge o+d} \left(\widetilde{Z}_{\mathsf{U}_y} \otimes e_0(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R) \right) \mathsf{QE}_o \right\|_{\mathrm{op}} \end{aligned} \tag{216e}$$

$$\begin{aligned} &\le \underbrace{\max_y \left\| \mathsf{QE}_{\ge o+d} \cdot \left(e_1(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R) \right) \cdot \mathsf{QE}_o \right\|_{\mathrm{op}}}_{(A)} \\ &\quad + \underbrace{\max_y \left\| \mathsf{QE}_{\ge o+d} \cdot \left(e_0(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R) \right) \cdot \mathsf{QE}_o \right\|_{\mathrm{op}}}_{(B)}. \end{aligned} \tag{216f}$$

Here, we first use the fact that $\mathsf{QE}_{\ge o+d}$ acts only on the purifying register, then use the fact that $\|A\|_{\mathrm{op}} \le 1$. Then we use the fact that for any operator that is block-diagonal, the operator norm is the max of the operator norms restricted to each block. Finally, we apply the triangle inequality. Now we bound parts (A) and (B) separately using Lemma 13.7 and Lemma 13.9. We have the following

$$(A) = \max_y \sqrt{2} \left\| \mathsf{QE}_{\ge o+d} \cdot \left(\mathrm{e}^{-(\kappa/2)\cdot\left(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R\right)} \sqrt{1 - \frac{1}{2}\mathrm{e}^{-\kappa\left(\mathsf{Con}_R \cdot \widetilde{\mathsf{G}}_y^2 \cdot \mathsf{Con}_R\right)}} \right) \cdot \mathsf{QE}_o \right\|_{\mathrm{op}} \tag{217a}$$

$$\leq \sqrt{2}(d+1)\left(\frac{64R^5}{\sqrt{\ell}}\right)^{d/4}. \tag{217b}$$

Here, for every y, we use Lemma 13.8 with $\lambda = d-1$, the same projectors $\Pi_1 = \mathsf{QE}_o \preceq \ldots \Pi_{d-1} = \mathsf{QE}_{o+d}$, and the 2 operators $A_2 = \mathrm{e}^{-(\kappa/2)\left(\mathsf{Con}_R\cdot\widetilde{\mathsf{G}}_y^2\cdot\mathsf{Con}_R\right)}$ and $A_1 = \sqrt{1-\frac{1}{2}\mathrm{e}^{-\kappa\left(\mathsf{Con}_R\cdot\widetilde{\mathsf{G}}_y^2\cdot\mathsf{Con}_R\right)}}$. Lemma 13.7 and Lemma 13.9 bound each $\|(\mathrm{id}-\Pi_{d-1})A_i\Pi_o\|_{\mathrm{op}} \leq (64R^5/\sqrt{\ell})^{d/4}$, completing the bound.

Similarly, we bound (B) as follows.

$$(B) \leq \max_y \left\| \mathsf{QE}_{\geq o+d}\cdot\left(\mathrm{id}-\mathrm{e}^{-\kappa\left(\mathsf{Con}_R\cdot\widetilde{\mathsf{G}}_y^2\cdot\mathsf{Con}_R\right)}\right)\cdot\mathsf{QE}_o\right\| \tag{218a}$$

$$\leq \left(\frac{2R^5}{\sqrt{\ell}}\right)^{d/4}. \tag{218b}$$

Here, we directly apply Lemma 13.7, together with the fact that $\mathsf{QE}_{\geq o+d}\cdot\mathsf{QE}_o = 0$ so we can remove the identity term. Putting things together, we have an upper bound on the operator norm of

$$(A)+(B) \leq 2d\left(\frac{64R^5}{\sqrt{\ell}}\right)^{d/4} \leq \left(\frac{2^{14}R^5}{\sqrt{\ell}}\right)^{d/4}. \tag{219}$$

Here, we take the upper bound $2d \leq 2^{d+1}$ in the final line. □

The previous lemma applies to a standard query y. However, if the verification algorithm applies a conditional query and in the case that the conditional query is not applied, it is clear to see that the algorithm cannot change the quasi-evenness of the S register. Formally, we have the following.

Corollary 13.11 (Single controlled query preserves quasi-evenness)**.** *For all o, $R \leq \ell^{1/10}/2$, $1 \leq d \leq R$, and unitaries A acting on A, the following inequality holds.*

$$\left\|\mathsf{QE}_{\geq o+d}\left(\sum_{y,b} A\cdot|b,y\rangle\langle b,y|\otimes\begin{pmatrix}\widetilde{X}_{\mathsf{U}_y}\otimes e_1(\mathsf{Con}_R\cdot\widetilde{\mathsf{G}}_y^2\cdot\mathsf{Con}_R)\\ +\widetilde{Z}_{\mathsf{U}_y}\otimes e_0(\mathsf{Con}_R\cdot\widetilde{\mathsf{G}}_y^2\cdot\mathsf{Con}_R)\end{pmatrix}^b\right)\mathsf{QE}_o\right\|_{\mathrm{op}} \leq \left(\frac{2^{14}R^5 d}{\sqrt{\ell}}\right)^{d/4}. \tag{220}$$

Here, we have omitted id_{A} *on the projectors* QE_o *and* $\mathsf{QE}_{\geq o+d}$*.*

