The Quantum Supremacy Tsirelson Inequality

William Kretschmer

Department of Computer Science, The University of Texas at Austin, Austin, TX 78712, USA

A leading proposal for verifying near-term quantum supremacy experiments on noisy random quantum circuits is linear cross-entropy benchmarking. For a quantum circuit C on n qubits and a sample $z \in \{0,1\}^n$, the benchmark involves computing $|\langle z|C|0^n\rangle|^2$, i.e. the probability of measuring z from the output distribution of C on the all zeros input. Under a strong conjecture about the classical hardness of estimating output probabilities of quantum circuits, no polynomial-time classical algorithm given C can output a string z such that $|\langle z|C|0^n\rangle|^2$ is substantially larger than $\frac{1}{2^n}$ (Aaronson and Gunn, 2019). On the other hand, for a random quantum circuit C, sampling z from the output distribution of C achieves $|\langle z|C|0^n\rangle|^2 \approx \frac{2}{2^n}$ on average (Arute et al., 2019).

In analogy with the Tsirelson inequality from quantum nonlocal correlations, we ask: can a polynomial-time quantum algorithm do substantially better than $\frac{2}{2^n}$? We study this question in the query (or black box) model, where the quantum algorithm is given oracle access to C. We show that, for any $\varepsilon \geq \frac{1}{\text{poly}(n)}$, outputting a sample z such that $|\langle z|C|0^n\rangle|^2 \geq \frac{2+\varepsilon}{2^n}$ on average requires at least $\Omega\left(\frac{2^{n/4}}{\text{poly}(n)}\right)$ queries to C, but not more than $O\left(2^{n/3}\right)$ queries to C, if C is either a Haar-random n-qubit unitary, or a canonical state preparation oracle for a Haar-random n-qubit state. We also show that when C samples from the Fourier distribution of a random Boolean function, the naive algorithm that samples from C is the optimal 1-query algorithm for maximizing $|\langle z|C|0^n\rangle|^2$ on average.

1 Introduction

A team based at Google has claimed the first experimental demonstration of quantum computational supremacy on a programmable device [10]. The experiment involved *random circuit sampling*, where the task is to sample (with nontrivial fidelity) from the output distribution of a quantum circuit containing random 1- and 2-qubit gates. To verify their experiment, they used the so-called *Linear Cross-Entropy Benchmark*, or Linear XEB. Specifically, for an n-qubit quantum circuit C and samples $z_1, \ldots, z_k \in \{0,1\}^n$, the benchmark is given by:

$$b = \frac{2^n}{k} \cdot \sum_{i=1}^{k} |\langle z_i|C|0^n\rangle|^2.$$

The goal is for b to be large with high probability over the choice of the random circuit and the randomness of the sampler, as this demonstrates that the observations tend to concentrate on the outputs that are more likely to be measured under the ideal distribution for C (i.e. the noiseless distribution in which z is measured with probability $|\langle z|C|0^n\rangle|^2$). We formalize this task as the b-XHOG task:

Problem 1 (b-XHOG, or Linear Cross-Entropy Heavy Output Generation). *Fix a distribution $\mathcal{D}$ over quantum circuits on n-qubits.[1] Given a quantum circuit C sampled from $\mathcal{D}$, output a sample $z \in \{0,1\}^n$ such that:*

$$\mathop{\mathbb{E}}_{C \sim \mathcal{D}} \left[|\langle z|C|0^n\rangle|^2 \right] \geq \frac{b}{2^n}.$$

We emphasize that in this work, b-XHOG is fundamentally an average-case problem: it is always defined with respect to a distribution $\mathcal{D}$ over the choice of circuit C, which is why we require that the Linear XEB score exceeds b *in expectation*. Note also that the procedure for solving XHOG could itself be randomized (e.g. if it involves sampling z from the output of a quantum device), so the expectation in the above definition is taken with respect to this randomness as well.

For the purposes of demonstrating quantum supremacy, we will typically think of "large" b as any b bounded away from 1, as guessing z uniformly at random achieves $b = 1$ (regardless of the distribution $\mathcal{D}$!), just because $\sum_{z \in \{0,1\}^n} |\langle z|C|0^n\rangle|^2 = 1$ for any n-qubit circuit C. On the other hand, sampling from the noiseless output distribution of C achieves $b \approx 2$ when $\mathcal{D}$ selects random circuits of polynomial size and sufficient depth. Indeed, sampling from C achieves $b \approx 2$ whenever the circuit distribution $\mathcal{D}$ empirically exhibits the *Porter-Thomas* distribution on circuit output probabilities, in which the output probabilities are approximately i.i.d. exponential random variables [3, 10].

Under a strong complexity-theoretic conjecture about the classical hardness of nontrivially estimating output probabilities of quantum circuits, Aaronson and Gunn showed that no classical polynomial-time algorithm can solve b-XHOG for any $b \geq 1 + \frac{1}{\text{poly}(n)}$ on random quantum circuits of polynomial size [3]. Thus, a physical quantum computer that solves b-XHOG for $b \geq 1 + \Omega(1)$ is considered strong evidence of quantum computational supremacy.

In this work, we ask: can an efficient quantum algorithm for b-XHOG do substantially better than $b = 2$?[2] That is, what is the largest b for which a polynomial-time quantum algorithm can solve b-XHOG on random circuits? Note that the largest b we could hope for is achieved by the optimal sampler that always outputs the string z maximizing $|\langle z|C|0^n\rangle|^2$. If the random circuits induce a Porter-Thomas distribution on output probabilities, then this solves b-XHOG for $b = \Theta(n)$ (see Fact 13 below). However, finding the largest output probability might be computationally difficult even on a quantum computer, which is why we restrict our attention to *efficient* quantum algorithms.

We refer to our problem as the "quantum supremacy Tsirelson inequality" in reference to the Bell [12] and Tsirelson [19] inequalities for quantum nonlocal correlations (for a modern overview, see [21]). Under this analogy, the quantity b in XHOG plays a similar role as the probability p of winning some nonlocal game. For example, the Bell inequality for the CHSH game [20] states that no classical strategy can win the game with probability $p > \frac{3}{4}$; we view this as analogous

---

[1] We will sometimes leave the choice of $\mathcal{D}$ implicit when it is clear from context.

[2] We thank Scott Aaronson [1] for raising this question, and for suggesting the analogy with the Tsirelson inequality.

to the conjectured inability of efficient classical algorithms to solve b-XHOG for any $b > 1$. By contrast, a quantum strategy with pre-shared entanglement allows players to win the CHSH game with probability $p = \cos^2\left(\frac{\pi}{8}\right) \approx 0.854 > \frac{3}{4}$. An experiment that wins the CHSH game with probability $p > \frac{3}{4}$, a violation of the Bell inequality, is analogous to an experimental demonstration of b-XHOG for $b > 1$ on a quantum computer that establishes quantum computational supremacy. Finally, the Tsirelson inequality for the CHSH game states that any quantum strategy involving arbitrary pre-shared entanglement wins with probability $p \leq \cos^2\left(\frac{\pi}{8}\right)$. Hence, an upper bound on b for efficient quantum algorithms is the quantum supremacy counterpart to the Tsirelson inequality. We emphasize that our choice to refer to this as a "Tsirelson inequality" is purely by analogy; we do not claim that the question involving quantum supremacy or the techniques one might use to answer it are otherwise related to quantum nonlocal correlations.

1.1 Our Results

We study the quantum supremacy Tsirelson inequality in the quantum query (or black box) model. That is, we consider b-XHOG with respect to distributions $\mathcal{D}$ that take the following form. We fix a quantum circuit C that queries a classical or quantum oracle $\mathcal{O}$. To sample C from $\mathcal{D}$, we choose $\mathcal{O}$ according to some distribution over oracles. We then ask what is the largest b such that b-XHOG with respect to $\mathcal{D}$ is solvable with polynomially many queries to $\mathcal{O}$.

Our motivation for studying this problem in the query model is twofold. First, quantum query results often give useful intuition for what to expect in the real world, and can provide insight into why naive algorithmic approaches fail. Second, we view this as an interesting quantum query complexity problem in its own right. Whereas most other quantum query lower bounds involve decision problems [6] or relation problems [13], XHOG is more like a weighted, average-case relation problem, because we only require that $|\langle z|C|0^n\rangle|^2$ be large *on average*. Contrast this with the relation problem considered in [2], where the task is to output a z such that $|\langle z|C|0^n\rangle|^2$ is greater than some threshold.

Note that there are known quantum query complexity lower bounds for relation problems [10], and even relation problems where the output is a quantum state [7, 27]. Yet, it is unclear whether existing quantum query lower bound techniques are useful here. Whereas the adversary method tightly characterizes the quantum query complexity of decision problems and state conversion problems [26], it is not known to characterize the query complexity of relation problems, unless they are efficiently verifiable [13]. The adversary method appears to be essentially useless for saying anything about XHOG, which is not efficiently verifiable and is not a relation problem in the traditional sense.[3]

The XHOG task is well-defined for any distribution of random quantum circuits, so this gives us a choice in selecting the distribution. We focus on three classes of oracle circuits that either resemble random circuits used in practical experiments, or that were previously studied in the context of quantum supremacy. Formal definitions of these oracles (and the associated versions of XHOG) are given in Section 2.2.

Canonical State Preparation Oracles   Because the linear cross-entropy benchmark for a circuit C depends only on the state $|\psi\rangle := C|0^n\rangle$ produced by the circuit on the all zeros input, it is natural to consider an oracle $\mathcal{O}_\psi$ that prepares a random state $|\psi\rangle$ without leaking additional

---

[3]As we will see later, however, the polynomial method [11] plays an important role in one of our results.

information about $|\psi\rangle$. Formally, we choose a Haar-random n-qubit state $|\psi\rangle$, and fix a canonical state $|\bot\rangle$ orthogonal to all n-qubit states.[4] Then, we take the oracle $\mathcal{O}_\psi$ that acts as $\mathcal{O}_\psi|\bot\rangle = |\psi\rangle$, $\mathcal{O}_\psi|\psi\rangle = |\bot\rangle$, and $\mathcal{O}_\psi|\varphi\rangle = |\varphi\rangle$ for any state $|\varphi\rangle$ that is orthogonal to both $|\bot\rangle$ and $|\psi\rangle$. Equivalently, $\mathcal{O}_\psi$ is the reflection about the state $\frac{|\psi\rangle - |\bot\rangle}{\sqrt{2}}$. Finally, we let C be the composition of $\mathcal{O}_\psi$ with any unitary that sends $|0^n\rangle$ to $|\bot\rangle$, so that $C|0^n\rangle = |\psi\rangle$. This model is often chosen when proving lower bounds for quantum algorithms that query state preparation oracles [4, 8, 14], in part because the ability to simulate $\mathcal{O}_\psi$ follows in a completely black box manner from the ability to prepare $|\psi\rangle$ unitarily without garbage (see Lemma 7 below). Hence, the oracle $\mathcal{O}_\psi$ is "canonical" in the sense that it is uniquely determined by $|\psi\rangle$ and is not any more powerful than any other oracle that prepares $|\psi\rangle$ without garbage.

Haar-Random Unitaries   A random polynomial-size quantum circuit C does not behave like a canonical state preparation oracle: $C|x\rangle$ looks like a random quantum state for *any* computational basis state $|x\rangle$, not just $x = 0^n$. Indeed, random quantum circuits are known to information-theoretically approximate the Haar measure in certain regimes [15, 22], and it seems plausible that they are also computationally difficult to distinguish from the Haar measure. Thus, one could alternatively model random quantum circuits by Haar-random n-qubit unitaries.

Fourier Sampling Circuits   Finally, we consider quantum circuits that query a random *classical* oracle. For this, we use FOURIER SAMPLING circuits, which Aaronson and Chen [2] previously studied in the context of proving oracular quantum supremacy for a problem related to XHOG. FOURIER SAMPLING circuits are defined as $H^{\otimes n} U_f H^{\otimes n}$, where U_f is a phase oracle for a uniformly random Boolean function $f : \{0,1\}^n \to \{-1, 1\}$. On the all-zeros input, FOURIER SAMPLING circuits output a string $z \in \{0,1\}^n$ with probability proportional to the squared Fourier coefficient $\hat{f}(z)^2$. This model has the advantage that in principle, one can prove the corresponding quantum supremacy Bell inequality for classical algorithms given query access to f, and that in some cases one can replace f by a pseudorandom function to base quantum supremacy on cryptographic assumptions [2].

Our first result is an exponential lower bound on the number of quantum queries needed to solve $(2 + \varepsilon)$-XHOG given either of the two types of quantum oracles that we consider:

Theorem 2 (Informal version of Theorem 17 and Theorem 20). *For any $\varepsilon \geq \frac{1}{\text{poly}(n)}$, any quantum query algorithm for $(2 + \varepsilon)$-XHOG with query access to either:*

(1) a canonical state preparation oracle $\mathcal{O}_\psi$ for a Haar-random n-qubit state $|\psi\rangle$, or

(2) a Haar-random n-qubit unitary,

requires at least $\Omega\left(\frac{2^{n/4}}{\text{poly}(n)}\right)$ queries.

Recall that, because Haar-random states induce a Porter-Thomas distribution on measurement probabilities, the naive algorithm that outputs a sample from the measurement distribution of the state solves b-XHOG for $b \approx 2$. Hence, in the black box setting, Theorem 2 implies that it is

---

[4]We can always assume that a convenient $|\bot\rangle$ exists by extending the Hilbert space, if needed. For example, if $|\psi\rangle$ is an n-qubit state, a natural choice is to encode $|\psi\rangle$ by $|\psi\rangle|1\rangle$ and to choose $|\bot\rangle = |0^n\rangle|0\rangle$.

computationally difficult to substantially beat the naive algorithm for XHOG. We do not know if Theorem 2 is quantitatively optimal, but we show in Theorem 18 that a simple algorithm based on the quantum collision finding algorithm [16] solves $(2 + \Omega(1))$-XHOG using $O\left(2^{n/3}\right)$ queries to either oracle.

Finally, we show that for FOURIER SAMPLING circuits, the naive algorithm of simply running the circuit is optimal among all 1-query algorithms:

Theorem 3 (Informal version of Theorem 22). *Any 1-query quantum algorithm for b-XHOG with FOURIER SAMPLING circuits achieves $b \leq 3$.*

Note that the value of b achieved by the naive quantum algorithm for XHOG depends on the distribution of circuits used. In contrast to Haar-random circuits that achieve $b \approx 2$, FOURIER SAMPLING circuits achieve $b \approx 3$ (see Proposition 21). This stems from the fact that the amplitudes of a Haar-random quantum state are approximately distributed as *complex* normal random variables, whereas the amplitudes of a state produced by a random FOURIER SAMPLING circuit are approximately distributed as *real* normal random variables.

1.2 Our Techniques

The starting point for our proof of the Tsirelson inequality with a canonical state preparation oracle $\mathcal{O}_\psi$ is a result of Ambainis, Rosmanis, and Unruh [8]. It shows that any algorithm that queries $\mathcal{O}_\psi$ can be approximately simulated by a different algorithm that makes no queries, but starts with copies of a resource state that depends on $|\psi\rangle$. This resource state consists of polynomially many (in the number of queries to $\mathcal{O}_\psi$) states of the form $\alpha|\psi\rangle + \beta|\bot\rangle$, i.e. copies of $|\psi\rangle$ in superposition with $|\bot\rangle$. Our strategy is to show that if any algorithm solves b-XHOG given this resource state, then a similar algorithm solves b-XHOG given copies of $|\psi\rangle$ alone. Then, we prove a lower bound on the number of copies of $|\psi\rangle$ needed to solve b-XHOG. To do so, we argue that if $|\psi\rangle$ is Haar-random, then the best algorithm for b-XHOG given copies of $|\psi\rangle$ is a simple collision-finding algorithm: measure all copies of $|\psi\rangle$ in the computational basis, and output whichever string $z \in \{0,1\}^n$ appears most frequently in the measurement results. For a Haar-random n-qubit state, the chance of seeing *any* collisions is exponentially unlikely, unless the number of copies of $|\psi\rangle$ is exponentially large in n, and so this does not do much better than measuring a single copy of $|\psi\rangle$ and outputting the result.

To prove the analogous lower bound for b-XHOG with a Haar-random unitary oracle, we show more generally that the canonical state preparation oracles and Haar-random unitary oracles are essentially equivalent as resources, which may be of independent interest. More specifically, we show that for an n-qubit state $|\psi\rangle$, using a constant number of queries to $\mathcal{O}_\psi$, one can approximately simulate (to exponential precision) queries to a random oracle that prepares $|\psi\rangle$. By "random oracle that prepares $|\psi\rangle$," we mean an n-qubit unitary U_ψ that acts as $U_\psi|0^n\rangle = |\psi\rangle$ but Haar-random everywhere else. We can construct such a U_ψ by taking an arbitrary n-qubit unitary that maps $|0^n\rangle$ to $|\psi\rangle$, then composing it with a Haar-random unitary on the $(2^n - 1)$-dimensional subspace orthogonal to $|0^n\rangle$.

Our lower bound for FOURIER SAMPLING circuits uses an entirely different technique. We use the polynomial method of Beals et al. [11], which shows that for any quantum algorithm that makes T queries to a classical oracle, the output probabilities of the algorithm can be expressed as degree-$2T$ polynomials in the variables of the classical oracle. Our key observation is that the average linear XEB score achieved by such a quantum query algorithm can *also* be expressed as a polynomial in the variables of the classical oracle. We further observe that this polynomial is

constrained by the requirement that the polynomials representing the output probabilities must be nonnegative and sum to 1. This allows us to upper bound the largest linear XEB score achievable by the maximum value of a certain linear program, whose variables are the coefficients of the polynomials that represent the output probabilities of the algorithm. To upper bound this quantity, we exhibit a solution to the dual linear program.

2 Preliminaries

2.1 Notation

We use $[N]$ to denote the set $\{1, 2, \ldots, N\}$. We use $\mathbb{1}$ to denote the identity matrix (of implicit size). We let $\mathrm{TD}(\rho, \sigma)$ denote the trace distance between density matrices ρ and σ, and let $||A||_\diamond$ denote the diamond norm of a superoperator A acting on density matrices (see [5] for definitions). For a unitary matrix U, we use $U \cdot U^\dagger$ to denote the superoperator that maps ρ to $U\rho U^\dagger$. In a slight abuse of notation, if A denotes a quantum algorithm (which may consist of unitary gates, measurements, oracle queries, and initialization of ancilla qubits), then we also use A to denote the superoperator corresponding to the action of A on input density matrices.

2.2 Oracles for Quantum States

We frequently consider quantum algorithms that query quantum oracles. In this model, a query to a unitary matrix U consists of a single application of either U, $U^\dagger$, or controlled versions of U or $U^\dagger$. We also consider quantum algorithms that make queries to *random* oracles. In analogue with the classical random oracle model, such calls are not randomized at each query. Rather, a unitary U is chosen randomly (from some distribution) at the start of the execution of the algorithm, and thereafter all queries for the duration of the algorithm are made to U.

We now define several types of unitary oracles that we will use. These definitions (and associated lemmas giving constructions of them) have appeared implicitly or explicitly in prior work, e.g. [4, 8, 9, 14]. For completeness, we provide proofs of the constructions.

Definition 4. *For an n-qubit quantum state $|\psi\rangle$, the* reflection about $|\psi\rangle$*, denoted $\mathcal{R}_\psi$, is the n-qubit unitary $\mathcal{R}_\psi := \mathbb{1} - 2|\psi\rangle\langle\psi|$.*

In other words, $|\psi\rangle$ is a -1 eigenstate of $\mathcal{R}_\psi$, and all states orthogonal to $|\psi\rangle$ are $+1$ eigenstates. Note that some authors define the reflection about $|\psi\rangle$ to be the negation of this operator (e.g. [9, 28, 31]), while others follow our convention (e.g. [4, 17, 24]). This makes little difference, as these definitions are equivalent up to a global phase (or, if using the controlled versions, equivalent up to a Pauli Z gate).

The following lemma shows that $\mathcal{R}_\psi$ can be simulated given any unitary that prepares $|\psi\rangle$ from the all-zeros state, possibly with unentangled garbage.

Lemma 5. *Let U be a unitary that acts as $U|0^n\rangle|0^m\rangle = |\psi\rangle|\varphi\rangle$, where $|\psi\rangle$ and $|\varphi\rangle$ are n- and m-qubit states, respectively. Then one can simulate T queries to the reflection $\mathcal{R}_\psi$ using $2T + 1$ queries to U.*

Proof. Consider the unitary $U(\mathbb{1} - 2|0^{m+n}\rangle\langle 0^{m+n}|)U^\dagger$. For any n-qubit state $|x\rangle$, the action of this unitary on $|x\rangle|\varphi\rangle$ is equivalent to the action of $\mathcal{R}_\psi$ on $|x\rangle|\varphi\rangle$. So, we can simulate $\mathcal{R}_\psi$ on $|x\rangle$ as follows: first use one query to U to prepare $|\psi\rangle|\varphi\rangle$ from $|0^{m+n}\rangle$, so that we have a copy of $|\varphi\rangle$.

6

Then, simulate each query to $\mathcal{R}_\psi$ using a query to U and $U^\dagger$ to perform $U(\mathbb{1} - 2|0^{m+n}\rangle\langle 0^{m+n}|)U^\dagger$ applied to $|x\rangle|\varphi\rangle$, using the copy of $|\varphi\rangle$ prepared in the first step. ∎

Definition 6. *For a quantum state $|\psi\rangle$, the* canonical state preparation oracle *for $|\psi\rangle$, denoted $\mathcal{O}_\psi$, is the reflection about the state $\frac{|\psi\rangle - |\bot\rangle}{\sqrt{2}}$, where $|\bot\rangle$ is some canonical state orthogonal to $|\psi\rangle$.*

Unless otherwise specified, we generally assume that if $|\psi\rangle$ is an n-qubit state, then $|\bot\rangle$ is orthogonal to the space of n-qubit states under a suitable encoding (see Footnote 4).

The next lemma shows that $\mathcal{O}_\psi$ can be simulated from *any* oracle that prepares $|\psi\rangle$ without garbage:

Lemma 7. *Let U be an n-qubit unitary that satisfies $U|0^n\rangle = |\psi\rangle$. Then one can simulate T queries to $\mathcal{O}_\psi$ using $4T + 2$ queries to U.*

Proof. $|\bot\rangle$ is known, so we may assume that a known unitary V acts as $V|0^n\rangle = |\bot\rangle$. Because $\mathcal{O}_\psi$ is defined as the reflection about $\frac{|\psi\rangle - |\bot\rangle}{\sqrt{2}}$, by Lemma 5, it suffices to construct a unitary that prepares any state of the form $\frac{|\psi\rangle - |\bot\rangle}{\sqrt{2}}|\varphi\rangle$ from $|0^n\rangle|0^m\rangle$ using 2 queries to U. The following circuit accomplishes this, with $|\varphi\rangle = |0\rangle$:

∎
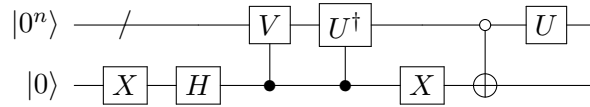
We introduce the notion of a *random* state preparation oracle, which, to our knowledge, is new.

Definition 8. *For an n-qubit state $|\psi\rangle$ we define a* random state preparation oracle *for $|\psi\rangle$, denoted U_ψ, as follows. We fix an arbitrary n-qubit unitary V that satisfies $V|0^n\rangle = |\psi\rangle$, then choose a Haar-random unitary W that acts on the $(2^n - 1)$-dimensional subspace orthogonal to $|0^n\rangle$ in the space of n-qubit states. Finally, we set $U_\psi = VW$.*

The invariance of the Haar measure guarantees that this distribution over U_ψ is independent of the choice of V, and hence this is well-defined. Note that while we often refer to U_ψ as a single unitary matrix, U_ψ really refers to a *distribution* over unitary matrices. Notice also that if $|\psi\rangle$ is distributed as a Haar-random n-qubit state, then U_ψ is distributed as a Haar-random n-qubit unitary.