Proof. Applying the triangle inequality, we have that

$$\left\|\mathsf{QE}_{\geq o+d}\left(\sum_{y,b} A\cdot|b,y\rangle\langle b,y|\otimes\begin{pmatrix}\widetilde{X}_{\mathsf{U}_y}\otimes e_1(\mathsf{Con}_R\cdot\widetilde{\mathsf{G}}_y^2\cdot\mathsf{Con}_R)\\ +\widetilde{Z}_{\mathsf{U}_y}\otimes e_0(\mathsf{Con}_R\cdot\widetilde{\mathsf{G}}_y^2\cdot\mathsf{Con}_R)\end{pmatrix}^b\right)\mathsf{QE}_o\right\|_{\mathrm{op}} \tag{221a}$$

$$\leq \left\|\mathsf{QE}_{\geq o+d}\left(\sum_y A\cdot|1,y\rangle\langle 1,y|\otimes\begin{pmatrix}\widetilde{X}_{\mathsf{U}_y}\otimes e_1(\mathsf{Con}_R\cdot\widetilde{\mathsf{G}}_y^2\cdot\mathsf{Con}_R)\\ +\widetilde{Z}_{\mathsf{U}_y}\otimes e_0(\mathsf{Con}_R\cdot\widetilde{\mathsf{G}}_y^2\cdot\mathsf{Con}_R)\end{pmatrix}\right)\mathsf{QE}_o\right\|_{\mathrm{op}} + \left\|\mathsf{QE}_{\geq o+d}\left(\sum_y A\cdot|0,y\rangle\langle 0,y|\otimes\mathrm{id}\right)\mathsf{QE}_o\right\|_{\mathrm{op}} \tag{221b}$$

$$\leq \left(\frac{2^{14}R^5 d}{\sqrt{\ell}}\right)^{d/4} + 0\,. \tag{221c}$$

Here, we use the fact that when $b=0$, the inner unitary is the identity, and then we use the fact that $\mathsf{QE}_{\geq o+d}$ and QE_o are projectors onto orthogonal subspaces since $d \geq 1$ and they commute past A and $|0,y\rangle\langle 0,y|$. When $b=1$, we use the fact that $\|A|1\rangle\langle 1|\|_{\mathrm{op}} \leq 1$ and apply Lemma 13.10. □

Theorem 13.12. *Let $\mathcal{A}$ be any T-query algorithm making queries to U and $v \in \mathbb{Z}_{\geq 0}$ be a multiple of 4. Then there exists a constant c such that for suitably large n, the purified state of the algorithm has squared overlap with the complement of the $(c \cdot T^{10}, v/4)$-low collision subspace that is lower bounded by*

$$\left\| \left(\mathrm{id} - \mathsf{QEC}_{(c\cdot T^{10}, v/4)}\right) |\psi_{\mathrm{PQ}}\rangle \right\|^2 \leq \left(\left(\frac{T^4}{\ell^{1/32}} \right)^v + \mathrm{e}^{-5T} \right)^2 . \tag{222}$$

Proof. Assume without loss of generality that $T \geq n$. Then we apply Theorem 12.2 with $\iota = \mathrm{e}^{-5T}$ to show that there exist integers $r = O(T^{10})$ and $R = O(T^{13})$ to get that

$$\| |\psi_{\mathrm{PQ}}\rangle - |\psi_{R,r}\rangle \| \leq \mathrm{e}^{-5T} . \tag{223}$$

Then, we apply Lemma 13.8 with $\lambda = v/4$, $\Pi_0 = \mathsf{QEC}_{r,0} \preceq \ldots \preceq \Pi_{v/4} = \mathsf{QEC}_{r,v/4}$ and operators A_i being

$$A_i = \left(\sum_{y,b} A \cdot |b, y\rangle\langle b, y| \otimes \begin{pmatrix} \widetilde{X}_{\mathsf{U}_y} \otimes e_1(\mathsf{Con}_R \cdot \widetilde{\mathbf{G}}_y^2 \cdot \mathsf{Con}_R) \\ + \widetilde{Z}_{\mathsf{U}_y} \otimes e_0(\mathsf{Con}_R \cdot \widetilde{\mathbf{G}}_y^2 \cdot \mathsf{Con}_R) \end{pmatrix}^b \right), \tag{224}$$

where A here is the unitary that the verification algorithm applies, and we note that there are $T \leq \ell^{1/10}/2$ of them. We also use the fact that the initial state is contained in QE_0. Then Corollary 13.11 gives us the bound on the individual norms of $\left(\frac{64R^5(v/4)}{\sqrt{\ell}}\right)^{(v/4+1)/4}$. Resolving Lemma 13.8, we have

$$\left\| \left(\mathrm{id} - \mathsf{QEC}_{(r,v/4)}\right) \cdot |\psi_{R,r}\rangle \right\| \leq \binom{T + v/4}{T-1} \left(\frac{2^{14}R^5}{\sqrt{\ell}} \right)^{(v/4+1)/4} \tag{225a}$$

$$\leq (2T)^{v/4} \left(\frac{2^{14}R^5}{\sqrt{\ell}} \right)^{v/16} \tag{225b}$$

$$\leq \left(\frac{2^{18}R^5T^4}{\sqrt{\ell}} \right)^{v/16} . \tag{225c}$$

Combining these two equations using the triangle inequality, we have that

$$\left\| \left(\mathrm{id} - \mathsf{QEC}_{(r,v/4)}\right) \cdot |\psi_{\mathrm{PQ}}\rangle \right\| \leq \left(\frac{2^{18}R^5T^4}{\sqrt{\ell}} \right)^{v/16} + \mathrm{e}^{-5T} \tag{226a}$$

$$\leq \left(\frac{4R^{5/16}T^{1/4}}{\ell^{1/32}} \right)^v + e^{-5T} . \tag{226b}$$

Substituting $R = c \cdot T^{13}$ and $r = c \cdot T^{10}$ for some constant c (for suitably large n), for sufficiently large T, we have that $4R^{5/16}T^{1/4} \leq T^4$ which completes the theorem. □

Proving the main sampling upper bound Combining the gentle measurement lemma, Theorem 13.12, and Theorem 10.2, which bounds the probability of sampling v many points given that the state is in $\mathsf{QEC}_{(c\cdot T^{10}, v/4)}$ (with $T = vt$ being the number of queries), we achieve the main result of Parts III and IV: Theorem 9.1.