With these definitions in hand, we can now formally define the three versions of the b-XHOG task (Problem 1) that we consider in this paper.

Problem 9 (b-XHOG with canonical state preparation)**.** *Let $|\psi\rangle$ be a Haar-random n-qubit state. Given oracle access to $\mathcal{O}_\psi$, output a sample $z \in \{0,1\}^n$ such that:*

$$\mathop{\mathbb{E}}_{|\psi\rangle}\left[|\langle z|\psi\rangle|^2\right] \geq \frac{b}{2^n}.$$

Problem 10 (b-XHOG with a Haar-random oracle)**.** *Let U be a Haar-random n-qubit unitary. Given oracle access to U, output a sample $z \in \{0,1\}^n$ such that:*

$$\mathop{\mathbb{E}}_{U}\left[|\langle z|U|0^n\rangle|^2\right] \geq \frac{b}{2^n}.$$

Equivalently, in the above definition, we can let $|\psi\rangle$ be a Haar-random state, choose $U = U_\psi$, and output a sample z such that:

$$\mathop{\mathbb{E}}_{|\psi\rangle}\left[|\langle z|\psi\rangle|^2\right] \geq \frac{b}{2^n}.$$

Problem 11 (*b*-XHOG with FOURIER SAMPLING circuits). *Let $f : \{0,1\}^n \to \{-1,1\}$ be a uniformly random Boolean function of n bits. Let U_f denote the phase oracle for f, meaning the unitary transformation which acts as $U_f|x\rangle = f(x)|x\rangle$ for any $x \in \{0,1\}^n$. Given oracle access to U_f, output a sample $z \in \{0,1\}^n$ such that:*

$$\mathop{\mathbb{E}}_{f}\left[|\langle z|H^{\otimes n}U_f H^{\otimes n}|0^n\rangle|^2\right] \geq \frac{b}{2^n},$$

where H is the Hadamard transformation.

2.3 Other Useful Facts

We use the following formula for the distance between unitary superoperators in the diamond norm.

Fact 12 ([5]). *Let V and W be unitary matrices, and suppose d is the distance between 0 and the polygon in the complex plane whose vertices are the eigenvalues of $VW^\dagger$. Then*

$$\left\|V \cdot V^\dagger - W \cdot W\right\|_\diamond = 2\sqrt{1 - d^2}.$$

Finally, we observe that for a Haar-random n-qubit quantum state, the information-theoretically largest linear XEB achievable is $O(n)$, and this is tight.

Fact 13. *Let $|\psi\rangle$ be a Haar-random n-qubit quantum state. Then:*

$$\mathop{\mathbb{E}}_{|\psi\rangle}\left[\max_{z\in\{0,1\}^n} |\langle z|\psi\rangle|^2\right] = \frac{\Theta(n)}{2^n}.$$

Proof sketch. For a Haar-random $|\psi\rangle$, the probabilities $|\langle z|\psi\rangle|^2$ follow a *Porter-Thomas distribution* [10], which is to say that they approach i.i.d. exponential random variables with mean $\frac{1}{2^n}$ in the limit. By a well-known result of Rényi [32], the maximum of N i.i.d. exponential random variables with mean μ is distributed as $\sum_{i=1}^{N} \frac{E_i}{i}$, where $E_1, \ldots, E_N$ are i.i.d. exponentially distributed with mean μ. In particular, the expected value of the maximum of N i.i.d. exponential random variables with mean μ is $H_N \cdot \mu$, where H_N is the Nth harmonic number. So, $\mathbb{E}\left[\max_{z\in\{0,1\}^n} |\langle z|\psi\rangle|^2\right]$ should approach $\frac{\Theta(n)}{2^n}$, because $H_N \approx \ln N$.

In reality, the probabilities $|\langle z|\psi\rangle|^2$ are not exactly i.i.d. exponentially distributed, but are distributed according to a Dirichlet distribution (in fact, uniform on the 2^n-dimensional probability simplex). This distribution can be sampled from as follows: sample $E_1, E_2, \ldots, E_{2^n}$ to be i.i.d. exponential random variables, and set $|\langle z|\psi\rangle|^2 = \frac{E_z}{\sum_{i=1}^{2^n} E_i}$. The same proof idea still works, essentially because the denominator $\sum_{i=1}^{2^n} E_i$ concentrates well (indeed, the denominator is exponentially close to 1 with high probability). ∎

3   Canonical State Preparation Oracles

In this section, we prove the quantum supremacy Tsirelson inequality for XHOG with a canonical state preparation oracle for a Haar-random state (Problem 9). We first sketch the important ideas in the proof. At the heart of our proof is the following lemma, due to Ambainis, Rosmanis, and Unruh [8]. It shows that any quantum algorithm that makes queries to a canonical state preparation oracle $\mathcal{O}_\psi$ can be approximately simulated by a quantum algorithm that makes no queries to $\mathcal{O}_\psi$, and instead receives various copies of $|\psi\rangle$ and superpositions of $|\psi\rangle$ with some canonical orthogonal state.

Lemma 14 ([8]). *Let A be a quantum query algorithm that makes T queries to $\mathcal{O}_\psi$. Then for any k, there is a quantum algorithm B that makes no queries to $\mathcal{O}_\psi$, and a quantum state $|R\rangle$ of the form:*

$$|R\rangle := \bigotimes_{j=1}^{k} \alpha_j |\psi\rangle + \beta_j |\perp\rangle$$

such that for any state $|\varphi\rangle$:

$$\mathrm{TD}(A(|\varphi\rangle\langle\varphi|), B(|R\rangle\langle R|, |\varphi\rangle\langle\varphi|)) \leq O\left(\frac{T}{\sqrt{k}}\right).$$

So long as $k \gg T^2$, the output of B will be arbitrarily close to the output of A in trace distance. We will use this and Fact 13 to show that if A solves b-XHOG for some $b > 2$, then so does B. Then, to prove a lower bound on the number of queries T to $\mathcal{O}_\psi$ needed to solve b-XHOG, it suffices to instead lower bound k, the number of states of the form $\alpha_j |\psi\rangle + \beta_j |\perp\rangle$ needed to solve b-XHOG.

When $|\psi\rangle$ is a Haar-random state, notice that the linear XEB depends only on the *magnitude* of the amplitudes in $|\psi\rangle$; the phases are irrelevant. So, when considering algorithms that attempt to solve b-XHOG given only a state $|R\rangle$ of the form used in Lemma 14, we might as well assume that the algorithm randomly reassigns the phases on $|\psi\rangle$. More formally, define the mixed state σ_R as

$$\sigma_R := \mathop{\mathbb{E}}_{\text{diagonal } U} \left[U^{\otimes k} |R\rangle\langle R| U^{\dagger \otimes k} \right], \tag{1}$$

where the expectation is over the diagonal unitaries U such that the entries $\langle i|U|i\rangle$ are i.i.d. uniformly random complex phases (and by convention, $\langle \perp|U|\perp\rangle = 1$). Then, the algorithm's average linear XEB score given σ_R for a random choice of $|\psi\rangle$ is identical to its average linear XEB score given $|R\rangle$ for a random choice of $|\psi\rangle$, because of the invariance of the Haar measure with respect to phases.

Next, we observe that one can prepare σ_R by measuring k copies of $|\psi\rangle$ in the computational basis. We prove this in Lemma 15. So, when considering algorithms for XHOG that start with $|R\rangle$, it suffices to instead consider algorithms that simply measure k copies of $|\psi\rangle$ in the computational basis. Such algorithms are much easier to analyze, because once we have measured the k copies of $|\psi\rangle$, we can assume (by convexity) that any optimal such algorithm for XHOG outputs a string z deterministically given the k measurement results. And in that case, clearly the optimal strategy is to output whichever z maximizes the posterior expectation of $|\langle z|\psi\rangle|^2$ given the measurement results. We analyze this strategy in Lemma 16, and show that roughly $2^{n/2}$ copies of $|\psi\rangle$ are needed to solve b-XHOG for b bounded away from 2. The intuition is that the posterior expectation of $|\langle z|\psi\rangle|^2$ increases only when we see z at least twice in the measurement results. However, the

probability that any two measurement results are the same is tiny—on the order of 2^{-n}—and so we need to measure at least $2^{n/2}$ copies of $|\psi\rangle$ to see any collisions with decent probability.

We now proceed to proving the necessary lemmas.

Lemma 15. *Let $|\psi\rangle = \sum_{i=1}^{N} \psi_i |i\rangle$ be an unknown quantum state, and consider a state $|R\rangle$ of the form:*

$$|R\rangle := \bigotimes_{j=1}^{k} \alpha_j |\psi\rangle + \beta_j |\perp\rangle,$$

where α_j, β_j are known for $j \in [k]$, and the vectors $\{|1\rangle, |2\rangle, \ldots, |N\rangle, |\perp\rangle\}$ form an orthonormal basis. Define the mixed state σ_R as above (1). Then there exists a protocol to prepare σ_R by measuring k copies of $|\psi\rangle$ in the computational basis.

To give some intuition, we note that it is simpler to prove Lemma 15 in the case where $\alpha_j = 1$ for all j. In that case, σ_R can be viewed as an $N^k \times N^k$ density matrix where both the rows and columns are indexed by strings in $[N]^k$. Then, the averaging over diagonal unitaries implies that σ_R is obtained from $(|R\rangle\langle R|)^{\otimes k}$ by zeroing out all entries where the index corresponding to the row is not a reordering of the index corresponding to the column. In fact, one can show that σ_R is expressible as a mixture of pure states, where each pure state is a uniform superposition over basis states that are reorderings of each other. Moreover, the probability associated with each pure state in this mixture is precisely the probability that one of the reorderings is observed when we measure k copies of $|\psi\rangle$ in the computational basis. So, to prepare σ_R, it suffices to measure $|\psi\rangle^{\otimes k}$ and then output the uniform superposition over reorderings of the measurement result.

The proof of Lemma 15 is similar, but we instead have to randomly set some of the measurement results to $\perp$ with probability $|\beta_j|^2$.

Proof of Lemma 15. We first describe the protocol. Define $[N_\perp] := [N] \cup \{\perp\}$. Measure $|\psi\rangle^{\otimes k}$ in the computational basis to obtain a string $\mathbf{x} \in [N]^k$. Then, sample a string $\overline{\mathbf{x}} \in [N_\perp]^k$ by setting

$$\overline{\mathbf{x}}_j = \begin{cases} \mathbf{x}_j & \text{with probability } |\alpha_j|^2 \\ \perp & \text{with probability } |\beta_j|^2 \end{cases}$$

independently for each $j \in [k]$. Let $\mathbf{Z} := \{z \in [N_\perp]^k : z \text{ is a reordering of } \overline{\mathbf{x}}\}$. For each $z \in \mathbf{Z}$ and $j \in [k]$, define

$$\gamma_{zj} := \begin{cases} \alpha_j & z_j \neq \perp \\ \beta_j & z_j = \perp. \end{cases}$$

Finally, prepare and output the state

$$|\zeta_{\mathbf{Z}}\rangle := \frac{\sum_{z \in \mathbf{Z}} \left(\prod_{j=1}^{k} \gamma_{zj} \right) |z\rangle}{\sqrt{\sum_{z \in \mathbf{Z}} \prod_{j=1}^{k} \gamma_{zj} \gamma_{zj}^*}}. \tag{2}$$

This allows us to express the density matrix ρ_R output by this protocol as follows:

$$\rho_R := \sum_{Z \subset [N_\perp]^k} \Pr[\mathbf{Z} = Z] \cdot |\zeta_Z\rangle\langle\zeta_Z|. \tag{3}$$

10

To complete the proof, we want to show that $\rho_R = \sigma_R$. To see that this holds, first consider an entry $\langle x|\sigma_R|y\rangle$ of σ_R, where $x, y \in [N_\perp]^k$. It is equal to

$$\langle x|\sigma_R|y\rangle = \underset{\text{diagonal } U}{\mathbb{E}} \left[\left(\prod_{j=1}^{k} \gamma_{xj}\gamma_{yj}^* \right) \cdot \left(\prod_{j:x_j \neq \perp} U_{x_j x_j}\psi_{x_j} \right) \cdot \left(\prod_{j:y_j \neq \perp} U_{y_j y_j}^* \psi_{y_j}^* \right) \right] \tag{4}$$

$$= \langle x|R\rangle\langle R|y\rangle \cdot \underset{\text{diagonal } U}{\mathbb{E}} \left[\left(\prod_{j:x_j \neq \perp} U_{x_j x_j} \right) \cdot \left(\prod_{j:y_j \neq \perp} U_{y_j y_j}^* \right) \right] \tag{5}$$

$$= \begin{cases} \langle x|R\rangle\langle R|y\rangle & x \text{ is a reordering of } y \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Here, (4) and (5) are simple calculations that follow from the definitions of $|R\rangle$ and σ_R. In (6), we use the fact that the entries U_{ii} are independent, uniformly random complex units, and so $\mathbb{E}[U_{ii}^a U_{ii}^{*b}] = \mathbb{E}[U_{ii}^{a-b}]$ is 1 if $a = b$ and 0 otherwise, for positive integers a, b. Also, if $i \neq j$, then $\mathbb{E}[U_{ii}^a U_{jj}^{*b}] = 0$ unless $a = b = 0$.