Theorem (Sampling probability upper bound,Theorem 9.1, restated)**.** *For all v, for all quantum algorithms $\mathcal{A}^U$ accessing an oracle U and outputting v distinct outputs, while making t queries per output, if a pair of oracles (S, U) are sampled according to distribution* Strong *(defined in Definition 7.2), then the probability that all v of the outputs of $\mathcal{A}^U$ are elements of S is at most*

$$\leq 2\left(\frac{4v((vt)^{30}+v(vt)^{20})\sqrt{\ell}}{2^{n/4}}\right)^{v}+\left(\left(\frac{(vt)^4}{\ell^{1/32}}\right)^{v}+e^{-5vt}\right)^{2}. \tag{227}$$

Proof of Theorem 9.1. Let $|\psi_{\mathrm{PQ}}\rangle$ be the state of the algorithm after making vt queries to U, then we have, for Π_{succ} being the success operator from eq. (56).

$$\left\|\Pi_{\mathrm{succ}}\left|\psi_{\mathrm{PQ}}\right\rangle\right\|^2 \leq \left\|\Pi_{\mathrm{succ}}\cdot \mathsf{QEC}_{(c\cdot v^{10}t^{10},v/4)}\cdot\left|\psi_{\mathrm{PQ}}\right\rangle\right\|^2+\left\|\Pi_{\mathrm{succ}}\cdot\left(\mathrm{id}-\mathsf{QEC}_{(c\cdot v^{10}t^{10},v/4)}\right)\left|\psi_{\mathrm{PQ}}\right\rangle\right\|^2 \tag{228a}$$

$$\leq 2\left(\frac{4v((vt)^{30}+v(vt)^{20})\sqrt{\ell}}{2^{n/4}}\right)^{v}+\left(\left(\frac{(vt)^4}{\ell^{1/32}}\right)^{v}+e^{-5vt}\right)^{2}. \tag{228b}$$

Here, we apply the triangle inequality and bound the second term using Theorem 13.12, and the first term using Theorem 10.2 with $T = vt$. □

Part V
Theorem statements and concluding remarks

14 Property-testing and oracle separations

We can now combine both our sampler success probability upper bound with the lower bound on the success probability implied by a QCMA algorithm to get a lower bound on the witness length of a successful QCMA algorithm.

Theorem 14.1. *Consider any choice of constants $a > 0$ and functions $t(n), q(n)$ that satisfy $t(n) \leq an^a$, $q(n) \leq an^a$ for all $n \geq n_0$. Then, let n_0 be the smallest integer such that for all $n \geq n_0$, both Theorem 6.2 and Theorem 9.1 apply using for $t = t(n), q = q(n)$, and $v = 1000q$. Then for any $n \geq n_0$, and $\mathcal{A}$ a binary-output quantum query algorithm with classical witness of length $q(n)$ and making $t(n)$ queries to the oracles (S, U) of size n, there exists a pair of oracle (S^*, U^*) of size n such that*

1. *either (S^*, U^*) are at least $\frac{59}{100}$-spectrally Forrelated, but for all witnesses w of length $q(n)$,*

$$\mathbb{P}[\mathcal{A}^{(S^*,U^*)}(w) = 1] < \frac{2}{3}. \tag{229}$$

2. *or (S^*, U^*) are at most $\frac{57}{100}$-spectrally Forrelated, but there exists a witness $\widetilde{w}$ of length $q(n)$ such that*

$$\mathbb{P}[\mathcal{A}^{(S^*,U^*)}(\widetilde{w}) = 1] > \frac{1}{3}. \tag{230}$$

Proof. If the stated consequence was false, then the algorithm $\mathcal{A}$ properly classifies all spectral Forrelation problems promised that either the instance is at least 59/100- or at most 57/100-spectrally Forrelated of size n. However, setting $\ell = 2^{n/10}$ and $v = 1000q$, then we can apply Theorem 6.2 and Lemma 7.3 with $\kappa = 1/10$ and $\rho = \frac{2\ell^2}{2^n} \ln\left(\frac{2^n}{2\ell^4}\right)$, yielding a sampler that outputs v points with probability $O(t^{-1000q})$ when (S, U) is sampled from the distribution Strong. Applying Theorem 9.1 yields a sampling probability upper bound of $O\left((\text{poly}(n)2^{-n/160})^{1000q}\right)$, whenever t and q are $\leq an^a$, for the distribution Strong. For suitably large n, $(\text{poly}(n)2^{-n/160})^{1000q} \ll t^{-1000q}$, concluding the proof by contradiction. □

We have shown in Theorem 14.1 that for every sufficiently large instance size n, there exists a property testing problem about size n oracles such that there exists an n-qubit quantum witness for the problem verifiable by an efficient quantum algorithm. Still, there does not exist any $t = t(n)$ time quantum algorithm accepting $q = q(n)$-length classical witnesses for the problem, where both t and q are polynomial in n. This is technically not yet an oracle separation between QCMA and QMA.

The oracle separation language Let $S, U : \{0,1\}^* \to \{0,1\}$ be oracles and let S_n, U_n be the restriction to n-bit inputs. We define the *unary* language $\mathcal{L}^{S,U}$ as follows:

$$1^n \in \mathcal{L}^{S,U} \iff S_n, U_n : \{0,1\}^n \to \{0,1\} \text{ are } \geq \frac{59}{100}\text{-spectrally Forrelated,} \tag{231a}$$

$$1^n \notin \mathcal{L}^{S,U} \iff S_n, U_n : \{0,1\}^n \to \{0,1\} \text{ are } \leq \frac{57}{100}\text{-spectrally Forrelated,} \tag{231b}$$

$$\text{and, all other strings } \in \{0,1\}^* \text{ are } \notin \mathcal{L}^{S,U}. \tag{231c}$$

This language will be what we use to prove Theorem 1.1. We note that 1^n being in the language is determined by the oracle of size $n+1$ bits, as the pair of oracles (S_n, U_n) is an $n+1$ bit oracle. Roughly speaking, we will choose the oracles S, U such that the k-th polynomial-time uniform quantum query algorithm in an enumeration of quantum query algorithms will incorrectly identify membership of 1^{n_k} in $\mathcal{L}$ of an appropriately chosen integer n_k.