Evidently, $\langle x|\rho_R|y\rangle = \langle x|\sigma_R|y\rangle = 0$ whenever x is not a reordering of y, because ρ_R is a mixture of pure states, each of which is a superposition of basis states that are reorderings of one another. So, it remains to show that $\langle x|\rho_R|y\rangle = \langle x|\sigma_R|y\rangle = \langle x|R\rangle\langle R|y\rangle$ whenever x is a reordering of y. Let $Z := \{z \in [N_\perp]^k : z \text{ is a reordering of } x\}$. Then:

$$\langle x|\rho_R|y\rangle = \Pr[\mathbf{Z} = Z] \cdot \langle x|\zeta_Z\rangle\langle\zeta_Z|y\rangle \tag{7}$$

$$= \left(\sum_{z \in Z} \Pr[\overline{\mathbf{x}} = z] \right) \cdot \langle x|\zeta_Z\rangle\langle\zeta_Z|y\rangle \tag{8}$$

$$= \left(\sum_{z \in Z} \left(\prod_{j=1}^{m} \gamma_{zj}\gamma_{zj}^* \right) \left(\prod_{j:z_j \neq \perp} \psi_{z_j}\psi_{z_j}^* \right) \right) \cdot \frac{\prod_{j=1}^{m} \gamma_{xj}\gamma_{yj}^*}{\sum_{z \in Z} \prod_{j=1}^{m} \gamma_{zj}\gamma_{zj}^*} \tag{9}$$

$$= \left(\prod_{j=1}^{m} \gamma_{xj}\gamma_{yj}^* \right) \cdot \left(\prod_{j:x_j \neq \perp} \psi_{x_j}\psi_{x_j}^* \right) \tag{10}$$

$$= \langle x|R\rangle\langle R|y\rangle \tag{11}$$

$$= \langle x|\sigma_R|y\rangle. \tag{12}$$

Here, (7) holds because $|\zeta_Z\rangle$ and $|\zeta_{Z'}\rangle$ have disjoint support when $Z \cap Z' = \emptyset$; (8) holds by definition of $\mathbf{Z}$; (9) holds by definitions of $\overline{\mathbf{x}}$ and $|\zeta_Z\rangle$; (10) is a simplification; (11) holds by definition of $|R\rangle$, and (12) follows from (6), because x was assumed to be a reordering of y. ∎

Combining Lemma 14 and Lemma 15, we have reduced the problem of lower bounding the number of $\mathcal{O}_\psi$ queries needed to solve b-XHOG, to lower bounding the number of copies of $|\psi\rangle$ needed to solve b-XHOG. The next lemma lower bounds this latter quantity.

Lemma 16. *Let $|\psi\rangle$ be a Haar-random n-qubit quantum state. Consider a quantum algorithm that receives as input $|\psi\rangle^{\otimes k}$ and outputs a string $z \in \{0,1\}^n$. Then:*

$$\underset{|\psi\rangle, z}{\mathbb{E}} \left[|\langle z|\psi\rangle|^2 \right] \leq \frac{2}{2^n} + \frac{O(k^2)}{4^n}.$$

11

Proof. Let $|R\rangle = |\psi\rangle^{\otimes k}$. As we have argued above, the algorithm achieves the same linear XEB score on average if it instead begins with the mixed state σ_R defined in (1), because of the invariance of the Haar measure with respect to phases. By Lemma 15, the algorithm can prepare σ_R by measuring $|R\rangle$ in the computational basis. By a convexity argument, we can assume that the algorithm outputs z deterministically given the measurement results.

Suppose the measurement results are $z_1, z_2, \ldots, z_k$. Clearly, the choice of z that maximizes $\mathbb{E}\left[|\langle z|\psi\rangle|^2\right]$ is whichever z appears most frequently in $z_1, z_2, \ldots, z_k$ (with ties broken arbitrarily): the probabilities $|\langle i|\psi\rangle|^2$ are distributed according to a $\mathrm{Dir}(1, 1, \ldots, 1)$ distribution, so we can easily compute the posterior expectations $\mathbb{E}\left[|\langle i|\psi\rangle|^2 \mid z_1, z_2, \ldots, z_k\right]$. So, it suffices to bound $\mathbb{E}\left[|\langle z|\psi\rangle|^2\right]$ for the algorithm that chooses z to be the most frequent measurement result.

Let m be a random variable that denotes the frequency of the chosen z. Then

$$\mathbb{E}\left[|\langle z|\psi\rangle|^2\right] = \mathbb{E}\left[\mathbb{E}\left[|\langle z|\psi\rangle|^2 \mid m\right]\right] \tag{13}$$

$$= \mathbb{E}\left[\frac{1+m}{2^n + k}\right] \tag{14}$$

$$\leq \frac{1}{2^n} + \mathbb{E}\left[\frac{m}{2^n}\right] \tag{15}$$

$$\leq \frac{1}{2^n} + \mathbb{E}\left[\frac{1 + \sum_{i \neq j} \mathbb{1}[z_i = z_j]}{2^n}\right] \tag{16}$$

$$= \frac{2}{2^n} + \sum_{i \neq j} \frac{\Pr[z_i = z_j]}{2^n} \tag{17}$$

$$= \frac{2}{2^n} + \binom{k}{2} \frac{2}{2^n(2^n + 1)} \tag{18}$$

$$\leq \frac{2}{2^n} + \frac{O(k^2)}{4^n}. \tag{19}$$

Here, (13) is valid by the law of total expectation. (14) substitutes the formula for the posterior expectation of a Dirichlet distribution. (15) is valid by linearity of expectation. In (16), we use the crude upper bound that m is at most one more than the number of pairwise collisions in $z_1, \ldots, z_k$ (which is tight when the number of collisions is 0 or 1). (17) is valid by linearity of expectation. (18) expands the sum, and computes the collision probabilities in terms of moments of the underlying $\mathrm{Dir}(1, 1, \ldots, 1)$ prior distribution. ∎

We note that one should not expect Lemma 16 to be tight for large k (say, $k = \Omega\left(2^{n/2}\right)$). For example, to achieve $b = 4$, we need at least enough samples to see $m \geq 3$ with good probability. But $\Pr[m \geq 3]$ is negligible unless $k = \Omega\left(2^{2n/3}\right)$. More generally, a tight bound on the number of copies of $|\psi\rangle$ needed to achieve a particular value of b seems closely related to the number of measurements of $|\psi\rangle$ needed to see $m \geq b-1$. This is like a sort of "balls into bins" problem [23, 30] with k balls and 2^n bins in which we want to bound the probability that the maximum load of any bin exceeds m, but where the probabilities associated to each bin follow a Dirichlet prior rather than being uniform.

We finally have the tools to prove the main result of this section.

Theorem 17. *Any quantum query algorithm for $(2 + \varepsilon)$-XHOG with query access to $\mathcal{O}_\psi$ for a Haar-random n-qubit state $|\psi\rangle$ requires $\Omega\left(\frac{2^{n/4}\varepsilon^{5/4}}{n}\right)$ queries.*

Proof. Consider a quantum algorithm A that makes T queries to $\mathcal{O}_\psi$ and solves $(2 + \varepsilon)$-XHOG. Choose $k = \frac{c^2 T^2 n^2}{\varepsilon^2}$ in Lemma 14 for a constant c to be chosen later. By Lemma 14, there is a quantum algorithm B that makes no queries to $\mathcal{O}_\psi$ and instead starts with a state $|R\rangle$ (depending on $|\psi\rangle$) such that the trace distance between the output of A and B is at most $O\left(\frac{\varepsilon}{cn}\right)$ for every $|\psi\rangle$. In particular, if we view $|\psi\rangle$ as fixed, then the total variation distance between the outputs z_A and z_B of A and B, respectively, (as probability distributions over $\{0,1\}^n$) is at most $O\left(\frac{\varepsilon}{cn}\right)$. Hence, for every $|\psi\rangle$, we may write:

$$
\begin{aligned}
\mathop{\mathbb{E}}_{z_A}\left[|\langle z_A|\psi\rangle|^2\right] - \mathop{\mathbb{E}}_{z_B}\left[|\langle z_B|\psi\rangle|^2\right] &= \sum_{z \in \{0,1\}^n} |\langle z|\psi\rangle|^2 \cdot (\Pr[z_A = z] - \Pr[z_B = z]) \\
&\leq \sum_{z \in \{0,1\}^n} |\langle z|\psi\rangle|^2 \cdot |\Pr[z_A = z] - \Pr[z_B = z]| \\
&\leq \max_{z \in \{0,1\}^n} |\langle z|\psi\rangle|^2 \cdot \sum_{z' \in \{0,1\}^n} |\Pr[z_A = z'] - \Pr[z_B = z']| \\
&\leq \max_{z \in \{0,1\}^n} |\langle z|\psi\rangle|^2 \cdot O\left(\frac{\varepsilon}{cn}\right),
\end{aligned}
$$

because the sum in the penultimate inequality is twice the total variation distance between z_A and z_B. Fact 13 states that for a Haar-random $|\psi\rangle$, $\mathbb{E}_{|\psi\rangle}\left[\max_{z \in \{0,1\}^n} |\langle z|\psi\rangle|^2\right] \leq \frac{O(n)}{2^n}$. So, for a Haar-random $|\psi\rangle$, we have

$$
\mathop{\mathbb{E}}_{|\psi\rangle, z_A}\left[|\langle z_A|\psi\rangle|^2\right] - \mathop{\mathbb{E}}_{|\psi\rangle, z_B}\left[|\langle z_B|\psi\rangle|^2\right] \leq O\left(\frac{\varepsilon}{c2^n}\right).
$$

In particular, if we choose c sufficiently large, then B solves $(2 + \frac{\varepsilon}{2})$-XHOG.

Because of the invariance of the Haar measure with respect to phases, B still solves $(2 + \frac{\varepsilon}{2})$-XHOG if the pure state $|R\rangle$ is replaced with the mixed state σ_R defined in (1). By Lemma 15, this implies the existence of an algorithm that solves $(2 + \frac{\varepsilon}{2})$-XHOG given k copies of $|\psi\rangle$. By Lemma 16, such an algorithm must satisfy:

$$
\frac{\varepsilon}{2} \leq \frac{O(k^2)}{2^n}.
$$

Plugging in k gives the desired lower bound on T:

$$
\frac{\varepsilon}{2} \leq O\left(\frac{T^4 n^4}{2^n \varepsilon^4}\right)
$$

$$
T \geq \Omega\left(\frac{2^{n/4} \varepsilon^{5/4}}{n}\right). \qquad \blacksquare
$$

Lastly, we give an upper bound on the number of queries needed to nontrivially beat the naive algorithm for XHOG with $\mathcal{O}_\psi$. In fact, the following algorithm works with *any* oracle that prepares a Haar-random state (including a Haar-random unitary), because the algorithm only needs copies of $|\psi\rangle$ and the ability to perform the reflection $\mathcal{R}_\psi$. We thank Scott Aaronson for suggesting this approach based on quantum collision-finding.