As one might suspect, we will use Theorem 14.1 to diagonalize against all polynomial-sized classical witness oracle algorithms. The challenge is that the algorithms can query the oracle at any length—intuitively, on input 1^n, querying at lengths $\neq n$ will not help the algorithm. However, formalizing this intuition is tedious. We do this in the following proof by identifying a family of appropriately chosen integers $n_1 < n_2 < \ldots$ such that the algorithm on input of length n_k will not query at lengths $\geq n_{k+1}$. This will be enough to inductively select the definition of the oracle at size n_k (based on the choices of the oracle at all sizes $< n_k$) to complete the diagonalization.

Proof of Theorem 1.1. For any choice of oracles S, U such that each restriction to size n inputs encodes either a $\geq 59/100$ or $\leq 57/100$ instance of spectral Forrelation, the containment of $\mathcal{L}^{S,U}$ in $\mathsf{QMA}^{S,U}$ follows from Theorem 5.10. Now we prove the lower bound for QCMA algorithms.

Let $M_1, M_2, \ldots$ be an enumeration of all possible Turing machines. Second, identify any surjective function $\iota : \mathbb{N} \twoheadrightarrow \mathbb{N}^2$ and define functions $j, a : \mathbb{N} \to \mathbb{N}$ by $(j(k), a(k)) = \iota(k)$. We will use this surjection to diagonalize against all possible polynomial-time quantum algorithms to prove a QCMA-lower bound.

Third, define a function $F : \mathbb{N} \to \mathbb{N}$ by $F(a)$ is the minimum value such that for all $n \geq F(a)$, any $t(n) = an^a$ query algorithm with $q(n) = an^a$ length classical witness must misclassify some pair (S, U) of size n. By Theorem 14.1, for every integer a, $F(a)$ is well-defined. In other words, $F(a)$ is the first integer n for which we can guarantee that an an^a query and an^a witness restricted algorithm must misclassify some pair (S, U) of size n.

Fourth, we identify integers $n_1, n_2, \ldots$ where the oracles will be defined to be non-zero. Define

$$n_1 \stackrel{\text{def}}{=} 1 + F(a(1)), \tag{232a}$$

$$\forall\ k > 1,\ n_k \stackrel{\text{def}}{=} 1 + \max\left\{F(a(k)),\ a(k-1)(n_{k-1})^{a(k-1)}\right\}. \tag{232b}$$

We now define the oracles S, U by defining the oracles at each length. For any $n \in \mathbb{N} \setminus \{n_1, n_2, \ldots\}$, let both oracles S_n and U_n equal 0 everywhere. Then the spectral-Forrelation problem defined by the n-th pair of oracles is trivially a no instance for $n \in \mathbb{N} \setminus \{n_1, n_2, \ldots\}$. We now go through and define the oracles at sizes remaining sizes: $n_1, n_2, \ldots,$.

For each $k = 1, 2, 3, \ldots$, run Turing machine $M_{j(k)}$ on input 1^{n_k} for $a(k)n_k^{a(k)}$ steps and interpret its output as a quantum query circuit $\mathcal{A}_{n_k}$ which takes as input a classical witness. By adding the halting conditions to the Turing machine, we have implicitly enforced that the witness length and total number of gates (elementary or oracle) are at most $a(k)n_k^{a(k)}$. Furthermore, by construction, the largest input queryable by this algorithm is size $a(k)n_k^{a(k)}$. This is a standard diagonalization trick to ensure that every polynomial-time and polynomial-query algorithm is considered and that no super-polynomial parameterized algorithms are accidentally considered.

Next, we take this algorithm $\mathcal{A}_{n_k}$ and we build from it a query algorithm $\mathcal{B}_{n_k}$ that only queries oracles of input length n_k. To construct $\mathcal{B}_{n_k}$, take the algorithm $\mathcal{A}_{n_k}$ which can make oracle queries of varying sizes and for every query it makes of length $< n_k$, use the previously generated definitions of the oracles S, U and hardcode these answers. This is well defined as we are defining the oracles S, U for progressively larger input sizes. For queries it makes of length $> n_k$, replace the oracle gates with identity circuits. The resulting circuit will be $\mathcal{B}_{n_k}$, which only makes queries of length $= n_k$. This new algorithm $\mathcal{B}_{n_k}$ can be used to derive a pair (S_{n_k}, U_{n_k}) by applying Theorem 14.1 on $\mathcal{B}_{n_k}$ to generate the pair (S_{n_k}, U_{n_k}).

This completes the construction of the oracles S, U everywhere. It remains to prove that no $\mathsf{QCMA}^{S,U}$ algorithm exists. Assume, for contradiction, there exists a P-uniform family of quantum oracle algorithms $\{\mathcal{A}_n\}$ that solve spectral Forrelation for a witness of length $q(n) = \mathrm{poly}(n)$ with $t(n) = \mathrm{poly}(n)$ queries. Then, the family appears in the Turing machine enumeration as some $M_{j^\star}$ and there exists some $a^\star$ such that $t(n), q(n) \leq a^\star n^{a^\star}$. As ι is a surjection, there exists a $k^\star$ such that $\iota(k^\star) = (j^\star, a^\star)$. We now prove that this algorithm will misclassify the string $1^{n_{k^\star}}$, thereby proving that it does not solve spectral Forrelation.

To prove this, let $\mathcal{A}_{n_{k^\star}}$ be the quantum circuit for inputs of length $n_{k^\star}$. Observe that since the oracles are defined as $= 0$ for inputs $\notin \{n_1, n_2, \ldots\}$, and the fact that $n_{k^\star+1} > a^\star n_{k^\star}^{a^\star}$ (by construction), each query gate for inputs of length $> n_{k^\star}$ is effectively an identity gate as it only makes queries on inputs of length at most $a^\star n_{k^\star}^{a^\star}$. Therefore, by hardcoding the behavior on input sizes $< n_{k^\star}$, the query algorithm $\mathcal{B}_{n_{k^\star}}$ (previously defined) has the exact same output as $\mathcal{A}_{n_{k^\star}}$ on inputs of size $n_{k^\star}$. However, using Theorem 14.1, we specifically constructed a pair (S_{n_k}, U_{n_k}) that $\mathcal{B}_{n_{k^\star}}$ will misclassify. Therefore, the family $\{\mathcal{A}_n\}$ will answer incorrectly on input $1^{n_{k^\star}}$, completing the proof.

□

15 Concluding remarks

Having concluded the proof, we take a moment to reflect on the path that led us here. In particular, we discuss the role of the bosonic purification of the oracle, its conceptual advantages, and the technical challenges that arose in this formulation.

Zhandry's observation. Zhandry's observation [Zha25] concerned the *use-once* nature of quantum witnesses versus the *reusability* of classical witnesses. This distinction allowed him to construct a sampler

that, given oracle access to S, could produce multiple distinct samples from S, contingent on a technical conjecture.

The sampler perspective was appealing, as it delineated the boundary between what a classical witness can explicitly list about S and what additional structure must be inferred by the verifier through oracle queries. However, the argument hinged on a conjecture asserting that queries to the unitary oracle U were *computationally* indistinguishable from random—a conjecture which ultimately fails. While Zhandry's setting involved the quantum Fourier transform and ours employs a related but distinct Fourier-analytic framework, the key issue was the same: queries to U are *not* random, and an algorithm could detect that they are far from uniformly random. The saving observation is that their behavior is amenable to precise analysis through Fourier and, ultimately, bosonic tools. This is made precise by our analysis in the bosonic framework, which exactly characterizes the action of queries to U. Our proof can circumvent Zhandry's issue as it proves a sampling probability upper bound based on the bosonic characterization. In some sense, our proof suggests that the QMA vs. QCMA separation problem is not as related to pseudorandomness as prior results [LMY25, Zha25] have suggested.

Recognizing this failure was serendipitous. It led us to the insight that one could instead adapt the sampler proof to operate solely through access to U, while still producing valid samples from S. This realization shifted our focus entirely to understanding the detailed structure of oracle queries in the U-picture.

Strong sampling upper bounds Recall that the sampling probability lower bound we derive is incredibly small. Its scaling is roughly $2^{-q} \cdot \Omega(t)^{-2v}$ for q, the length of the witness, t the number of queries, and v the number of samples produced. However, as long as t is polynomial, this bound is only quasi-exponentially small. Therefore, a contradicting upper bound should be even smaller in order to contradict the existence of a QCMA algorithm in all polynomial parameter regimes. This stringent requirement means that it does not suffice to approximate the behavior of the state after each query or guess. In particular, what's important to study is the post-selected state of the first guesses $z_1, \ldots, z_k$ being correct. Each post-selection is conditioned on an event of exponentially small probability, and therefore, the post-selection state will be very far from the original state.

In particular, we spent considerable time chasing the idea that the post-measurement state after the first k guesses have been verified as correct will almost certainly be supported on Fock states with total momentum 0. Total momentum 0, by Noether's theorem, implies that any position guess will only succeed with negligible probability. While one can show using the $\widetilde{\mathbf{H}}_y$ formalism that the first guess will be on a total momentum 0 state, for future states this can only be guaranteed up to negligible additive error; the issue is that this error compounds, leading to a trivial probability upper bound. This issue forced us to construct a sampling upper bound technique that constructed a probability upper bound for all v samples at once using the quasi-even condensate structure.

The bosonic perspective. Once we accepted that the heart of the matter was to analyze queries to U, we turned to a wide array of standard quantum query techniques—compressed oracles, polynomial methods, adversary arguments—only to find each approach mired in technical complexity. The crux of the difficulty was purification: expressing S and U as parts of a single coherent quantum system.

In compressed oracle techniques, one would ideally purify S via independent amplitudes, e.g., sampling each element's inclusion in S with probability $\ell/2^n$. Yet this fails spectacularly for our setting, since each evaluation of U depends non-locally on *all* elements of S. The natural purification $(\cos\theta\,|0\rangle + \sin\theta\,|1\rangle)^{\otimes 2^n}$ for $\sin^2\theta \approx \ell/2^n$ simply cannot encode this dependency. We needed a new formalism.

After much frustration with indices and combinatorial cases (often arising from identical versus distinct indices in sums), we found relief in the bosonic perspective. The key insight was to treat the oracle's purification as a system of indistinguishable bosons whose spatial locations encode membership in S. This eliminated the combinatorial clutter of index management and replaced it with clean algebraic manipulations via creation and annihilation operators.

Within this framework, we were able to reinterpret and re-derive earlier results—such as those of Hamoudi and Magniez [HM23]—in the language of bosons, gaining both conceptual clarity and technical leverage. Moreover, the physical picture aligned beautifully with the intended semantics of the quantum witness: the prover's ideal witness is $|S\rangle$, the uniform superposition over elements of S. In the bosonic purification, S directly corresponds to the positions of the bosons, so the verifier's measurement effectively "grabs, in superposition," a uniformly random boson and records its position. This unifies the intuition that a quantum witness can represent the entire set S at once, while a classical witness must enumerate its elements.

Challenges within the bosonic formulation. Despite its elegance, the bosonic formalism introduced its own difficulties. Our initial hope was that the oracle queries would assume a particularly simple and structured form in the momentum basis. Early calculations suggested that queries to U could be expressed via the *double momentum hopping operator* $\widetilde{\mathbf{H}}_y$, which preserves total momentum. By analogy with Noether's theorem [Noe18], we could then argue that since the total momentum remains conserved (and initially zero), the success probability of the first "guess" at a boson's position should equal $\ell/2^n$. This was encouraging—it suggested an inductive argument might succeed.