Theorem 18. *There is a quantum algorithm for $(2 + \Omega(1))$-XHOG that makes $O\left(2^{n/3}\right)$ queries to a state preparation oracle for a Haar-random n-qubit state $|\psi\rangle$.*

Proof. The quantum algorithm is essentially equivalent to the collision-finding algorithm of Brassard, Høyer, and Tapp [16]. We proceed by measuring $k = 2^{n/3}$ copies of $|\psi\rangle$ in the computational basis, with results $z_1, z_2, \ldots, z_k \in \{0,1\}^n$. If any string appears twice in $z_1, z_2, \ldots, z_k$, we output the first such collision. Otherwise, we perform quantum amplitude amplification [17] on another copy of $|\psi\rangle$, where the "good" subspace is spanned by $z_1, z_2, \ldots, z_k$. This uses the reflection $\mathcal{R}_\psi$, which can be simulated using a constant number of queries to any oracle that prepares $|\psi\rangle$ (see Lemma 5). Finally, we measure and output the result of the amplitude amplification; call this result z_{k+1}. For the purpose of analyzing this algorithm, we say that the algorithm "succeeds" if it either finds a collision in $z_1, z_2, \ldots, z_k$, or if z_{k+1} is contained in the good subspace.

We first argue that for any $|\psi\rangle$, $O\left(2^{n/3}\right)$ queries to $\mathcal{R}_\psi$ (for amplitude amplification) are sufficient for the algorithm to succeed with high probability. For this part of the analysis, we view $|\psi\rangle$ as fixed, and consider only the randomness of the algorithm. Notice that for each $1 \leq i \leq k$, $\Pr\left[|\langle z_i|\psi\rangle|^2 \geq \frac{1}{2^{n+1}}\right] \geq \frac{1}{2}$, because at most half of the probability mass of the output distribution of $|\psi\rangle$ can be placed on inputs for which the output probability is less than $\frac{1}{2^{n+1}}$, because there are only 2^n possible outputs. Thus, by a Chernoff bound, we have that:

$$\Pr\left[\sum_{i=1}^{k} |\langle z_i|\psi\rangle|^2 \geq \frac{k}{2^{n+2}}\right] \geq 1 - \exp(O(k)).$$

In particular, with probability $1 - \exp(O(k))$, either the algorithm finds a collision in $z_1, z_2, \ldots, z_k$, or else $O\left(\sqrt{\frac{2^n}{k}}\right) = O\left(2^{n/3}\right)$ applications of $\mathcal{R}_\psi$ within the amplitude amplification subroutine are sufficient to measure a good string with arbitrarily high constant probability. So overall, the algorithm can be assumed to succeed with arbitrarily high constant probability.

Next, we argue that the algorithm outputs a string z such that $\mathbb{E}\left[|\langle z|\psi\rangle|^2\right] \geq \frac{2+\Omega(1)}{2^n}$. Suppose that instead of performing amplitude amplification at the end, we just measured one additional copy of $|\psi\rangle$ (still calling the result z_{k+1}) and output z_{k+1} if there were no collisions in $z_1, z_2, \ldots, z_k$. Then notice that, conditional on this modified algorithm's success, the expected XEB score is at least $\frac{3-o(1)}{2^n}$. In symbols, we claim that:

$$\mathbb{E}\left[|\langle z|\psi\rangle|^2 \mid \text{success}\right] \geq \frac{3}{2^n + k + 1}$$

for this modified algorithm, because conditional on success, z was observed at least twice in $z_1, z_2, \ldots, z_{k+1}$, so $\frac{3}{2^n+k+1}$ is a lower bound on the posterior expectation of the underlying Dirichlet prior distribution on the output probabilities of $|\psi\rangle$. But now, we claim that $\mathbb{E}\left[|\langle z|\psi\rangle|^2 \mid \text{success}\right]$ is the same for both the modified algorithm and the original algorithm that uses amplitude amplification. The reason is that amplitude amplification preserves conditional probabilities: the conditional probability distribution of z_{k+1} is exactly the same in both algorithms, when conditioned on measuring in the good subspace. So overall, we have that:

$$\mathbb{E}\left[|\langle z|\psi\rangle|^2\right] = \mathbb{E}\left[|\langle z|\psi\rangle|^2 \mid \text{success}\right] \cdot \Pr[\text{success}] + \mathbb{E}\left[|\langle z|\psi\rangle|^2 \mid \text{failure}\right] \cdot \Pr[\text{failure}]$$

$$\geq \mathbb{E}\left[|\langle z|\psi\rangle|^2 \mid \text{success}\right] \cdot \Pr[\text{success}]$$

$$= \frac{3 - o(1)}{2^n} \cdot (1 - p)$$
$$\geq \frac{2 + \Omega(1)}{2^n},$$

where p is the arbitrarily small constant failure probability of amplitude amplification. ∎

We remark that a sharper analysis could most likely improve the above algorithm from solving $(2+\Omega(1))$-XHOG to solving $(3-o(1))$-XHOG, while still using the same number of queries. For most Haar-random states $|\psi\rangle$, the probability of measuring in the "good" subspace should concentrate very well. As a result, it should be possible to fix some $T(n)$ such that running exactly $T(n)$ iterations of Grover's algorithm ensures finding a marked with high probability, rather than constant probability.

4   Random State Preparation Oracles

In this section, we show that a canonical state preparation oracle and a random state preparation oracle are essentially equivalent, and use it to prove the quantum supremacy Tsirelson inequality for XHOG with a Haar-random oracle (Problem 10).

By Lemma 7, for a state $|\psi\rangle$, query access to a random state preparation oracle U_ψ implies query access to the canonical state preparation oracle $\mathcal{O}_\psi$ with constant overhead. The reverse direction is less obvious. We know from the definition of U_ψ (Definition 8) that one can simulate U_ψ given *any* n-qubit unitary V that prepares $|\psi\rangle$ from $|0^n\rangle$. So, it is tempting to let $V = \mathcal{O}_\psi$ with $|\bot\rangle = |0^n\rangle$ to argue that $\mathcal{O}_\psi$ allows simulating U_ψ. However, this is only possible if $|0^n\rangle$ is orthogonal to $|\psi\rangle$. And while we previously argued that we can always find a canonical state $|\bot\rangle$ that is orthogonal to $|\psi\rangle$ (Footnote 4), this requires extending the Hilbert space, so that $\mathcal{O}_\psi$ no longer acts on n qubits!

To address this, imagine that we knew an explicit n-qubit state $|\psi^\bot\rangle$ orthogonal to $|\psi\rangle$. Notice that we could perfectly swap $|\psi\rangle$ and $|\psi^\bot\rangle$: the composition $\mathcal{O}_\psi \mathcal{O}_{\psi^\bot} \mathcal{O}_\psi$ sends $|\psi\rangle$ to $|\psi^\bot\rangle$, $|\psi^\bot\rangle$ to $|\psi\rangle$, and acts trivially on all states orthogonal to $|\psi\rangle$ and $|\psi^\bot\rangle$. In particular, this swaps $|\psi\rangle$ and $|\psi^\bot\rangle$ while acting only on the space of n-qubit states. Next, if we know $|\psi^\bot\rangle$ explicitly, we can certainly come up with an n-qubit unitary that sends $|0^n\rangle$ to $|\psi^\bot\rangle$. By composing such a unitary with $\mathcal{O}_\psi \mathcal{O}_{\psi^\bot} \mathcal{O}_\psi$, we are left with an n-qubit unitary that sends $|0^n\rangle$ to $|\psi\rangle$. This is sufficient to construct U_ψ, by Definition 8.

While we do not necessarily have such a state $|\psi^\bot\rangle$, a *random* n-qubit state $|\varphi\rangle$ will be exponentially close to such a $|\psi^\bot\rangle$ with overwhelming probability. The next theorem shows that we can use this observation to *approximately* simulate U_ψ given $\mathcal{O}_\psi$, by going through the steps above and keeping track of deviation from the ideal construction in terms of $\langle\psi|\varphi\rangle$.

Theorem 19. *Let $|\psi\rangle$ be an n-qubit state. Consider a quantum query algorithm A that makes T queries to U_ψ. Then there is a quantum query algorithm B that makes $2T$ queries to $\mathcal{O}_\psi$ such that:*

$$\left\| \mathbb{E}_{U_\psi} [A] - B \right\|_\diamond \leq \frac{10T + 4}{2^{n/2}}.$$

Proof. Without loss of generality, assume $|\bot\rangle$ is orthogonal to all n-qubit states. Let $|\varphi\rangle$ be a Haar-random n-qubit state, and let V be an arbitrary n-qubit unitary that satisfies $V|0^n\rangle = |\varphi\rangle$. Write $|\varphi\rangle = \alpha|\psi^\bot\rangle + \beta|\psi\rangle$, where $|\psi^\bot\rangle$ is some n-qubit state orthogonal to $|\psi\rangle$, with the phase chosen so that α is real and nonnegative. Note that $\beta = \langle\psi|\varphi\rangle$.

Suppose we had an oracle V' acting on n qubits such that $V'|0^n\rangle = |\psi^\perp\rangle$. Then we could appeal to Lemma 7 to simulate an oracle $\mathcal{O}_{\psi^\perp}$ that reflects about the state $\frac{|\psi^\perp\rangle - |\perp\rangle}{\sqrt{2}}$ using queries to V'. Then the composition $\mathcal{O}_\psi \mathcal{O}_{\psi^\perp} \mathcal{O}_\psi$ would swap $|\psi\rangle$ and $|\psi^\perp\rangle$, while acting only on the space of n-qubit states. Furthermore, we would have that $\mathcal{O}_\psi \mathcal{O}_{\psi^\perp} \mathcal{O}_\psi V'|0^n\rangle = |\psi\rangle$, where $\mathcal{O}_\psi \mathcal{O}_{\psi^\perp} \mathcal{O}_\psi V'$ acts on n qubits. So, by Definition 8, we could simulate U_ψ perfectly by choosing a random $(2^n - 1)$-dimensional unitary W and replacing calls to U_ψ with $\mathcal{O}_\psi \mathcal{O}_{\psi^\perp} \mathcal{O}_\psi V'W$.

Unfortunately, we do not have such an oracle V'; we only have V. However, we can show that there exists an oracle V' that is *close* to V, so if we replace all occurrences of V' with V, the resulting unitary we get is close to a random state preparation oracle for $|\psi\rangle$. Specifically, we take R to be a rotation in the 2-dimensional space spanned by $|\psi\rangle$ and $|\psi^\perp\rangle$ that satisfies $R|\varphi\rangle = |\psi^\perp\rangle$. Then, we let $V' = RV$.

R is a rotation by angle $\theta = \arccos(\alpha)$ in this 2-dimensional subspace, and acts as the identity elsewhere. So, R has eigenvalues $e^{i\theta}$, $e^{-i\theta}$, and 1. The assumption that $\alpha \geq 0$ implies $\theta \leq \frac{\pi}{2}$, so by Fact 12,

$$\left\| V \cdot V^\dagger - V' \cdot V'^\dagger \right\|_\diamond = 2\sqrt{1 - \cos^2(\theta)} = 2\sin\theta = 2|\langle\psi|\varphi\rangle|. \tag{20}$$

Lemma 7 shows that V' (or more precisely, controlled-V' or its inverse) is used $4T + 2$ times in implementing T queries to $\mathcal{O}_{\psi^\perp}$, which means we need $5T + 2$ applications of V' to implement T queries to $\mathcal{O}_\psi \mathcal{O}_{\psi^\perp} \mathcal{O}_\psi V'$.

Let $B_{\psi^\perp}$ denote the quantum algorithm that simulates A using $\mathcal{O}_\psi \mathcal{O}_{\psi^\perp} \mathcal{O}_\psi V'W$ (for a random choice of W) in place of U_ψ, and let B_φ denote the quantum algorithm that simulates $B_{\psi^\perp}$ using V in place of V'. Then

$$\left\| \mathbb{E}_{U_\psi}[A] - B_\varphi \right\|_\diamond = \left\| B_{\psi^\perp} - B_\varphi \right\|_\diamond \tag{21}$$

$$\leq (5T + 2)\left\| V \cdot V^\dagger - V' \cdot V'^\dagger \right\|_\diamond \tag{22}$$

$$= (10T + 4)|\langle\psi|\varphi\rangle|, \tag{23}$$

where (21) holds because $\mathbb{E}_{U_\psi}[A]$ and $B_{\psi^\perp}$ are equivalent as superoperators; (22) holds by the subadditivity of the diamond norm under composition, because $B_{\psi^\perp}$ queries V' a total of $5T + 2$ times; and (23) substitutes (20).