However, post-selection proved fatal to this line of reasoning. Conditioning on a successful guess perturbs the total momentum, breaking the conservation structure and rendering straightforward induction impossible. We therefore needed a more delicate understanding of how the oracle queries affect the system's momentum distribution.

A central obstacle was the definition of U itself: each element y must be included with probability proportional to $\gamma_y^{(S)}$, the squared Hadamard amplitude of $|S\rangle$. A first question that needed resolving was why $\gamma_y^{(S)}$ was the correct parameter to use for sampling U. In particular, as previously noted, using the "raw" Hadamard amplitudes of $|S\rangle$ leaks the sign information in the Fourier basis, which is the crucial information for approximately synthesizing $|S\rangle$ in the Fourier basis. We suspect that our choice of $\gamma_y^{(S)}$ isn't unique, but we do not have a second example. $\gamma_y^{(S)}$ gave us enough trouble as it is. While $\gamma_y^{(S)}$ behaves roughly like the square of a Gaussian random variable, its unbounded support complicates any attempt to interpret it as a probability in $[0, 1]$. We explored several approaches—thresholding, randomized truncations, and fully quantum purifications—but most were unsatisfactory. Truncating isn't smooth, which yields difficulties in analysis, as we only knew of techniques for handling polynomials in the hopping operators. Another option is conditioning the distribution Strong on the $\gamma_y^{(S)}$ being bounded by say n^{100}. Chernoff bounds argue that this event occurs with probability only exponentially small. Unfortunately, conditioning the distribu-

tion on this event changes the initial state from a state of ℓ bosons in 0-momentum to a state exponentially close in additive error. However, as previously discussed, additive approximations are not strong enough to prove exponentially small probability upper bounds as the errors compound.

Ultimately, we adopted an exponential function, guided by the "flat-tail" polynomial approximations introduced by Narayanan [Nar24]. This choice was technically crucial, as it allowed us to approximate the exponential behavior using low-degree polynomials in the double momentum hopping operators. With this analytic machinery in hand, we could begin to generalize beyond the specific oracle construction. Standard intuition from quantum uncertainty principles suggests that any state supported on a few momentum modes should yield a sampling upper bound. Yet, since the algorithm can query all y simultaneously, this intuition breaks down. The key insight was that the double momentum hopping operator typically acts by moving a *pair* of bosons, especially when starting from a condensate. A useful way to reinterpret this phenomenon is through momentum conservation. Each oracle query preserves total momentum, which—by analogy to translational invariance in many-body systems—acts as a global symmetry constraining how amplitude can flow between momentum modes. In the initial condensate, the total momentum is zero, and so any admissible process must preserve this invariance. This observation already implies a tight bound on the probability of sampling a single correct element from S, as any non-zero-momentum component must arise from a compensating momentum elsewhere in the system. Extending this reasoning to many guesses required recognizing that the double-hopping operator effectively creates and annihilates *pairs* of equal-and-opposite momenta—"Cooper-like" pairs in condensed-matter language—so that quasi-even condensates retain their global symmetry across queries. In this case, since the Fourier transform is considered, a "Cooper-like" pair is a pair of bosons in the same mode. It is precisely this paired structure, rather than a literal recording of queries in U, that allowed us to control the growth of odd momentum modes and prove the quasi-even property.

Finally, establishing that the quasi-even condensate property persists after a polynomial number of queries required substantial new machinery. Our bounds are almost certainly suboptimal, but they suffice to complete the proof. To our knowledge, this represents one of the first quantum query lower bounds addressing an oracle with as much internal quantum structure as U. Consequently, the classical tools of quantum query complexity—those tailored to unstructured oracles such as Grover's search or collision-finding—were insufficient, necessitating the new bosonic framework developed here.

Connections to set-size estimation and average-case NP In Section §7, we define the Strong distribution, a distribution over "yes" instances of spectral Forrelation, and implicitly define "no" instances of spectral Forrelation to have S with size $\leq \ell/100$. In addition to solving the spectral Forrelation problem, an algorithm that could distinguish between sets of size ℓ and size $\ell/100$ would be able to distinguish between the particular yes and no instances we use to separate QMA from QCMA. This puts our problem in line with the previous attempts of Fefferman and Kimmel [FK15] and Natarajan and Nirkhe [NN24], who more directly used the problem of subset size checking to separate QMA from QCMA relative to non-standard oracles. However, [FK15] notes that the problem of subset size checking is in AM due to Stockmeyer counting[14]. This further implies that the problem of set size checking (and the related problem

[14]Roughly speaking, in a Stockmeyer approximate counting protocol [Sto83], Arthur's public key message is interpreted as the key k for a hash function $h : \{0,1\}^n \to \{0,\ldots,\ell-1\}$ and Merlin is tasked with answering with $x \in S$ such that $h(x) = 0$. Merlin will

of distinguishing between our choice of yes and no instances) is in average-case NP, because by fixing the randomness of Stockmeyer counting to some optimal string, a deterministic polynomial-time classical verifier distinguishes between small and large sets for most small and large sets.

On the surface, this seems to imply a problem with our proof technique, since Theorem 9.1 is an average-case statement over the Strong distribution, and there is an average-case NP verifier for that distribution. We note that the prior works do not run into this problem as they directly apply the adversary method and diagonalization to the problem of subset size checking (given an in-place permutation oracle), and thus do not talk about a fixed distribution over yes and no instances of the subset size checking problem.

The resolution for this paper is the following subtle point: while Theorem 9.1 is an average case statement, the construction of the sampler from Theorem 6.2 requires that the QCMA distinguishing algorithm is able to distinguish between *all* yes and no instances. Even if the average-case NP verifier correctly outputs "no" on a $1 - 2^{-n}$ fraction of small sets, this will not be enough to get a successful sampler. To see this, we note that the successful sampler that we get from a QCMA verifier only samples v points with probability $\left(\frac{1}{36t^2}\right)^v$ – an incredibly small probability. This is consistent with the verifier only outputting v many points from that fixed small set Δ that is misclassified by the average-case NP verifier. The fact that our reduction to a sampler *requires* a worst-case distinguishing algorithm for our version of the set size checking problem (where the verifier is additionally given access to U) is the reason we avoid this problem, but also requires us to use a more elaborate diagonalization argument (than that of Aaronson and Kuperberg [AK07]) in to prove Theorem 1.1, as our property-testing contradiction only proves the existence of one misclassified instance per input size n.

16 Acknowledgments

We thank James Bartusek, Andrea Coladangelo, Uma Girish, Andrew Huang, Jonas Helsen, William Kretschmer, Anand Natarajan, Barak Nehoran, Fermi Ma, Er-cheng Tang, Umesh Vazirani, and Henry Yuen for helpful discussions. We additionally thank Fermi Ma and Andrew Huang for suggesting comments on an early draft of the result that greatly improved the presentation. We thank Nicholas Kocurek and Joe Slote for assistance in creating figures.

JB is supported by Henry Yuen's AFORS award FA9550-21-1-036 and NSF CAREER award CCF2144219. JH acknowledges funding from the Harvard Quantum Initiative postdoctoral fellowship. This work was partially completed while all authors were participants in the *Simons Institute for the Theory of Computing* Summer Clustering on Quantum Computing.

be able to answer with probability $\approx 1 - e^{-1} \geq 0.63$ when $|S| \approx \ell$ and will only be able to answer with probability $1 - e^{-1/100} \leq 0.01$ when $|S| \leq \ell/100$.

As the hard instances we construct are based on set-size estimation, the AM containment holds for our construction. (Meanwhile, we do not know (nor believe) that the generic problem of spectral Forrelation is in AM as not all spectral Forrelation instances are variants of set-size estimation.) However, since the problem is in AM, the problem is also in average-case NP, by simply amplifying the AM protocol and then fixing the choice of randomness of Arthur. So, there will exist no instances that are falsely accepted; however, this will only occur for an $\exp(-n)$ small fraction of the no instances.

References

[AA09] Scott Aaronson and Andris Ambainis. The need for structure in quantum speedups. *arXiv preprint arXiv:0911.0996*, 2009.

[AA15] Scott Aaronson and Andris Ambainis. Forrelation: A problem that optimally separates quantum from classical computing. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 307–316, 2015.

[Aar10] Scott Aaronson. BQP and the polynomial hierarchy. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 141–150, 2010.

[Aar21] Scott Aaronson. Open problems related to quantum query complexity. *ACM Transactions on Quantum Computing*, 2(4):1–9, 2021.

[ABN23] 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.

[AC13] Scott Aaronson and Paul Christiano. Quantum Money from Hidden Subspaces. *Theory of Computing*, 9(9):349–401, 2013.

[AGL25] Prabhanjan Ananth, Aditya Gulati, and Yao-Ting Lin. On the Limitations of Pseudorandom Unitaries. Cryptology ePrint Archive, Paper 2025/1785, 2025.

[AK07] 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. IEEE, 2007.

[AK25] Avantika Agarwal and Srijita Kundu. A Cautionary Note on Quantum Oracles, 2025.

[AKKT20] Scott Aaronson, Robin Kothari, William Kretschmer, and Justin Thaler. Quantum Lower Bounds for Approximate Counting via Laurent Polynomials. In Shubhangi Saraf, editor, *35th Computational Complexity Conference (CCC 2020)*, volume 169 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 7:1–7:47, Dagstuhl, Germany, 2020. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[Amb02] Andris Ambainis. Quantum Lower Bounds by Quantum Arguments. *J. Comput. Syst. Sci.*, 64(4):750–767, jun 2002.

[AN02] Dorit Aharonov and Tomer Naveh. Quantum NP-a survey. *arXiv preprint quant-ph/0210077*, 2002.

[Bar25] Mohammed Barhoush. Separating pseudorandom generators from logarithmic pseudorandom states. Cryptology ePrint Archive, Paper 2025/1994, 2025.

[BB14] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical Computer Science*, 560:7–11, 2014. Theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84.

[BBBV97] 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.

[BCN25] John Bostanci, Boyang Chen, and Barak Nehoran. Oracle Separation Between Quantum Commitments and Quantum One-wayness. In *Eurocrypt 2025*, 2025.

[BDK24] Shalev Ben-David and Srijita Kundu. Oracle separation of QMA and QCMA with bounded adaptivity. *arXiv preprint arXiv:2402.00298*, 2024.

[BLMT24] Ainesh Bakshi, Allen Liu, Ankur Moitra, and Ewin Tang. Learning quantum Hamiltonians at any temperature in polynomial time. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1470–1477, 2024.

[BMM+25] Amit Behera, Giulio Malavolta, Tomoyuki Morimae, Tamer Mour, and Takashi Yamakawa. A New World in the Depths of Microcrypt: Separating OWSGs and Quantum Money from QEFID. In *Eurocrypt 2025*, 2025.

[CCS25] Boyang Chen, Andrea Coladangelo, and Or Sattath. The power of a single Haar random state: constructing and separating quantum pseudorandomness. In *Eurocrypt 2025*, 2025.

[CM16] Toby Cubitt and Ashley Montanaro. Complexity Classification of Local Hamiltonian Problems. *SIAM Journal on Computing*, 45(2):268–316, 2016.

[FK15] Bill Fefferman and Shelby Kimmel. Quantum vs classical proofs and subset verification. *arXiv preprint arXiv:1510.06750*, 2015.