Finally, let $B = \mathbb{E}_{|\varphi\rangle}[B_\varphi]$ (i.e. run B_φ for a Haar-random choice of $|\varphi\rangle$). Then

$$\left\| \mathbb{E}_{U_\psi}[A] - B \right\|_\diamond = \left\| \mathbb{E}_{U_\psi}[A] - \mathbb{E}_{|\varphi\rangle}[B_\varphi] \right\|_\diamond \tag{24}$$

$$\leq \mathbb{E}_{|\varphi\rangle}\left[\left\| \mathbb{E}_{U_\psi}[A] - B_\varphi \right\|_\diamond \right] \tag{25}$$

$$\leq \mathbb{E}_{|\varphi\rangle}\left[(10T + 4)|\langle\psi|\varphi\rangle| \right] \tag{26}$$

$$= (10T + 4)\mathbb{E}_{|\varphi\rangle}\left[\frac{\sum_{i=1}^{2^n} |\langle i|\varphi\rangle|}{2^n} \right] \tag{27}$$

$$\leq \frac{10T + 4}{2^n} \max_{|\varphi\rangle}\left[\sum_{i=1}^{2^n} |\langle i|\varphi\rangle| \right] \tag{28}$$

16

$$\leq \frac{10T + 4}{2^{n/2}}, \tag{29}$$

where (24) holds by the definition of B; (25) holds by Jensen's inequality because the diamond norm, like every norm, is convex; (26) substitutes (23); (27) holds by symmetry (the choice of orthonormal basis $\{|i\rangle : i \in [2^n]\}$ is arbitrary); (28) trivially upper bounds (27); and (29) holds because the 1-norm of an n-qubit quantum state is at most $2^{n/2}$ (maximized by a uniform superposition). ∎

The above theorem implies that the oracle $\mathcal{O}_\psi$ in Theorem 17 can be replaced by a Haar-random n-qubit unitary.

Theorem 20. *Any quantum query algorithm for $(2 + \varepsilon)$-XHOG with query access to U_ψ for a Haar-random n-qubit state $|\psi\rangle$ (i.e. a Haar-random n-qubit unitary) requires $\Omega\left(\frac{2^{n/4}\varepsilon^{5/4}}{n}\right)$ queries.*

Proof. Consider a quantum algorithm A that makes T queries to U_ψ and solves $(2 + \varepsilon)$-XHOG. Let c be a constant to be chosen later. If $T > c\frac{2^{n/2}\varepsilon}{n}$, then we are done, because we can always assume that $\varepsilon \leq O(n)$ (Fact 13), so $\frac{2^{n/2}\varepsilon}{n} \geq \frac{2^{n/4}\varepsilon^{5/4}}{n}$ for sufficiently large n. In the complementary case, suppose that $T \leq c\frac{2^{n/2}\varepsilon}{n}$. By Theorem 19 and the definition of the diamond norm, there is a quantum query algorithm B that makes $2T$ queries to $\mathcal{O}_\psi$ such that the trace distance between the output of A (averaged over the choice of U_ψ) and B is at most $\frac{10T+4}{2^{n/2}} \leq \frac{14T}{2^{n/2}} \leq \frac{14c\varepsilon}{n}$ for every $|\psi\rangle$. By an argument involving Fact 13 similar to the one used in the proof of Theorem 17, we conclude that if c is a sufficiently small constant, then B solves $(2 + \frac{\varepsilon}{2})$-XHOG (with a canonical state preparation oracle for a Haar-random state). By Theorem 17, this implies $T = \Omega\left(\frac{2^{n/4}\varepsilon^{5/4}}{n}\right)$. ∎

5 Fourier Sampling Circuits

In this section, we prove the quantum supremacy Tsirelson inequality for single-query algorithms over FOURIER SAMPLING circuits (Problem 11).

Throughout this section, we let $N = 2^n$, and let $\mathcal{F}_n := \{f : \{0,1\}^n \to \{-1,1\}\}$ denote the set of all n-bit Boolean functions. Given a function $f \in \mathcal{F}_n$, we define the Fourier coefficient

$$\hat{f}(z) := \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x)(-1)^{x \cdot z}$$

for each $z \in \{0,1\}^n$. We also define the characters $\chi_z : \{0,1\}^n \to \{-1,1\}$ for each $z \in \{0,1\}^n$:

$$\chi_z(x) := (-1)^{x \cdot z}.$$

Given oracle access to a function $f \in \mathcal{F}_n$, the FOURIER SAMPLING quantum circuit for f consists of a layer of Hadamard gates, then a single query to f, then another layer of Hadamard gates, so that the resulting circuit samples a string $z \in \{0,1\}^n$ with probability $\hat{f}(z)^2$. In the context of XHOG, we consider the distribution of FOURIER SAMPLING circuits where the oracle f is chosen uniformly at random from $\mathcal{F}_n$.

Proposition 21. *FOURIER SAMPLING circuits over n qubits solve $(3 - \frac{2}{2^n})$-XHOG.*

Proof. Because the circuit samples z with probability $\hat{f}(z)^2$, the expected linear XEB score is:

$$\mathbb{E}_{f \in \mathcal{F}_n} \left[\sum_{z \in \{0,1\}^n} \hat{f}(z)^4 \right] = \mathbb{E}_{f \in \mathcal{F}_n} \left[\sum_{z \in \{0,1\}^n} \widehat{f \cdot \chi_z}(0^n)^4 \right] \tag{30}$$

$$= 2^n \mathbb{E}_{f \in \mathcal{F}_n} \left[\hat{f}(0^n)^4 \right] \tag{31}$$

$$= 2^n \left(\frac{2}{2^n} \right)^4 \mathbb{E} \left[(B(2^n, 1/2) - \mathbb{E}\left[B(2^n, 1/2) \right])^4 \right] \tag{32}$$

$$= \frac{3 - \frac{2}{2^n}}{2^n}, \tag{33}$$

where (30) applies the substitution $\hat{f}(z) = \widehat{f \cdot \chi_z}(0^n)$; (31) is valid because if f is uniform over $\mathcal{F}_n$ then so is $f \cdot \chi_z$; (32) uses the fact that $\frac{2^n}{2}(\hat{f}(0^n) + 1)$ is binomially distributed with 2^n trials and success probability $\frac{1}{2}$; and (33) uses the formula $Np(1-p)(1 + (3N-6)p(1-p))$ for the 4th central moment of a binomial distribution with N trials and success probability p. ■

The remainder of this section constitutes the proof of the following theorem, which shows the optimality of the 1-query algoritm for XHOG with FOURIER SAMPLING circuits:

Theorem 22. *Any 1-query algorithm for b-XHOG over n-qubit FOURIER SAMPLING circuits satisfies $b \leq 3 - \frac{2}{2^n}$.*

To prove Theorem 22, we use the polynomial method of Beals et al. [11]. Consider a quantum query algorithm that makes T queries to $f \in \mathcal{F}_n$ and outputs a string $z \in \{0,1\}^n$. The polynomial method implies that for each $z \in \{0,1\}^n$, the probability that the algorithm outputs z can be expressed as a real multilinear polynomial of degree $2T$ in the bits of f. We write such a polynomial as:

$$p_z(f) = \sum_{S \subset \{0,1\}^n, |S| \leq 2T} c_{z,S} \cdot \prod_{x \in S} f(x).$$

Then, the expected XEB score of this quantum query algorithm is given by:

$$\frac{1}{2^N} \sum_{f \in \mathcal{F}_n} \sum_{z \in \{0,1\}^n} p_z(f) \cdot \hat{f}(z)^2. \tag{34}$$

Our key observation is that the quantity (34) is linear in the coefficients $c_{z,S}$. This allows us to express the largest XEB score achievable by polynomials of degree $2T$ as a linear program, with the constraints that the polynomials $\{p_z(f) : z \in \{0,1\}^n\}$ must represent a probability distribution. Then, the objective value of the linear program can be upper bounded by giving a solution to the dual linear program. We can write the linear program as follows:

$$\begin{array}{lll}
\max & \dfrac{1}{2^N} \displaystyle\sum_{f \in \mathcal{F}_n} \sum_{z \in \{0,1\}^n} p_z(f) \cdot \hat{f}(z)^2 & \\
\text{subject to} & p_z(f) \geq 0 & \text{for each } f \in \mathcal{F}_n; z \in \{0,1\}^n \\
& \displaystyle\sum_{z \in \{0,1\}^n} p_z(f) = 1 & \text{for each } f \in \mathcal{F}_n \\
& c_{z,S} \in \mathbb{R} & \text{for each } z \in \{0,1\}^n; 0 \leq |S| \leq 2T
\end{array} \tag{35}$$

Before giving a solution to (or even writing down) the dual linear program, we will first show that the primal linear program can be simplified considerably.

We first argue that one can apply a sort of symmetrization to reduce the number of variables. Consider a solution to the linear program (35) in terms of polynomials p_z, and define:

$$p_z'(f) = \frac{1}{N} \sum_{y \in \{0,1\}^n} p_{y \oplus z}(f \cdot \chi_y).$$

Then we claim that the polynomials p_z' are also a solution to the linear program with the same objective value. The intuition is that $\hat{f}(z) = \widehat{f \cdot \chi_y}(y \oplus z)$, so we might as well assume that the probability of outputting z on f is the same as the probability of outputting $y \oplus z$ on $f \cdot \chi_y$, by averaging over the possible choices of y. We verify that the objective value is:

$$
\begin{aligned}
\frac{1}{2^N} \sum_{f \in \mathcal{F}_n} \sum_{z \in \{0,1\}^n} p_z'(f) \cdot \hat{f}(z)^2 &= \frac{1}{N 2^N} \sum_{f \in \mathcal{F}_n} \sum_{z \in \{0,1\}^n} \sum_{y \in \{0,1\}^n} p_{y \oplus z}(f \cdot \chi_y) \cdot \hat{f}(z)^2 \\
&= \frac{1}{N 2^N} \sum_{f \in \mathcal{F}_n} \sum_{z \in \{0,1\}^n} \sum_{y \in \{0,1\}^n} p_{y \oplus z}(f \cdot \chi_y) \cdot \widehat{f \cdot \chi_y}(y \oplus z)^2 \\
&= \frac{1}{N 2^N} \sum_{y \in \{0,1\}^n} \sum_{f \in \mathcal{F}_n} \sum_{z \in \{0,1\}^n} p_{y \oplus z}(f \cdot \chi_y) \cdot \widehat{f \cdot \chi_y}(y \oplus z)^2 \\
&= \frac{1}{2^N} \sum_{f \in \mathcal{F}_n} \sum_{z \in \{0,1\}^n} p_z(f) \cdot \hat{f}(z)^2.
\end{aligned}
$$

The nonnegativity constraint on each $p_z'(f)$ is satisfied by convexity, and the polynomials sum to 1 for each f because

$$
\begin{aligned}
\sum_{z \in \{0,1\}^n} p_z'(f) &= \frac{1}{N} \sum_{z \in \{0,1\}^n} \sum_{y \in \{0,1\}^n} p_{y \oplus z}(f \cdot \chi_y) \\
&= \frac{1}{N} \sum_{y \in \{0,1\}^n} \sum_{z \in \{0,1\}^n} p_{y \oplus z}(f \cdot \chi_y) \\
&= \frac{1}{N} \sum_{y \in \{0,1\}^n} \sum_{z \in \{0,1\}^n} p_z(f \cdot \chi_y) \\
&= \frac{1}{N} \sum_{y \in \{0,1\}^n} 1 \\
&= 1.
\end{aligned}
$$

Notice that the p_z's satisfy $p_z'(f) = p_{0^n}'(f \cdot \chi_z)$. So, we can rewrite the linear program in terms of $p_{0^n}'(f)$ alone. Define $p(f) = p_{0^n}'(f)$ and define the coefficients of $p(f)$ by:

$$p(f) = \sum_{S \subset \{0,1\}^n, |S| \leq 2T} c_S \cdot \prod_{x \in S} f(x).$$

19

Now, we can rewrite the linear program (35) in terms of $p(f)$ as:

$$
\begin{aligned}
\max \quad & \frac{1}{2^N} \sum_{f \in \mathcal{F}_n} \sum_{z \in \{0,1\}^n} p(f \cdot \chi_z) \cdot \hat{f}(z)^2 \\
\text{subject to} \quad & p(f) \geq 0 & \text{for each } f \in \mathcal{F}_n \\
& \sum_{z \in \{0,1\}^n} p(f \cdot \chi_z) = 1 & \text{for each } f \in \mathcal{F}_n \\
& c_S \in \mathbb{R} & \text{for each } 0 \leq |S| \leq 2T
\end{aligned}
\tag{36}
$$

We can simplify the objective function of the linear program (36) even further:

$$
\begin{aligned}
\frac{1}{2^N} \sum_{f \in \mathcal{F}_n} \sum_{z \in \{0,1\}^n} p(f \cdot \chi_z) \cdot \hat{f}(z)^2 &= \frac{1}{2^N} \sum_{f \in \mathcal{F}_n} \sum_{z \in \{0,1\}^n} p(f \cdot \chi_z) \cdot \widehat{f \cdot \chi_z}(0^n)^2 \\
&= \frac{1}{2^N} \sum_{z \in \{0,1\}^n} \sum_{f \in \mathcal{F}_n} p(f \cdot \chi_z) \cdot \widehat{f \cdot \chi_z}(0^n)^2 \\
&= \frac{1}{2^N} \sum_{z \in \{0,1\}^n} \sum_{f \in \mathcal{F}_n} p(f) \cdot \hat{f}(0^n)^2 \\
&= \frac{N}{2^N} \sum_{f \in \mathcal{F}_n} p(f) \cdot \hat{f}(0^n)^2.
\end{aligned}
$$

Notice that we can also assume $p(f) = p(-f)$ without loss of generality, because the squared Fourier coefficient of f is the same as the squared Fourier coefficient of its negation, and because replacing $p(f)$ by $\frac{p(f) + p(-f)}{2}$ still satisfies all of the constraints. In particular, we can assume $c_S = 0$ if $|S|$ is odd.

Next, we turn to simplifying the equality constraint. Define $q(f)$ by:

$$
\begin{aligned}
q(f) &:= \sum_{z \in \{0,1\}^n} p(f \cdot \chi_z) \\
&= \sum_{z \in \{0,1\}^n} \sum_{|S| \leq 2T} c_S \cdot \prod_{x \in S} f(x) \cdot (-1)^{x \cdot z},
\end{aligned}
$$

which is also a multilinear polynomial in f of degree $2T$. Then the equality constraint reads as $q(f) = 1$ for every $f \in \mathcal{F}_n$. Because q is multilinear, this implies $q(f) = 1$ in fact holds identically over all $f : \{0,1\}^n \to \mathbb{R}$, and not just Boolean-valued f. So, the coefficient on the monomial of the set S in q must be 1 if $S = \emptyset$ and 0 otherwise. For S empty, we have:

$$
\sum_{z \in \{0,1\}^n} c_\emptyset = 1,
$$

which is to say that $c_\emptyset = \frac{1}{N}$. Otherwise, for nonempty S we have:

$$
\sum_{z \in \{0,1\}^n} c_S \prod_{x \in S} (-1)^{x \cdot z} = 0.
$$

We can rewrite the equation above as:

$$
\sum_{z \in \{0,1\}^n} c_S (-1)^{\left(\sum_{x \in S} x \right) \cdot z} = 0.
$$

Now, there are two cases:

20

- If $\bigoplus_{x \in S} x = 0^n$, then the equation holds if and only if $c_S = 0$.

- If $\bigoplus_{x \in S} x \neq 0^n$, then the terms in the sum are c_S half of the time and $-c_S$ the other half of the time, so equality always holds.

Putting this altogether, we can conclude:

- $c_\emptyset = \frac{1}{N}$.

- $c_S = 0$ if $\bigoplus_{x \in S} x = 0^n$.

- $c_S = 0$ if $|S|$ is odd.

- c_S is otherwise unconstrained by $q(f) = 1$.

After all of this, our linear program now has the much simpler form:[5]

$$
\boxed{
\begin{array}{lll}
\max & \dfrac{N}{2^N} \displaystyle\sum_{f \in \mathcal{F}_n} p(f) \cdot \hat{f}(0^n)^2 & \\[2ex]
\text{subject to} & p(f) \geq 0 & \text{for each } f \in \mathcal{F}_n \\[1ex]
& c_\emptyset = \dfrac{1}{N} & \\[1ex]
& c_S \in \mathbb{R} & \text{for each } 2 \leq |S| \leq 2T \text{ with } |S| \text{ even, } \bigoplus S \neq 0^n
\end{array}
}
\tag{37}
$$

Alternatively, we can express the linear program (37) purely in terms of the variables c_S, rather than leaving them implicit in $p(f)$. In the objective function, the coefficient on c_S is given by:

$$
k_S := \frac{N}{2^N} \sum_{f \in \mathcal{F}_n} \hat{f}(0^n)^2 \cdot \prod_{x \in S} f(x).
$$

We compute k_S depending on the size of S:

- If $S = \emptyset$, then $k_S = N \cdot \mathbb{E}\left[\hat{f}(0^n)^2\right] = N \cdot \mathbb{E}\left[\hat{f}(0^n)^2 - \mathbb{E}\left[\hat{f}(0^n)\right]^2\right] = N \cdot \mathrm{Var}\left[\hat{f}(0^n)\right] = 1$, because for a random f, $\hat{f}(0^n)$ is a sum of 2^n independent $\pm \frac{1}{2^n}$ variables.

- If $S \neq \emptyset$, then:

$$
\begin{aligned}
k_S &= \frac{N}{2^N} \sum_{f \in \mathcal{F}_n} \frac{1}{2^{2n}} \sum_{x_1 \in \{0,1\}^n} \sum_{x_2 \in \{0,1\}^n} f(x_1) f(x_2) \prod_{x \in S} f(x) \\
&= \frac{N}{2^N} \frac{1}{2^{2n}} \sum_{x_1 \in \{0,1\}^n} \sum_{x_2 \in \{0,1\}^n} \sum_{f \in \mathcal{F}_n} f(x_1) f(x_2) \prod_{x \in S} f(x) \\
&= \begin{cases} \frac{2}{N} & |S| = 2 \\ 0 & |S| > 2 \end{cases},
\end{aligned}
$$

because in the second line, the innermost sum is 0 unless $\{x_1, x_2\} = S$.

---

[5]Here, $\bigoplus S$ is shorthand for $\bigoplus_{x \in S} x$.

So, the final primal linear program takes the form:

$$
\begin{aligned}
\text{max} \quad & c_\emptyset + \frac{2}{N} \sum_{|S|=2} c_S \\
\text{subject to} \quad & \sum_S c_S \cdot \prod_{x \in S} f(x) \geq 0 \quad \text{for each } f \in \mathcal{F}_n \\
& c_\emptyset = \frac{1}{N} \\
& c_S \in \mathbb{R} \quad\quad\quad\quad\quad\quad \text{for each } 2 \leq |S| \leq 2T \text{ with } |S| \text{ even, } \bigoplus S \neq 0^n
\end{aligned}
\tag{38}
$$

Standard manipulations reveal the dual linear program of (38):

$$
\begin{aligned}
\text{min} \quad & \frac{b}{N} \\
\text{subject to} \quad & b - \sum_{f \in \mathcal{F}_n} \psi_f = 1 \\
& - \sum_{f \in \mathcal{F}_n} \psi_f \prod_{x \in S} f(x) = \frac{2}{N} \quad \text{for each } |S| = 2 \\
& - \sum_{f \in \mathcal{F}_n} \psi_f \prod_{x \in S} f(x) = 0 \quad \text{for each } 4 \leq |S| \leq 2T \text{ with } |S| \text{ even, } \bigoplus S \neq 0^n \\
& \psi_f \geq 0 \quad\quad\quad\quad\quad\quad \text{for each } f \in \mathcal{F}_n \\
& b \in \mathbb{R}
\end{aligned}
\tag{39}
$$

We now construct a solution to the dual linear program (39) for $T = 1$ query. Our dual solution is motivated by complementary slackness, which guarantees that a variable in (39) of the optimal dual solution is nonzero if and only if the corresponding constraint in (38) is tight in the optimal primal solution. The naive XHOG algorithm solves the primal linear program with $p(f) = \hat{f}(0^n)^2$, so $p(f) = 0$ if and only if $\hat{f}(0^n) = 0$. Thus, if we think that the naive algorithm is optimal, then we should look for a dual solution where ψ_f is nonzero if and only if $\hat{f}(0^n) = 0$.

For some κ to be chosen later, we choose $\psi_f = \kappa$ if $\hat{f}(0^n) = 0$ and $\psi_f = 0$ otherwise. In other words, we let $\psi_f = \kappa \cdot \text{Half}_N(f)$, where $\text{Half}_N : \{-1, 1\}^N \to \{0, 1\}$ is the 0-1 indicator of the set of functions in $\mathcal{F}_n$ (viewed as N-bit strings) with exactly $\frac{N}{2}$ coordinates equal to -1.

Viewing $\psi_f = \psi(f)$ as a function $\psi : \{-1, 1\}^N \to \mathbb{R}$, it will be convenient to rewrite the constraints of the dual linear program in terms of the Fourier coefficients of ψ. We use the inner product formula below for Fourier coefficients, with the understanding that we identify a set $S \subseteq [N]$ with its characteristic string in $\{0, 1\}^N$:

$$
\hat{\psi}(S) = \frac{1}{2^N} \sum_{f \in \{-1,1\}^N} \psi_f \prod_{x \in S} f(x).
$$

Now, the dual linear program reads as:

$$
\begin{aligned}
&\min && \frac{b}{N} \\
&\text{subject to} && b - 2^N \hat{\psi}(\emptyset) = 1 \\
&&& 2^N \hat{\psi}(S) = -\frac{2}{N} && \text{for each } |S| = 2 \\
&&& \hat{\psi}(S) = 0 && \text{for each } 4 \le |S| \le 2T \text{ with } |S| \text{ even, } \bigoplus S \neq 0^n \\
&&& \psi_f \ge 0 && \text{for each } f \in \mathcal{F}_n \\
&&& b \in \mathbb{R}
\end{aligned}
\tag{40}
$$

The Fourier coefficients of Half_N are well known [29, Theorem 5.19], though they are also easy to compute by hand for sets of small size. For $|S| = 2j$, they are given by:

$$
\widehat{\mathrm{Half}_N}(S) = (-1)^j \frac{\binom{N/2}{j}}{\binom{N}{2j}} \cdot \frac{\binom{N}{N/2}}{2^N}.
$$

The $|S| = 2$ equality constraint of the dual linear program (40) implies

$$
2^N \cdot \kappa \cdot \widehat{\mathrm{Half}_N}(S) = -\frac{2}{N}
$$

$$
\kappa = \frac{1}{2^N} \cdot \frac{2}{N} \cdot \frac{\binom{N}{2}}{\binom{N/2}{1}} \cdot \frac{2^N}{\binom{N}{N/2}} = \frac{4\binom{N}{2}}{N^2 \binom{N}{N/2}} = \frac{2(N-1)}{N\binom{N}{N/2}}.
$$

Plugging this value of κ into the constraint on $\hat{\psi}(\emptyset)$ gives:

$$
b - 2^N \cdot \kappa \cdot \widehat{\mathrm{Half}_N}(\emptyset) = 1
$$

$$
b = 1 + 2^N \cdot \kappa \cdot \frac{\binom{N/2}{0}}{\binom{N}{0}} \cdot \frac{\binom{N}{N/2}}{2^N} = 1 + 2\frac{N-1}{N} = 3 - \frac{2}{N}.
$$

This completes the proof, as we have shown a solution to the dual linear program with objective value $\frac{3 - \frac{2}{N}}{N}$.