[GLLZ21] Siyao Guo, Qian Li, Qipeng Liu, and Jiapeng Zhang. Unifying presampling via concentration bounds. In *Theory of Cryptography Conference*, pages 177–208. Springer, 2021.

[GLMY25] Aditya Gulati, Yao-Ting Lin, Tomoyuki Morimae, and Shogo Yamada. Black-Box Separation Between Pseudorandom Unitaries, Pseudorandom Isometries, and Pseudorandom Function-Like States. Cryptology ePrint Archive, Paper 2025/1864, 2025.

[GS86] S Goldwasser and M Sipser. Private coins versus public coins in interactive proof systems. In *Proceedings of the Eighteenth Annual ACM Symposium on Theory of Computing*, STOC '86, page 59–68, New York, NY, USA, 1986. Association for Computing Machinery.

[GZ25] Eli Goldin and Mark Zhandry. Translating Between the Common Haar Random State Model and the Unitary Model. In *CRYPTO 2025*, 2025.

[Hei27] W. Heisenberg. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, 43(3):172–198, 1927.

[HM23] Yassine Hamoudi and Frédéric Magniez. Quantum Time–Space Tradeoff for Finding Multiple Collision Pairs. *ACM Transactions on Computation Theory*, 15(1-2):1–22, 2023.

[INN+21] Sandy Irani, Anand Natarajan, Chinmay Nirkhe, Sujit Rao, and Henry Yuen. Quantum search-to-decision reductions and the state synthesis problem. *arXiv preprint arXiv:2111.02999*, 2021.

[JKNN12] Stephen P. Jordan, Hirotada Kobayashi, Daniel Nagaj, and Harumichi Nishimura. Achieving perfect completeness in classical-witness quantum merlin-arthur proof systems. *Quantum Info. Comput.*, 12(5–6):461–471, May 2012.

[Kre21] William Kretschmer. Quantum Pseudorandomness and Classical Complexity. In Min-Hsiu Hsieh, editor, *16th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2021)*, volume 197 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 2:1–2:20, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[Liu22] Qipeng Liu. Non-uniformity and Quantum Advice in the Random Oracle Model. *Cryptology ePrint Archive*, 2022.

[LLPY23] Xingjian Li, Qipeng Liu, Angelos Pelecanos, and Takashi Yamakawa. Classical vs Quantum Advice and Proofs under Classically-Accessible Oracle. *arXiv preprint arXiv:2303.04298*, 2023.

[LMY25] Jiahui Liu, Saachi Mutreja, and Henry Yuen. QMA vs QCMA and Pseudorandomness. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 1520–1531, New York, NY, USA, 2025. Association for Computing Machinery.

[Lut11] Andrew Lutomirski. Component mixers and a hardness result for counterfeiting quantum money. *arXiv preprint arXiv:1107.0321*, 2011.

[Mah18] Urmila Mahadev. *Classical Verification and Blind Delegation of Quantum Computations*. PhD thesis, EECS Department, University of California, Berkeley, Jun 2018.

[Mer90] N David Mermin. Simple unified form for the major no-hidden-variables theorems. *Physical review letters*, 65(27):3373, 1990.

[MN25] Henry Ma and Anand Natarajan. Two bases suffice for QMA1-completeness, 2025.

[MW05] Chris Marriott and John Watrous. Quantum arthur–merlin games. *computational complexity*, 14(2):122–152, Jun 2005.

[Nar24] Shyam Narayanan. Improved algorithms for learning quantum Hamiltonians, via flat polynomials. *arXiv preprint arXiv:2407.04540*, 2024.

[NN24] Anand Natarajan and Chinmay Nirkhe. A distribution testing oracle separation between QMA and QCMA. *Quantum*, 8:1377, 2024.

[Noe18] E. Noether. Invariante Variationsprobleme. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1918:235–257, 1918.

[NZ24] Barak Nehoran and Mark Zhandry. A computational separation between quantum no-cloning and no-telegraphing. In *15th Innovations in Theoretical Computer Science Conference (ITCS 2024)*. Schloss-Dagstuhl-Leibniz Zentrum für Informatik, 2024.

[Per90] Asher Peres. Incompatible results of quantum measurements. *Physics Letters A*, 151(3-4):107–108, 1990.

[Sto83] Larry Stockmeyer. The complexity of approximate counting. In *Proceedings of the Fifteenth Annual ACM Symposium on Theory of Computing*, STOC '83, page 118–126, New York, NY, USA, 1983. Association for Computing Machinery.

[SV+14] Sushant Sachdeva, Nisheeth K Vishnoi, et al. Faster algorithms via approximation theory. *Foundations and Trends® in Theoretical Computer Science*, 9(2):125–210, 2014.

[Wie83] Stephen Wiesner. Conjugate coding. *SIGACT News*, 15(1):78–88, January 1983.

[Wik25] Wikipedia. Stars and bars (combinatorics) — Wikipedia, the free encyclopedia, 2025. [Online; accessed 16-October-2025].

[Wil17] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2 edition, 2017.

[Yao93] Andrew Chi-Chih Yao. Quantum Circuit Complexity. In *34th Annual Symposium on Foundations of Computer Science, Palo Alto, California, USA, 3-5 November 1993*, pages 352–361. IEEE Computer Society, 1993.

[YZ24] Takashi Yamakawa and Mark Zhandry. Verifiable quantum advantage without structure. *Journal of the ACM*, 71(3):1–50, 2024.

[Zha19] Mark Zhandry. How to record quantum queries, and applications to quantum indifferentiability. In *Advances in Cryptology–CRYPTO 2019: 39th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 18–22, 2019, Proceedings, Part II 39*, pages 239–268. Springer, 2019.

[Zha25] Mark Zhandry. Toward Separating QMA from QCMA with a Classical Oracle. In Raghu Meka, editor, *16th Innovations in Theoretical Computer Science Conference (ITCS 2025)*, volume 325 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 95:1–95:19, Dagstuhl, Germany, 2025. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il