6   Discussion

The most natural question left for future work is whether our bounds could be improved. Our lower bounds for b-XHOG with $\mathcal{O}_\psi$ or U_ψ show that for constant ε, $(2 + \varepsilon)$-XHOG requires $\Omega\left(\frac{2^{n/4}}{\mathrm{poly}(n)}\right)$ queries to either oracle, while the best upper bound we know of solves $(2 + \varepsilon)$-XHOG in $O\left(2^{n/3}\right)$ queries. We conjecture that this upper bound is tight.

One possible approach towards improving the lower bound for b-XHOG with $\mathcal{O}_\psi$ (and by extension, U_ψ) is to use the polynomial method, as we did for the FOURIER SAMPLING lower bound. Indeed, the output probabilities of an algorithm that makes T queries to $\mathcal{O}_\psi$ can be expressed as degree-$2T$ polynomials in the entries of $\mathcal{O}_\psi$. If we write $|\psi\rangle = \sum_{i=1}^N \alpha_i |i\rangle$, then these are polynomials in the amplitudes $\alpha_1, \ldots, \alpha_N$ and the conjugates of the amplitudes $\alpha_1^*, \ldots, \alpha_N^*$. Because

of the invariance of the Haar measure with respect to phases, and because the linear XEB score depends only on the magnitudes of the amplitudes, we can further assume without loss of generality that the output probabilities are polynomials in the variables $|\alpha_1|^2, \ldots, |\alpha_N|^2$, which are equivalently the measurement probabilities of $|\psi\rangle$ in the computational basis. We can also assume that these polynomials are homogeneous, because the input variables satisfy $\sum_{i=1}^N |\alpha_i|^2 = 1$, so we can multiply any lower-degree terms by this sum to make all terms have the same degree. Like in our FOURIER SAMPLING lower bound, the polynomials are constrained to represent a probability distribution for all valid inputs. However, unlike the FOURIER SAMPLING lower bound, this introduces uncountably many constraints in the primal linear program: the polynomials representing the output probabilities must be nonnegative for *all* inputs $(|\alpha_1|^2, \ldots, |\alpha_N|^2)$ on the probability simplex. It may still be possible to exhibit a solution to the dual linear program if only finitely many of the constraints are relevant (such an approach was used in [18], for example). Put another way, it might suffice to throw out all but finitely many of the primal constraints to obtain a nontrivial upper bound on the value of the linear program.

Our b-XHOG bound for FOURIER SAMPLING circuits is tight, but it only applies to single-query algorithms. In principle, our lower bound approach via the polynomial method could be generalized to algorithms that make additional queries, by increasing the degree of the polynomials in the linear program (38) and exhibiting another dual solution. The challenge seems to be that the parity constraint on the monomials with nonzero coefficients becomes unwieldy when working with polynomials of larger degree.

Beyond possible improvements to the query complexity bounds, it would be interesting to give some evidence that beating the naive XHOG algorithm is hard in the real world. Aaronson and Gunn [3] showed that $(1 + \varepsilon)$-XHOG is *classically* hard, assuming the classical hardness of nontrivially estimating the output probabilities of random quantum circuits. It is not clear whether a similar argument could work for quantum algorithms, though, because sampling from a random quantum circuit gives a better-than-trivial algorithm for estimating its output probabilities.

Acknowledgements

References

[1] Scott Aaronson. Random circuit sampling: Thoughts and open problems. The Quantum Wave in Computing, 2020. URL https://simons.berkeley.edu/talks/tbd-124. 2

[2] Scott Aaronson and Lijie Chen. Complexity-Theoretic Foundations of Quantum Supremacy Experiments. In Ryan O'Donnell, editor, *32nd Computational Complexity Conference (CCC 2017)*, volume 79 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 22:1–22:67, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. ISBN 978-3-95977-040-8. DOI: 10.4230/LIPIcs.CCC.2017.22. URL http://drops.dagstuhl.de/opus/volltexte/2017/7527. 3, 4

[3] Scott Aaronson and Sam Gunn. On the classical hardness of spoofing linear cross-entropy benchmarking. *Theory of Computing*, 16(11):1–8, 2020. DOI: 10.4086/toc.2020.v016a011. URL http://www.theoryofcomputing.org/articles/v016a011. 2, 24

[4] 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. ISBN 978-3-95977-156-6. DOI: 10.4230/LIPIcs.CCC.2020.7. URL https://drops.dagstuhl.de/opus/volltexte/2020/12559. 4, 6

[5] Dorit Aharonov, Alexei Kitaev, and Noam Nisan. Quantum circuits with mixed states. In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC '98, page 20–30, New York, NY, USA, 1998. Association for Computing Machinery. ISBN 0897919629. DOI: 10.1145/276698.276708. URL https://doi.org/10.1145/276698.276708. 6, 8

[6] Andris Ambainis. Understanding quantum algorithms via query complexity. In *Proceedings of the 2018 International Congress of Mathematicians*, volume 3, pages 3249–3270, 2018. DOI: 10.1142/9789813272880_0181. 3

[7] Andris Ambainis, Loïck Magnin, Martin Roetteler, and Jeremie Roland. Symmetry-assisted adversaries for quantum state generation. In *Proceedings of the 2011 IEEE 26th Annual Conference on Computational Complexity*, CCC '11, page 167–177, USA, 2011. IEEE Computer Society. ISBN 9780769544113. DOI: 10.1109/CCC.2011.24. URL https://doi.org/10.1109/CCC.2011.24. 3

[8] Andris Ambainis, Ansis Rosmanis, and Dominique Unruh. Quantum attacks on classical proof systems: The hardness of quantum rewinding. In *Proceedings of the 2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, FOCS '14, page 474–483, USA, 2014. IEEE Computer Society. ISBN 9781479965175. DOI: 10.1109/FOCS.2014.57. URL https://doi.org/10.1109/FOCS.2014.57. 4, 5, 6, 9

[9] Srinivasan Arunachalam, Aleksandrs Belovs, Andrew M. Childs, Robin Kothari, Ansis Rosmanis, and Ronald de Wolf. Quantum Coupon Collector. In Steven T. Flammia, editor, *15th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2020)*, volume 158 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 10:1–10:17, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. ISBN 978-3-95977-146-7. DOI: 10.4230/LIPIcs.TQC.2020.10. URL https://drops.dagstuhl.de/opus/volltexte/2020/12069. 6

[10] Frank Arute, Kunal Arya, Ryan Babbush, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019. DOI: 10.1038/s41586-019-1666-5. URL https://doi.org/10.1038/s41586-019-1666-5. 1, 2, 3, 8

[11] Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald de Wolf. Quantum lower bounds by polynomials. *J. ACM*, 48(4):778–797, July 2001. ISSN 0004-5411. DOI: 10.1145/502090.502097. URL https://doi.org/10.1145/502090.502097. 3, 5, 18

[12] John Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, Nov 1964. DOI: 10.1103/PhysicsPhysiqueFizika.1.195. URL https://link.aps.org/doi/10.1103/PhysicsPhysiqueFizika.1.195. 2

[13] Aleksandrs Belovs. Variations on quantum adversary, 2015. 3

[14] Aleksandrs Belovs and Ansis Rosmanis. Tight quantum lower bound for approximate counting with quantum states, 2020. 4, 6

[15] Fernando G. S. L. Brandão, Aram W. Harrow, and Michał Horodecki. Local random quantum circuits are approximate polynomial-designs. *Communications in Mathematical Physics*, 346(2):397–434, 2016. DOI: 10.1007/s00220-016-2706-8. URL https://doi.org/10.1007/s00220-016-2706-8. 4

[16] Gilles Brassard, Peter Høyer, and Alain Tapp. Quantum cryptanalysis of hash and claw-free functions. *SIGACT News*, 28(2):14–19, June 1997. ISSN 0163-5700. DOI: 10.1145/261342.261346. URL https://doi.org/10.1145/261342.261346. 5, 14

[17] Gilles Brassard, Peter Høyer, Michele Mosca, and Alain Tapp. Quantum amplitude amplification and estimation. In *Quantum Computation and Quantum Information*, volume 305 of *Contemporary Mathematics*, pages 53–74. American Mathematical Society, 2002. ISBN 9780821821404. DOI: 10.1090/conm/305. 6, 14

[18] Mark Bun and Justin Thaler. Dual lower bounds for approximate degree and Markov-Bernstein inequalities. *Inf. Comput.*, 243(C):2–25, August 2015. ISSN 0890-5401. DOI: 10.1016/j.ic.2014.12.003. URL https://doi.org/10.1016/j.ic.2014.12.003. 24

[19] Boris Cirel'son (Tsirelson). Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics*, 4(2):93–100, 1980. DOI: 10.1007/BF00417500. URL https://doi.org/10.1007/BF00417500. 2

[20] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, Oct 1969. DOI: 10.1103/PhysRevLett.23.880. URL https://link.aps.org/doi/10.1103/PhysRevLett.23.880. 2

[21] Richard Cleve, Peter Høyer, Benjamin Toner, and John Watrous. Consequences and limits of nonlocal strategies. In *Proceedings of the 19th IEEE Annual Conference on Computational Complexity*, CCC '04, page 236–249, USA, 2004. IEEE Computer Society. ISBN 0769521207. DOI: 10.1109/CCC.2004.1313847. 2

[22] Aram Harrow and Saeed Mehraban. Approximate unitary t-designs by short random quantum circuits using nearest-neighbor and long-range gates, 2018. 4

[23] Norman L. Johnson and Samuel Kotz. *Urn models and their application: an approach to modern discrete probability theory*. Wiley, 1977. ISBN 9780471446309. 12

[24] Shelby Kimmel, Cedric Yen-Yu Lin, Guang Hao Low, Maris Ozols, and Theodore J. Yoder. Hamiltonian simulation with optimal sample complexity. *npj Quantum Information*, 3(1):13, 2017. DOI: 10.1038/s41534-017-0013-7. URL https://doi.org/10.1038/s41534-017-0013-7. 6

[25] William Kretschmer. The Quantum Supremacy Tsirelson Inequality. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, volume 185 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 13:1–13:13, Dagstuhl, Germany, 2021. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. ISBN 978-3-95977-177-1. DOI: 10.4230/LIPIcs.ITCS.2021.13. URL https://drops.dagstuhl.de/opus/volltexte/2021/13552. 1

[26] Troy Lee, Rajat Mittal, Ben W. Reichardt, Robert Špalek, and Mario Szegedy. Quantum query complexity of state conversion. In *Proceedings of the 2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, FOCS '11, page 344–353, USA, 2011. IEEE Computer Society. ISBN 9780769545714. DOI: 10.1109/FOCS.2011.75. URL https://doi.org/10.1109/FOCS.2011.75. 3

[27] Nathan Lindzey and Ansis Rosmanis. A Tight Lower Bound For Non-Coherent Index Erasure.

In Thomas Vidick, editor, *11th Innovations in Theoretical Computer Science Conference (ITCS 2020)*, volume 151 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 59:1–59:37, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. ISBN 978-3-95977-134-4. DOI: 10.4230/LIPIcs.ITCS.2020.59. URL https://drops.dagstuhl.de/opus/volltexte/2020/11744. 3

[28] Frederic Magniez, Ashwin Nayak, Jeremie Roland, and Miklos Santha. Search via quantum walk. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '07, page 575–584, New York, NY, USA, 2007. Association for Computing Machinery. ISBN 9781595936318. DOI: 10.1145/1250790.1250874. URL https://doi.org/10.1145/1250790.1250874. 6

[29] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, USA, 2014. ISBN 1107038324. DOI: 10.1017/CBO9781139814782. 23

[30] Martin Raab and Angelika Steger. "Balls into bins" - a simple and tight analysis. In *Proceedings of the Second International Workshop on Randomization and Approximation Techniques in Computer Science*, RANDOM '98, pages 159–170, Berlin, Heidelberg, 1998. Springer-Verlag. ISBN 354065142X. DOI: 10.1007/3-540-49543-6_13. 12

[31] Ben W. Reichardt. Reflections for quantum query algorithms. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '11, page 560–569, USA, 2011. Society for Industrial and Applied Mathematics. DOI: 10.1137/1.9781611973082.44. 6

[32] Alfréd Rényi. On the theory of order statistics. *Acta Mathematica Academiae Scientiarum Hungarica*, 4(3):191–231, 1953. DOI: 10.1007/BF02127580. URL https://doi.org/10.1007/BF02127580. 8